Jens Vygen

New approximation algorithms for the TSP

The traveling salesman problem (TSP) is probably the best known combinatorial optimization problem. Although studied intensively for sixty years, the TSP continues to pose grand challenges. Cook's [2012] recent book gives an excellent introduction.

Since the TSP is *NP*-hard (Karp [1972]), it is natural to ask for approximation algorithms. How good solutions can we guarantee to find in polynomial time? Christofides [1976] devised a $\frac{3}{2}$-approximation algorithm for the SYMMETRIC TSP: it always finds a solution that is at most 50% longer than optimum.

Can we do better? Can we do similarly well for the ASYMMETRIC TSP? These questions – still unsolved – belong to the most intriguing open problems in our field. Recently, there has been progress that makes us hope that we will learn more in the near future.

In this survey we try to describe the state of the art, in particular the recent progress, mostly from 2010–2012. We use standard notation and some basic terminology and well-known results from combinatorial optimization; see Korte and Vygen [2012] or Schrijver [2003] if necessary.

1 Introduction

Let us first review different formulations of the problems and basic approximation algorithms.

1.1 ASYMMETRIC TSP *and* SYMMETRIC TSP

The ASYMMETRIC TSP can be defined as follows. Given a finite set V (of *cities*) and distances $c(v,w) \geq 0$ for all $v,w \in V$ (also called *length* or *cost*), find a closed walk of minimum total length visiting each city at least once. More precisely, we look for a sequence $v_0, v_1, \ldots, v_k$ with $v_k = v_0$ and $\{v_0, \ldots, v_k\} = V$ (also called a *tour*) such that $\sum_{i=1}^{k} c(v_{i-1}, v_i)$ is minimum.

The SYMMETRIC TSP is the special case in which the distances are symmetric: $c(v,w) = c(w,v)$ for all $v,w \in V$.

Often these problems are formulated such that each city must be visited *exactly* once instead of at least once (except, of course, that we must end in the same city where we start). This is equivalent if the distances obey the *triangle inequality*

$$c(u,w) \;\leq\; c(u,v) + c(v,w) \quad \text{for all } u,v,w \in V \tag{1}$$

because then we can shortcut whenever we visit a city a second time.

On the other hand, given an instance of the ASYMMETRIC TSP (or SYMMETRIC TSP) as defined above, we can set

$$\bar{c}(v,w) \;:=\; \min\left\{\textstyle\sum_{e \in E(P)} c(e) \;:\; P \text{ path from } v \text{ to } w,\; V(P) \subseteq V\right\}$$

and consider $(V, \bar{c})$ instead. Note that $\bar{c}$ obeys (1). The instance $(V, \bar{c})$ is equivalent to (V, c) because we can move from v to w at cost $\bar{c}(v,w)$ via a shortest v-w-path. The pair $(V, \bar{c})$ is called the *metric closure* of (V, c).

1.2 Approximation algorithms

A *ρ-approximation algorithm* (for a minimization problem) is an algorithm that runs in polynomial time and always computes a solution (here: a tour) that costs at most ρ times the optimum. Here ρ can be a constant or a function of n; here and henceforth $n = |V|$ denotes the number of cities.

If we do not require the triangle inequality but still want to visit every city exactly once, the problems look hopeless: any approximation algorithm would allow us to decide in polynomial time whether a given graph contains a Hamiltonian circuit, and thus imply *P* = *NP* (this easy observation is due to Sahni and Gonzalez [1976]).

So we assume henceforth that we may visit cities more than once.

1.3 Euler's theorem

For the total length of a tour, all that matters is how many times we move from v to w for each ordered pair (v, w), or how many times we move between v and w for each unordered pair $\{v, w\}$ in the undirected case. Hence we can represent a tour by a directed or undirected graph (possibly with parallel edges) with vertex set V. As observed by Euler [1736], this graph has two properties:

(a) it is connected;

and for every city:

(b′) the in-degree equals the out-degree in the directed case;
(b″) the degree is even in the undirected case.

Digraphs with the properties (a) and (b′) and undirected graphs with properties (a) and (b″) are called *Eulerian*. The conditions are equivalent to the existence of an *Eulerian walk*: a closed walk traversing every edge exactly once and every vertex at least once. Given an Eulerian graph or digraph, one can find an Eulerian walk in linear time (Hierholzer [1873]).

Therefore, one can reformulate the ASYMMETRIC TSP and the SYMMETRIC TSP by asking for a (multi)set F such that (V, F) is an Eulerian (directed or undirected, respectively) graph with minimum total length $c(F)$. We call such an F a *tour*, too. Here and in the following we abbreviate $c(F) := \sum_{e \in F} c(e)$ and $c(e) := c(v, w)$ for any edge e from v to w.

1.4 GRAPHIC TSP

A natural special case of the SYMMETRIC TSP arises when we are given a connected undirected graph G and let $V = V(G)$ and $c(v, w) = 1$ if $\{v, w\} \in E(G)$ and $c(v, w) = \infty$ otherwise. This problem is called the GRAPHIC TSP. The metric closure $(V, \bar{c})$ of (V, c) is also called the metric closure of G. Functions $\bar{c}$ arising in this way are called *graphic metrics*.

By the observations in Section 1.3, the GRAPHIC TSP can be reformulated as follows. Given a connected graph G, find an Eulerian spanning multi-subgraph (V, F) with minimum $|F|$. Here a multi-subgraph arises from a subgraph by doubling a subset of its edges. Again, we call an Eulerian multi-subgraph of G simply a *tour in G*.

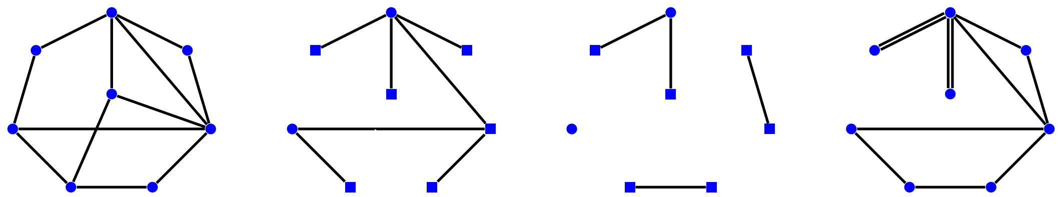
The GRAPHIC TSP has also been called GRAPH TSP by some authors. It is easy to see that, without loss of generality, we may assume that G is 2-vertex-connected (otherwise find a tour in each block separately).

1.5 Double tree and Christofides' algorithm

A 2-approximation algorithm for the SYMMETRIC TSP is easy: take a tree on vertex set V with minimum total edge length (it is well-known that such a *minimum spanning tree* can be found efficiently, e.g., by the greedy algorithm), and double all its edges. Since any tour is connected and thus contains a spanning tree, a minimum spanning tree cannot be longer than an optimum tour. Hence we have a 2-approximation algorithm.

Christofides [1976] showed how to improve this. His algorithm also begins by computing a minimum spanning tree (V, S). But then, to correct the parities, it adds a minimum weight T_S-join, where $T_S = \{v \in V : |\delta_S(v)| \text{ odd}\}$ is the set of odd-degree vertices of (V, S). For a set $T \subseteq V$, a *T-join* is a subset F of edges such that $|\delta_F(v)|$ is odd for $v \in T$ and even for $v \in V \setminus T$. See Figure 1. Of course S itself is a T_S-join, but we can do better: since every tour contains two disjoint T_S-joins (color the edges of a tour red and blue, changing the color whenever visiting an element of T_S for the first time, and finally delete pairs of parallel edges of the same color), the minimum weight of a T_S-join is at most half the length of an optimum tour.

Figure 1. *Christofides' algorithm. From left to right: an instance of the* GRAPHIC TSP, *a spanning tree (V, S) whose odd-degree vertices (elements of T_S) are shown as squares, a minimum T_S-join, and the resulting tour.*



So we have a $\frac{3}{2}$-approximation algorithm for the SYMMETRIC TSP. Its running time is $O(n^3)$, dominated by the subroutine to find a minimum weight T_S-join.

This bound on the approximation ratio of Christofides' algorithm is tight even for the GRAPHIC TSP: for a Hamiltonian graph (so a tour of length n exists) that contains a spanning tree all whose vertices have odd degree, if we take such a spanning tree, we end up with $\frac{3}{2}n - 1$ edges.

2 Relaxations

For *NP*-hard problems it is often useful to study relaxations that are easier to solve. For the TSP, there are several interesting relaxations.

As explained in Section 1.3, it is often useful to view a tour as a (multi)set F of edges. Now we associate a vector $x \in \mathbb{Z}^E_{\geq 0}$ with each tour, where E is the set of ordered or (in the symmetric case) unordered pairs of elements of V and x_e is the number of copies of e in F. Then the tour has length $c(F) = c(x) := \sum_{e \in E} c(e) x_e$. Given a vector $x \in \mathbb{R}^E_{\geq 0}$, the graph $(V, \{e \in E : x_e > 0\})$ is called the *support graph* of x. For any subset $E' \subseteq E$ we will write $x(E') := \sum_{e \in E'} x_e$.

2.1 Subtour LP

Let (V, c) be an instance of the SYMMETRIC TSP, $n = |V| \geq 3$, and $E = \binom{V}{2}$. The following LP, first formulated by Dantzig, Fulkerson and Johnson [1954], has often been called *subtour elimination LP* or simply *subtour LP* or *Held-Karp relaxation*:

$$\begin{array}{llcll} \min & c(x) & & & \\ \text{subject to} & x(\delta(U)) & \geq & 2 & (\emptyset \neq U \subset V) \\ & x(\delta(v)) & = & 2 & (v \in V) \\ & x_e & \leq & 1 & (e \in E) \\ & x_e & \geq & 0 & (e \in E) \end{array} \tag{2}$$

Note that the constraints $x_e \leq 1$ $(e \in E)$ could be omitted as they are implied by the other constraints: for $e = \{u, v\}$ we have $2x_e = x(\delta(u)) + x(\delta(v)) - x(\delta(\{u, v\})) \leq 2 + 2 - 2$.

The set of feasible solutions of the subtour LP (2) is called the *subtour polytope*. The integral feasible solutions of (2) are exactly the incidence vectors of Hamiltonian circuits. Their convex hull is called the *TSP polytope*.

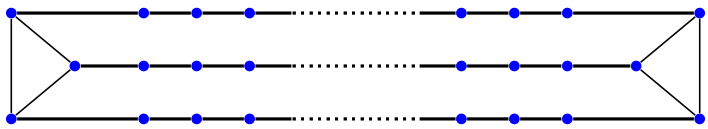
So (2) is a relaxation of the SYMMETRIC TSP if the triangle inequality holds. For a general instance, we can consider (2) for its metric closure.

This relaxation has been tightened by many classes of additional valid inequalities. It is also the basis of branch-and-cut algorithms (with exponential worst-case running time) that made impressive progress over the last four decades and have found optimum solutions to TSP instances with up to 85 900 cities; see Applegate et al. [2006].

2.2 Integrality ratio

Obviously, the integer solutions of (2) are exactly the incidence vectors of Hamiltonian circuits. However, this LP has no integral optimum solution in general. Figure 2 shows a well-known example. We have an infinite family of graphs G; each is an instance of the GRAPHIC TSP. The subtour LP of the metric closure (V, c) has a unique optimum solution: $x_e = 1$ on the horizontal edges and $x_e = \frac{1}{2}$ on the six other edges. Its LP value is n. However, an optimum tour has length $\frac{4}{3}n - 2$.

Figure 2. *Examples showing a lower bound of $\frac{4}{3}$ on the integrality ratio of the subtour polytope*



The *integrality ratio* of a family of polytopes $(P \subseteq \mathbb{R}^{E_P})_{P \in \mathcal{P}}$ is the supremum of $\min\{c(x) : x \in P \cap \mathbb{Z}^{E_P}\} / \min\{c(x) : x \in P\}$ over all $P \in \mathcal{P}$ and all $c : E_P \to \mathbb{R}_{>0}$. Often the weight functions are restricted in the supremum, e.g. to metrics or to graphic metrics. The above family of examples shows that the integrality ratio of the family of subtour polytopes for graphic metrics (and hence for general metrics) is at least $\frac{4}{3}$. Worse examples are not known. The fact that the worst known examples are instances of the GRAPHIC TSP raised interest in this special case. See also Sections 2.5 and 7.2.

2.3 Spanning trees

The difficulty of the SYMMETRIC TSP lies in the combination of connectivity and parity requirements. If we require only connectivity, a minimum spanning tree does the job. Edmonds [1970] gave the following polyhedral description:

Proposition 1. *The convex hull of incidence vectors of trees with vertex set V and edges in E is the set of vectors $x \in \mathbb{R}^E$ with*

$$\begin{array}{rcll} x(E) & = & n-1 & \\ \sum_{e=\{v,w\}\in E: v,w\in U} x_e & \leq & |U|-1 & (\emptyset \neq U \subset V) \\ x_e & \geq & 0 & (e \in E) \end{array} \tag{3}$$

This set is called the *spanning tree polytope* of the graph (V, E). The following easy observation was made by Asadpour et al. [2010], strengthening a result of Held and Karp [1970]:

Proposition 2. *If x is a feasible solution of (2), then $\frac{n-1}{n}x$ is in the relative interior of the spanning tree polytope of the support graph.*

Proof. We have $\frac{n-1}{n}x(E) = \frac{n-1}{2n}\sum_{v\in V} x(\delta(v)) = n-1$ as well as $\sum_{e=\{v,w\}\in E:v,w\in U} \frac{n-1}{n}x_e = \frac{n-1}{2n}(\sum_{v\in U} x(\delta(v)) - x(\delta(U))) = \frac{n-1}{2n}(2|U| - x(\delta(U))) \leq \frac{n-1}{n}(|U|-1)$ for any $\emptyset \neq U \subset V$. $\square$

2.4 T-joins

Now consider the parity aspect. Edmonds and Johnson [1973] proved:

Proposition 3. *The minimum weight of a T-join in a graph (V, E) with weights $c \in \mathbb{R}^E_{\geq 0}$ and $T \subseteq V$ equals the optimum value of the LP:*

$$\begin{array}{lrcll} \min & c(x) & & & \\ \text{subject to} & x(\delta(U)) & \geq & 1 & (U \subseteq V, |U \cap T| \text{ odd}) \\ & x_e & \geq & 0 & (e \in E) \end{array} \tag{4}$$

The cuts $\delta(U)$ with $|U \cap T|$ odd are called *T-cuts*.

For negative weights the LP (4) cannot be used directly. We will also need:

Proposition 4. *The convex hull of incidence vectors of T-joins in (V, E) is the set of vectors $x \in [0, 1]^E$ with*

$$|F| - x(F) + x(\delta(U) \setminus F) \;\; \geq \;\; 1 \quad \begin{array}{l}(U \subseteq V,\ F \subseteq \delta(U), \\ \;|U \cap T| + |F| \text{ odd})\end{array} \tag{5}$$

This is called the *T-join polytope* of (V, E). A minimum weight T-join can be found in $O(n^3)$ time via weighted matching. See Schrijver [2003] or Korte and Vygen [2012] for details and proofs of Propositions 1, 3, and 4.

2.5 Wolsey's analysis

Wolsey [1980] proved that Christofides' algorithm computes a tour of length at most $\frac{3}{2}L$, where L is the LP value of (2). In fact, this is easy to see from the above: By Proposition 2, the minimum weight of a spanning tree is at most L. By Proposition 3, the minimum weight of a T-join is at most $\frac{L}{2}$ for any $T \subseteq V$ with $|V|$ even.

This shows that the integrality ratio of (2) is at most $\frac{3}{2}$. No better upper bound is known in general.

2.6 Two-edge-connected spanning subgraphs

Every tour is 2-edge-connected, so a relaxation of the SYMMETRIC TSP is to find a minimum weight 2-edge-connected spanning subgraph (if the triangle inequality holds) or multi-subgraph. Unfortunately, these problems are also *NP*-hard (cf. Section 7.6).

Let $G = (V, E)$ be a 2-edge-connected undirected graph. Then the incidence vectors of the 2-edge-connected spanning subgraphs (2ECSS) of G are the integral feasible solutions of the following LP:

$$\begin{array}{lrcll} \min\ c(x) & & & & \\ \text{subject to} & x(\delta(U)) & \geq & 2 & (\emptyset \neq U \subset V) \\ & x_e & \leq & 1 & (e \in E) \\ & x_e & \geq & 0 & (e \in E) \end{array} \tag{6}$$

This LP arises from (2) by omitting the equality constraints. If we allow using edges twice, the LP (6) can be simplified further by omitting the upper bounds:

$$\begin{array}{lrcll} \min\ c(x) & & & & \\ \text{subject to} & x(\delta(U)) & \geq & 2 & (\emptyset \neq U \subset V) \\ & x_e & \geq & 0 & (e \in E) \end{array} \tag{7}$$

Cunningham (see Monma, Munson and Pulleyblank [1990]) and Goemans and Bertsimas [1993] observed:

Proposition 5. *If (V, E) is a complete graph, $|V| \geq 3$, and c obeys the triangle inequality, then the optimum values of (2), (6), and (7) are the same.*

Proof. Let x be a rational feasible solution of (7). Choose $k \in \mathbb{N}$ such that kx_e is an even integer for each $e \in E$. If there is a vertex $v \in V$ with $x(\delta(v)) > 2$, choose incident edges $e = \{v, w\}$ and $e' = \{v, w'\}$ with $x_e > 0$ and $x_{e'} > 0$, reduce x_e and $x_{e'}$ each by $\frac{1}{k}$ and increase $x_{\{w,w'\}}$ by $\frac{1}{k}$ while maintaining feasibility (the existence of two such edges e, e' follows from applying Lovász' [1976] splitting theorem to the Eulerian graph with kx_e copies of each edge e). Note that we maintain the property that $kx(\delta(v))$ is an even integer for all $v \in V$, so we end up with a feasible solution of (7) also satisfying $x(\delta(v)) = 2$ for all $v \in V$. Then also $x_e = \frac{1}{2}(x(\delta(v)) + x(\delta(w)) - x(\delta(\{v, w\}))) = \frac{1}{2}(2 + 2 - x(\delta(\{v, w\}))) \leq 1$ for all $e = \{v, w\} \in E$, so we have a feasible solution of (2). Due to the triangle inequality we never increased $c(x)$. $\square$

We also note:

Proposition 6. *If (V, E) is a 2-edge-connected graph and $c(e) = 1$ for all $e \in E$, then the optimum values of (6) and (7) are the same.*

Proof. Let x be an optimum solution of (7). Let $f = \{v, w\} \in E$ with $x_f > 1$. Call two vertices a and b close if $x(\delta(\{a\} \cup S) \setminus \{f\}) \geq 1$ for all $S \subseteq V \setminus \{a, b\}$. This is a transitive relation. If v and w are close, then we can reduce x_f to 1 and maintain feasibility. Otherwise each vertex is either close to v or close to w; so there is an edge $f' = \{v', w'\}$ such that v and v' are close, w and w' are close, and $x_{f'} < 1$. Then increasing $x_{f'}$ to $\min\{1, x_{f'} + x_f - 1\}$ and decreasing x_f to 1 maintains feasibility. $\square$

The same holds for integral solutions: we never need to take two copies of any edge, except of course for bridges of G.

The constraints of (7) define facets of the *graphical traveling salesman polyhedron*: the convex hull of vectors $x \in \mathbb{Z}^E_{\geq 0}$ for which $x(\delta(U)) \in \{2, 4, 6, \dots\}$ for all $\emptyset \neq U \subset V$. This was studied by Cornuéjols, Fonlupt and Naddef [1985].

2.7 Asymmetric subtour LP

Let (V, c) be an instance of the $\textsc{Asymmetric TSP}$ with (1) and $E = \{(v, w) : v, w \in V, v \neq w\}$. The following is the natural analogon of the subtour LP in this case:

$$
\begin{array}{lrcrcl}
\min\ c(y) & & & & & \\
\text{subject to} & & & y(\delta^+(U)) & \geq & 1 \quad (\emptyset \neq U \subset V) \\
 & y(\delta^+(v)) & = & y(\delta^-(v)) & = & 1 \quad (v \in V) \\
 & & & y_e & \geq & 0 \quad (e \in E)
\end{array}
\tag{8}
$$

Again, the integral feasible solutions to this LP are exactly the Hamiltonian circuits. Vectors $y \in \mathbb{R}^E_{\geq 0}$ with $y(\delta^+(v)) = y(\delta^-(v))$ for all $v \in V$ are called *circulations* in (V, E).

From a feasible solution y to (8) one can obtain a feasible solution to (2) by setting $x_{\{v,w\}} := y_{(v,w)} + y_{(w,v)}$ for all $\{v, w\} \in \binom{V}{2}$.

2.8 Solving the linear programs

All linear programs above have exponentially many constraints, but they can all be solved in polynomial time; in fact an optimum basic solution can be found in polynomial time. One way to show this is via the equivalence of optimization and separation. The LPs for spanning trees and T-joins can be solved by combinatorial algorithms. For the LPs (2), (6), (7), and (8), there are straightforward polynomial-size extended formulations (by introducing flow variables and using the max-flow min-cut theorem), but combinatorial algorithms to solve these LPs are not known. Held and Karp [1970], however, showed how to solve (2) fast approximately.

2.9 Optimum basic solutions

Any optimum basic solution x^* of any of the LPs (2), (6), and (7) has at most $2n - 3$ nonzero variables; in fact the subgraph of the support graph induced by U has at most $2|U| - 3$ edges for any $U \subseteq V$ with $|U| \geq 2$ (cf. Cornuéjols, Fonlupt and Naddef [1985] and Goemans [2006]).

Any optimum basic solution x^* of the LP (8) has at most $3n - 4$ nonzero variables (and the subgraph of the support graph induced by U has at most $3|U| - 4$ edges for any $U \subseteq V$ with $|U| \geq 2$); this was also shown by Goemans [2006].

The basic feasible solutions of (8) arise as the unique solutions of a linear equation system with all coefficients 0 or 1, so by Cramer's rule each of their components can be written as $\frac{a}{b}$ for integers a and $b \leq (3n-4)!/2$. The same holds for (2), (6), and (7), even with $b \leq (2n-3)!/2$.

2.10 Covering the cities by disjoint circuits

Another relaxation works both in the directed and undirected case. We ignore connectivity and look for a graph in which each city belongs to a circuit and the circuits are pairwise vertex-disjoint.

In the undirected case, such an edge set is called a *perfect 2-matching* because every vertex must have degree 2. A minimum weight perfect 2-matching can be found in polynomial time (this is essentially equivalent to nonbipartite weighted matching), but it did not prove useful in the design of approximation algorithms so far.

In the directed case, the analogous relaxation is even easier to solve. Here every vertex must have in-degree and out-degree 1. A minimum weight spanning subgraph with this property can easily be found by solving a bipartite weighted matching problem. This was used for the first nontrivial approximation algorithm for the ASYMMETRIC TSP, to be described next.

2.11 $O(\log n)$-approximation for ASYMMETRIC TSP

For the ASYMMETRIC TSP no constant-factor approximation algorithm is known, so we will consider $f(n)$-approximation algorithms, where $f(n)$ is a function of the number n of cities. It is trivial to give an n-approximation algorithm: order the cities arbitrarily, say $V = \{v_1, \ldots, v_n\}$, and take a shortest v_n-v_1-path and a shortest v_{i-1}-v_i-path for $i = 2, \ldots, n$.

The first nontrivial approximation algorithm was found by Frieze, Galbiati and Maffioli [1982]. It assumes that the triangle inequality holds and works as follows.

Begin with $W := V$. Find a minimum weight subset F of edges with $|\delta_F^+(v)| = |\delta_F^-(v)| = 1$ for all $v \in W$ and $|\delta_F^+(v)| = |\delta_F^-(v)| = 0$ for all $v \in V \setminus W$ (cf. Section 2.10). Then pick one vertex from each connected component of (W, F), and replace W by the set of these vertices. Iterate until (W, F) is connected.

The set of all edges chosen in this algorithm forms an Eulerian graph. Due to the triangle inequality, the total cost of edges picked in each iteration is at most the length of an optimum tour. Since $|W|$ decreases by a factor of two in each iteration, we are done after $\lfloor \log_2 n \rfloor$ iterations. Hence we have a $(\log_2 n)$-approximation algorithm.

Bläser [2003], Kaplan et al. [2005], and Feige and Singh [2007] improved this by a constant factor.

3 Random sampling

It is quite natural to first take a spanning tree to guarantee connectivity, and then add a minimum cost set of edges in order to make the graph Eulerian. This idea (underlying Christofides' algorithm) works also in the directed case as we shall see soon (in the proof of Theorem 8).

A minimum spanning tree does often not give the best overall result. A certain kind of random sampling led to better approximation algorithms for the ASYMMETRIC TSP and the GRAPHIC TSP.

3.1 Thin trees

Asadpour, Goemans, Mądry, Oveis Gharan and Saberi [2010] were the first to obtain an $o(\log n)$-approximation algorithm for the ASYMMETRIC TSP. Their algorithm is randomized. The main ingredient is the following result, which, interestingly, applies to an undirected instance:

Theorem 7. *There is a randomized polynomial-time algorithm which, given a feasible solution x of the LP (2), computes a spanning tree (V, S) such that with probability at least $\frac{1}{2}$ we have $c(S) \leq 2c(x)$ and $|\delta_S(U)| \leq \alpha x(\delta(U))$ for all $U \subseteq V$, where $\alpha = 4 \log n / \log \log n$.*

The last property is called α-thinness. By Proposition 2, $\frac{n-1}{n}x$ is a convex combination of spanning trees, i.e. $\frac{n-1}{n}x_e = \sum_{S \in \mathcal{S}: e \in S} p_S$ for all $e \in E$, where $\mathcal{S}$ is the set of (edge sets of) spanning trees, $p_S \geq 0$ for all $S \in \mathcal{S}$ and $\sum_{S \in \mathcal{S}} p_S = 1$. Such an explicit convex combination can be obtained in polynomial time. If we pick each tree $S \in \mathcal{S}$ with probability p_S, the expected cost is $\sum_{e\in E} c(e)x_e$, and hence the cost at most is twice as much with probability at least $\frac{1}{2}$.

The difficulty is that such a random spanning tree will in general not be thin enough. Therefore Asadpour et al. [2010] choose the probability distribution carefully, namely such that it maximizes the entropy $\sum_{S\in\mathcal{S}} p_S \log \frac{1}{p_S}$. Equivalently, $p_S = \gamma \Pi_{e\in S} \lambda_e$ for all $S \in \mathcal{S}$, for suitable positive numbers γ and λ_e ($e \in E$). Such a distribution is also called λ-*uniform*.

Asadpour et al. [2010] show how to sample trees efficiently from approximately this distribution. Moreover, they prove that the random variables indicating for each edge whether it is part of the selected tree are negatively correlated; then thinness is implied by a Chernoff bound together with the fact that in any graph there are less than $2n^{2\gamma}$ many γ-approximate minimum cuts, for any $\gamma \geq 1$ (Karger and Stein [1996]). See Asadpour et al. [2010] for the details. (Alternatively, a thin tree can be obtained by the dependent randomized rounding approach of Chekuri, Vondrák and Zenklusen [2010].)

3.2 The $O(\log n/\log\log n)$-approximation algorithm

Using Theorem 7, the randomized $O(\log n/\log\log n)$-approximation algorithm of Asadpour et al. [2010] and its analysis can be described easily. We work in the metric closure, so c satisfies the triangle inequality.

First solve the LP relaxation (8) to obtain a vector y. Then get a solution x of (2) by setting $x_{\{v,w\}} := y_{(v,w)} + y_{(w,v)}$ for all $\{v,w\} \in \binom{V}{2}$. Note that $x_e = 0$ or $x_e \geq \frac{1}{(3n-4)!}$ for all $e \in \binom{V}{2}$ (cf. Section 2.9); so x can be stored with $O(n^2 \log n)$ bits.

Next apply Theorem 7 to obtain a spanning tree (V, S). Orient the edges of this tree by replacing each $\{v,w\} \in S$ by the cheaper one of (v,w) and (w,v). Setting $c'(\{v,w\}) := \min\{c(v,w), c(w,v)\}$, we get with probability at least $\frac{1}{2}$ that the resulting arc set R satisfies $c(R) = c'(S) \leq 2c'(x) \leq 2c(y)$ as well as $|\delta_R^-(U)| \leq |\delta_S(U)| \leq \alpha x(\delta(U)) = \alpha(y(\delta^-(U)) + y(\delta^+(U))) = 2\alpha y(\delta^+(U))$ for all $U \subseteq V$.

Finally apply the following Theorem 8 to R and y. With probability at least $\frac{1}{2}$ we obtain a tour of length at most $(2\alpha + 2)c(y)$.

Theorem 8. *Let (V, R) be a connected spanning subgraph of the complete digraph (V, E), $y \in \mathbb{R}^E_{\geq 0}$, and $\beta > 0$ such that $|\delta_R^-(U)| \leq \beta y(\delta^+(U))$ for all $U \subseteq V(G)$. Then we can find a tour F with $c(F) \leq c(R) + \beta c(y)$ in polynomial time.*

Proof. Let $l(e) := 1$ for $e \in R$ and $l(e) := 0$ for $e \in E \setminus R$. Any integral circulation f in (V, E) with $f \geq l$ corresponds to a tour. We compute an integral minimum cost circulation $f^* \geq l$ and note that the resulting tour has cost $c(f^*)$.

To prove that such a circulation (and hence an integral circulation) of cost at most $c(R) + \beta c(y)$ exists, we let $u(e) := \max\{l(e), \beta y_e\}$ for all $e \in E$ and observe that a circulation g with $l \leq g \leq u$ exists; then $c(f^*) \leq c(g) \leq \sum_{e\in E} c(e)u(e) \leq c(R) + \beta c(y)$.

The existence of g follows from Hoffman's [1960] circulation theorem: we have $l \leq u$ and $l(\delta^-(U)) = |\delta_R^-(U)| \leq \beta y(\delta^+(U)) \leq u(\delta^+(U))$ for all $U \subseteq V$. □

3.3 Random sampling for the Symmetric TSP

The random sampling of Asadpour et al. [2010] was also used by Oveis Gharan, Saberi and Singh [2011] for the first improvement over Christofides' algorithm for the Graphic TSP.

They proposed the following algorithm. Take the metric closure and solve the subtour LP (2) to obtain an optimum basic solution x. Again, $\frac{n-1}{n}x$ is a convex combination of spanning trees, and we pick one at random according to a maximum entropy distribution; call it (V, S). Let T_S be again the set of odd degree vertices of (V, S). Finally add a minimum-weight T_S-join to S as in Christofides' algorithm.

Oveis Gharan, Saberi and Singh [2011] conjectured that this algorithm has better approximation ratio than $\frac{3}{2}$, but they could prove this only for the GRAPHIC TSP, and only for a slight variant of this algorithm. Their main structure theorem is the following. (The constants below are not best possible, but the improvement is tiny anyway.)

Theorem 9. *Let (V, c) be an instance of the* SYMMETRIC TSP *satisfying the triangle inequality. Let x be an optimum solution of (2), and let (V, S) be a spanning tree picked at random according to the maximum entropy distribution $(p_S)_{S\in\mathcal{S}}$ with $\sum_{S\in\mathcal{S}:e\in S} p_S = \frac{n-1}{n}x_e$ for all $e \in E$. Let T_S be the set of odd degree vertices of (V, S). Call an edge e* good *if e does not belong to any T_S-cut $\delta(U)$ with $x(\delta(U)) \leq 2 + 10^{-15}$. Then at least one of the following holds:*

(a) there is a subset E^* of edges with $x(E^*) \geq 10^{-12}n$ such that for each $e \in E^*$ the probability that e is good *is at least 10^{-24};*

(b) there are at least $\frac{19}{20}n$ edges e with $x_e \geq 1 - 10^{-7}$.

The proof of this theorem is very long. It uses deeper results about random spanning trees and the structure of near-minimum cuts.

3.4 First improvement over Christofides for the GRAPHIC TSP

Following Oveis Gharan, Saberi and Singh [2011], we show now that Theorem 9 implies a better approximation ratio for the GRAPHIC TSP.

In case (a), Wolsey's analysis can be improved: let $y_e := x_e/(2 + 10^{-15})$ for good edges e and $y_e := x_e/2$ for other edges e. Since $y(\delta(U)) \geq 1$ for every T_S-cut $\delta(U)$, we conclude that y is a feasible solution to (4). Therefore the expected cost of a minimum weight T_S-join is at most $c(y) \leq \frac{1}{2}c(x) - 10^{-16}\sum_{e\in E:e\,\text{good}} c(e)x_e \leq \frac{1}{2}c(x) - 10^{-40}\sum_{e\in E^*} c(e)x_e$. If c is a graphic metric, we have $c(e) \geq 1$ for all $e \in E$ and $c(x) \leq 2n$. Then we get $c(y) \leq \frac{1}{2}c(x) - 10^{-40}x(E^*) \leq \frac{1}{2}c(x) - 10^{-52}n \leq \frac{1}{2}(1 - 10^{-52})c(x)$.

In case (b), the approximation ratio is better, and the proof is also easy. Let I be the set of edges e with $x_e \geq 1 - 10^{-7}$. The edges in I form vertex-disjoint paths and circuits, and each circuit has length at least 10^7 (or is Hamiltonian). Remove one edge from each circuit and add edges of cost 1 to obtain a spanning tree (V, S). Note that $c(S\setminus I) = |S\setminus I| < (\frac{1}{20}+10^{-7})n \leq (\frac{1}{20}+10^{-7})c(x)$. We get $c(S) = c(S\cap I)+c(S\setminus I) \leq \sum_{e\in S} c(e)x_e/(1-10^{-7})+(\frac{1}{20}+10^{-7})c(x)$.

Finally we add a minimum weight T_S-join J, where T_S is the set of vertices with odd degree in (V, S). To bound $c(J)$, let $y_e := \frac{1}{3}$ for $e \in S$ and $y_e := \frac{2}{3}x_e$ for other edges e. We show that y is a feasible solution to (4).

For any set U with $|U\cap T_S|$ odd we have $|\delta(U)\cap S|$ odd. If $|\delta(U)\cap S| = 1$, then $y(\delta(U)) \geq \frac{1}{3} + y(\delta(U)\setminus S) \geq \frac{1}{3} + \frac{2}{3}(x(\delta(U)) - 1) \geq 1$. If $|\delta(U)\cap S| \geq 3$, then $y(\delta(U)) \geq 3\cdot\frac{1}{3} = 1$.

Hence $c(J) \leq c(y) = \frac{1}{3}c(S) + \frac{2}{3}\sum_{e\in E\setminus S} c(e)x_e$. We conclude $c(S \,\dot\cup\, J) \leq \frac{4}{3}c(S) + \frac{2}{3}\sum_{e\in E\setminus S} c(e)x_e \leq \frac{4}{3}c(x)/(1 - 10^{-7}) + \frac{4}{3}(\frac{1}{20} + 10^{-7})c(x) \leq (\frac{7}{5} + 10^{-6})c(x)$.

Note that we used properties of the GRAPHIC TSP in both cases, (a) and (b). Although the improvement over Christofides' algorithm is tiny (in case (a)), this result received a lot of interest.

4 Correcting parity by adding and removing edges

So far, all algorithms began with a spanning tree and then added edges to make the graph Eulerian. Mömke and Svensson [2011] had a brilliant idea: if we begin with a 2-connected graph, we may also delete some edges for making it Eulerian, and this may be cheaper overall.

4.1 Removable pairings

The following definition of Mömke and Svensson [2011] is very interesting. A *removable pairing* in a 2-vertex-connected graph (V, E) is a pair $(R, \mathcal{P})$ with the following properties:

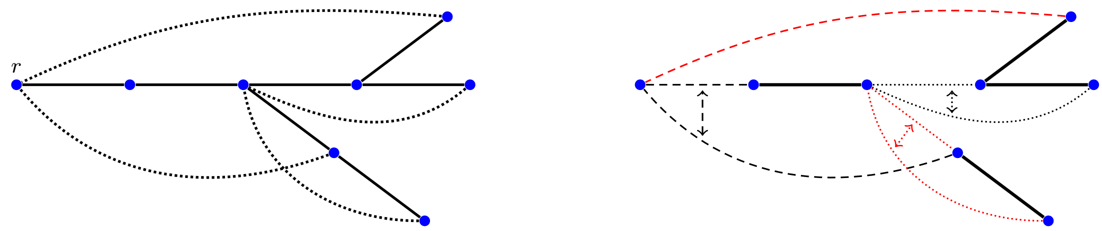
(a) $R \subseteq E$;
(b) for each $P \in \mathcal{P}$ there exists a vertex $v \in V$ and three distinct edges e_1, e_2, e_3 incident to v such that $P = \{e_1, e_2\}$;
(c) the elements of $\mathcal{P}$ are pairwise disjoint;
(d) for any set $F \subseteq R$ with $|F \cap P| \leq 1$ for all $P \in \mathcal{P}$, the graph $(V, E \setminus F)$ is connected.

Mömke and Svensson [2011] proposed to obtain a removable pairing as follows.

Lemma 10. *Let $G = (V, E)$ be a 2-vertex-connected graph and (V, S) a DFS-tree in G, rooted at $r \in V$. For each edge $e = \{v, w\} \in E \setminus S$, let w.l.o.g. be v on the r-w-path in (V, S), and let v' be the successor of v on this path. Add e to R; moreover if $|\delta(v)| \geq 3$ and $e' = \{v, v'\}$ has not yet been added to R, then add also e' to R and $\{e, e'\}$ to $\mathcal{P}$ (cf. Figure 3). Then $(R, \mathcal{P})$ is a removable pairing in G.*

Proof. (a)–(c) are easy to see. To show that condition (d) holds, take $F \subseteq R$ with $|F \cap P| \leq 1$ for all $P \in \mathcal{P}$. For each $v \in V$ we consider the set W_v of vertices w for which v is on the r-w-path in (V, S). We show that for each $v \in V$ the vertex set W_v induces a connected subgraph of $(V, E \setminus F)$. Indeed, this follows from a straightforward induction on $|W_v|$. □

Figure 3. *A 2-connected graph with a DFS tree (left, solid edges) and a removable pairing (right: dashed and dotted edges are in R; arrows indicate pairs).*



4.2 The Mömke-Svensson lemma

Now we can formulate and prove the key lemma of Mömke and Svensson [2011]. It works for general weights, although it has been used so far only for $c \equiv 1$. We follow the proof of Sebő and Vygen [2012], a variant of the original proof:

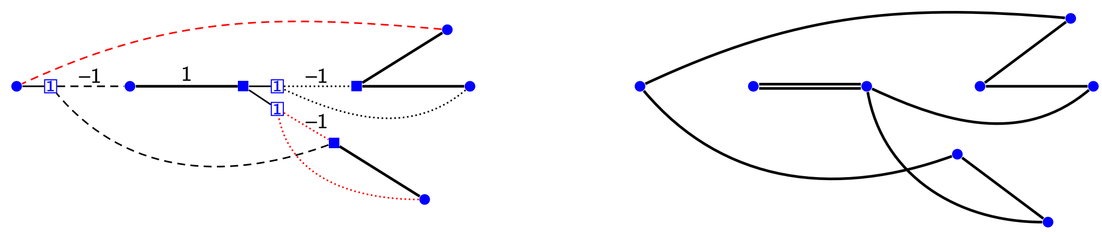
Theorem 11. *Let $G = (V, E)$ be a 2-vertex-connected graph, $c : E(G) \to \mathbb{R}$, and $(R, \mathcal{P})$ a removable pairing in G. Then one can find a tour in G of length at most $\frac{4}{3}c(E) - \frac{2}{3}c(R)$ in $O(n^3)$ time.*

Proof. Let T_G be the set of odd degree vertices of G. Let $c'(e) = c(e)$ for $e \in E \setminus R$ and $c'(e) = -c(e)$ for $e \in R$. For any T_G-join J in G that intersects each pair $P \in \mathcal{P}$ in at most one edge, we construct a tour from E by doubling the edges in $J \setminus R$ and deleting the edges in $J \cap R$. This tour has length $c(E) + c'(J)$.

To compute a T_G-join of weight at most $\frac{1}{3}c(E) - \frac{2}{3}c(R) = \frac{1}{3}c'(E)$, intersecting each pair at most once, we construct an auxiliary graph G' with weights c' from (G, c') as follows (cf. Figure 4). For each pair $P = \{\{v,w\},\{v,w'\}\} \in \mathcal{P}$ we add a vertex v_P and an edge $\{v, v_P\}$ of weight zero, and replace the two edges in P by $\{v_P, w\}$ and $\{v_P, w'\}$, keeping their weight.

Let $T_{G'}$ be the set of odd degree vertices of G'. G' is 2-edge-connected. Hence every $T_{G'}$-cut contains at least three edges, and the vector with all components $\frac{1}{3}$ is in the $T_{G'}$-join polytope of G' (cf. Proposition 4), and even in its face defined by $x(\delta(v_P)) = 1$ for all $P \in \mathcal{P}$. Hence there is a $T_{G'}$-join J' in G' with $|\delta_{J'}(v_P)| = 1$ for all $P \in \mathcal{P}$ and with weight at most $\frac{1}{3}c'(E)$. Such a J' can be found in $O(n^3)$ time (by adding a large weight to edges incident to v_P, for all $P \in \mathcal{P}$). It corresponds to a T_G-join J in G that intersects each pair at most once and has weight at most $\frac{1}{3}c'(E)$. $\square$

Figure 4. *Proof of Theorem 11. The graph G' on the left results from G and $(R, \mathcal{P})$ in Figure 3. Squares denote odd-degree vertices. Here $|E| = 11$ and $|R| = 7$. As J' one could choose, e.g., the four edges whose weight is shown. This leads to the tour shown on the right.*



4.3 Subcubic graphs

Boyd, Sitters, van der Ster and Stougie [2011] devised a $\frac{4}{3}$-approximation algorithm for cubic graphs. Mömke and Svensson [2011] gave a simpler proof for this result and extended it to subcubic graphs (i.e., graphs with maximum degree 3). Indeed, Lemma 10 yields a removable pairing with $|R| \geq 2(|E| - |S|) - 1$, because all non-tree edges, except possibly one incident to the root, can be paired with tree edges if the graph is subcubic. Theorem 11 yields a tour with at most $\frac{4}{3}|E| - \frac{2}{3}|R| = \frac{4}{3}n - \frac{2}{3}$ edges. This is best possible, e.g. for graphs that consist only of three internally vertex-disjoint paths of the same length and with the same endpoints.

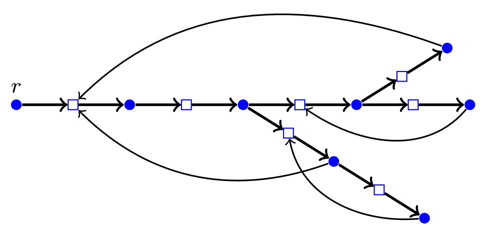
Correa, Larré and Soto [2012] refined the techniques of Boyd et al. [2011]; they can compute a tour of length less than $(\frac{4}{3} - \frac{1}{61236})n$ in any cubic graph in polynomial time.

4.4 Removable pairing via circulation

Mömke and Svensson [2011] showed how to find a good removable pairing in general graphs by a network flow approach, somewhat similar to Theorem 8. The idea again to start with a DFS tree and include some of the non-tree edges to make the subgraph 2-vertex-connected, but use as few as possible non-pairable edges.

First the input graph G is transformed into a flow network D as follows (cf. Figure 5). Let (V, S) be again a DFS tree, rooted at r. Note that r has degree 1 because G is 2-connected. Orient all tree edges away from r and all non-tree edges towards r. Subdivide each arc $e \in S$ by a vertex i_e. For each non-tree arc (v, w) add an arc (v, i_e), where e is the first edge on the w-v-path in (V, S).

Figure 5. *Flow network D in which we look for a circulation. Tree arcs (except the one incident to r) require at least one unit of flow. New vertices i_e ($e \in S$) are shown as squares.*



Let $l((v, i_e)) := 1$ for each $v \in V \setminus \{r\}$ and $e \in \delta^+_S(v)$, and $l(e) := 0$ for all other arcs in D. Let $c(f) := \sum_{e \in S} \max\{0, f(\delta^-(i_e)) - 1\}$. Mömke and Svensson [2011] proved:

Lemma 12. *Let f be an integral circulation in D with $f \geq l$. Then one can construct a tour with at most $\frac{4}{3}n + \frac{2}{3}c(f) - \frac{2}{3}$ edges in $O(n^3)$ time.*

Proof. Let B be the set of non-tree edges that correspond to edges in D with positive flow. $(V, S \cup B)$ is 2-vertex-connected. Let $C := \{e = (v, w) \in S : v \neq r,\ f(\delta^-(i_e) \setminus \{(v, i_e)\}) > 0\}$ be the set of tree edges that can be paired (with a non-tree edge). Define a removable pairing in G by $R := B \cup C$ and letting $\mathcal{P}$ contain a pair P for each element of C: for $e \in C$ choose an $e' \in B$ that corresponds to an edge in $\delta^-(i_e)$ and let $P = \{e, e'\}$. By Lemma 10, $(R, \mathcal{P})$ is indeed a removable pairing.

Now we apply Theorem 11 and obtain a tour with at most $\frac{4}{3}|S \cup B| - \frac{2}{3}|R| = \frac{4}{3}|S| + \frac{2}{3}|B| - \frac{2}{3}|C| = \frac{4}{3}(n-1) + \frac{2}{3}c(f) + \frac{2}{3}$ edges. $\square$

An integral circulation in D with $f \geq l$ and $c(f)$ minimum can be found in $O(n^3)$ time. To bound the cost, Mömke and Svensson [2011] (and then also Mucha [2012]) proceeded as follows.

1. Compute an optimum basic solution x of (6), with $c \equiv 1$ (in fact, Mömke and Svensson [2011] and Mucha [2012] used (7) instead, but using (6) simplifies the proof; cf. Proposition 6).
2. Compute a DFS tree (V, S) by choosing a root arbitrarily and following always an edge e with maximum x_e to an unvisited vertex.
3. Define the following fractional circulation f' in the associated flow network (D, l): For each $e \in E \setminus S$ send x_e units of flow along the fundamental cycle of e (the circuit in D corresponding to the unique circuit in $(V, S \cup \{e\})$).
4. For each $v \in V \setminus r$ and $e \in \delta^+(v)$ with $f'(v, i_e) < 1$, send $1 - f'(v, i_e)$ units of flow along any fundamental cycle containing e; this circulation is called f''. Let $f := f' + f''$. Then $f \geq l$.

Mömke and Svensson [2011] proved $c(f) \le (4\sqrt{2}-3)x(E) - (6\sqrt{2}-6)n$. Mucha [2012] improved the analysis and obtained $c(f) \le \frac{5}{3}x(E) - \frac{3}{2}n$. The heart of his proof consists of showing that for each $v \in V \setminus \{r\}$ the contribution of the edges in $B_v := \{e \in \delta^-(v) : x_e > 0\}$ to $c(f)$ plus the extra flow added for v in step 4 is at most $\frac{1}{6}|B_v| + \frac{5}{6}(x(\delta(v)) - 2)$; the result then follows from summation, using the fact that x has at most $2n-3$ nonzero variables (cf. Section 2.9).

We do not know whether Mucha's bound is tight. Together with Lemma 12 it directly yields a $\frac{13}{9}$-approximation algorithm for the $\textsc{Graphic TSP}$: we get a tour with at most $\frac{4}{3}n + \frac{10}{9}x(E) - n$ edges.

In the case that x in Step 1 is half-integral, we actually get a $\frac{4}{3}$-approximation (as observed by Mömke and Svensson [2011]): we may assume that the support graph is 2-connected (otherwise consider its blocks separately); then $f'' \equiv 0$ and $c(f) = 0$. This is particularly interesting because Schalekamp, Williamson and van Zuylen [2012] conjectured that the worst case for the integrality ratio occurs when x is an optimum fractional perfect simple 2-matching (and hence w.l.o.g. half-integral).

5 Using ear-decompositions and matroids

5.1 Ear-decompositions

An *ear-decomposition* of a connected graph $G = (V, E)$ is a sequence $P_0, P_1, \dots, P_k$ of subgraphs of G such that P_0 consists of a single vertex, $\{E(P_1), \dots, E(P_k)\}$ is a partition of E, and for $i = 1, \dots, k$, either P_i is a path with exactly its endpoints in $V(P_0) \cup \dots \cup V(P_{i-1})$ or P_i is a circuit with exactly one of its vertices (called its endpoint) in $V(P_0) \cup \dots \cup V(P_{i-1})$.

The vertices of an ear that are not endpoints are called its *internal vertices*. The length of an ear is the number of its edges; this is always the number of internal vertices plus one. An ear is called *trivial* if it has length 1, otherwise *nontrivial*. We call an ear *short* if it has length 2 or 3, otherwise *long*. An ear is called *odd* if its length is odd, otherwise *even*. The number of ears is always $|E| - |V| + 1$. See Figure 6, left-hand side, for an example.

Whitney [1932] observed that a graph is 2-edge-connected if and only if it has an ear-decomposition. Hence computing an ear-decomposition with minimum number of nontrivial ears is equivalent to finding the smallest 2-edge-connected spanning subgraph (2ECSS); this problem is *NP*-hard. However, the number of even ears can be minimized in polynomial time. This is a fundamental result of Frank [1993] (also proved in Schrijver's [2003] book):

Theorem 13. *Let $G = (V, E)$ be a 2-edge-connected graph. Let $\varphi(G)$ denote the minimum number of even ears in any ear-decomposition of G. Then for any $T \subseteq V$ such that $|T|$ is even, there exists a T-join in G with at most $\frac{1}{2}(|V| + \varphi(G) - 1)$ edges. Moreover, there exists a $T \subseteq V$ such that $|T|$ is even and the minimum cardinality of a T-join in G is $\frac{1}{2}(|V| + \varphi(G) - 1)$. Such a T and an ear-decomposition with $\varphi(G)$ even ears can be found in $O(|V||E|)$ time.*

Any 2-edge-connected spanning subgraph (2ECSS) of G has at least $\varphi(G)$ ears in any ear-decomposition. Hence any 2ECSS, and thus any tour, has at least $n - 1 + \varphi(G)$ edges. Cheriyan, Sebő and Szigeti [2001] used Theorem 13 to strengthen this statement. Let

$$\mathrm{LP}(G) \;:=\; \min\{x(E) : x \ge 0, x(\delta(U)) \ge 2 \; (\emptyset \ne U \subset V)\}. \tag{9}$$

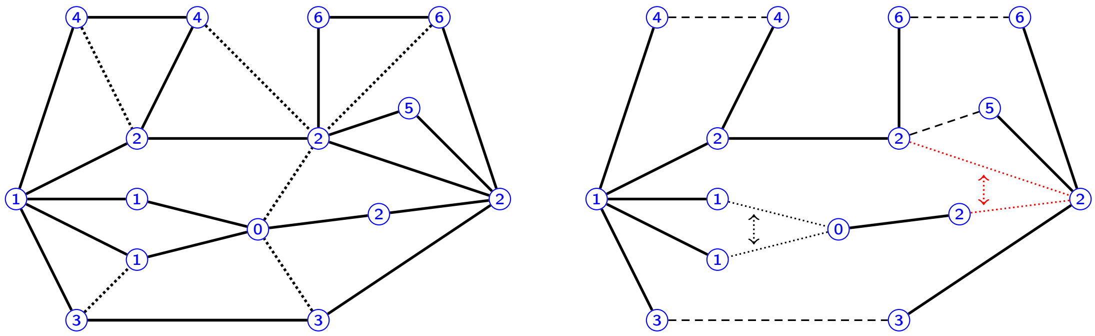
Note that (9) is the special case of (7) for $c(e) = 1$ for all $e \in E$, and $\mathrm{LP}(G)$ is a lower bound on the length of an optimum tour.

Corollary 14. *For any 2-edge-connected graph G we have*

$$L_\varphi \; := \; n-1+\varphi(G) \; \leq \; \mathrm{LP}(G).$$

Proof. By Theorem 13 there exists a set T of vertices such that $|T|$ is even and $\frac{1}{2}(n-1+\varphi(G))$ is the minimum cardinality of a T-join in G. Now observe that (4) is at most half of (7). $\square$

Figure 6. *Left: a graph G with an ear-decomposition. The internal vertices of the i-th ear are labelled i. We have $\varphi(G)=2$; ears 1 and 5 are even. Ears 3, 4, 5, and 6 are short; they are all pendant. Dotted edges are trivial ears. Right: trivial ears are deleted, and a removable pairing $(R,\mathcal{P})$ with $|R|=8$ and $|\mathcal{P}|=2$ as in the proof of Theorem 15 is shown.*



5.2 Applying the Mömke-Svensson lemma to ear-decompositions

Let us call an ear *pendant* if none of its internal vertices is endpoint of any nontrivial ear. By applying Theorem 11 to an ear-decomposition of a graph G, Sebő and Vygen [2012] observed:

Theorem 15. *Given a 2-vertex-connected graph G with an ear-decomposition with π pendant ears and no trivial ears, one can construct a tour with at most $\frac{4}{3}(n-1)+\frac{2}{3}\pi$ edges in $O(n^3)$ time.*

Proof. Define a removable pairing by taking an arbitrary edge of each pendant ear, and for each other ear a pair of its edges, incident to a common vertex that is endpoint of another ear. If $k=|E|-|V|+1$ denotes the number of ears, we have $|R|=2k-\pi$. Theorem 11 yields a tour with at most $\frac{4}{3}|E|-\frac{2}{3}|R|=\frac{4}{3}(n-1)+\frac{2}{3}\pi$ edges. $\square$

See Figure 6 for an example. This bound is good if there are few pendant ears. Otherwise we need something else. It turns out that long pendant ears are easy to deal with, but short pendant ears require care.

5.3 Nice and nicer ear-decompositions

Frank's Theorem 13 was also used by Cheriyan, Sebő and Szigeti [2001] and Sebő and Vygen [2012] as a starting point to obtain a nice ear-decomposition. An ear-decomposition is called *nice* if it has $\varphi(G)$ even ears, all short ears are pendant, and there is no edge joining internal vertices of different short ears. (Figure 6, left-hand side, displays a nice ear-decomposition.)

Lemma 16. *Given a 2-vertex-connected graph, one can compute a nice ear-decomposition in polynomial time.*

A nice ear-decomposition allows for optimizing the short ears in the following sense. Let M contain for each short ear the set of its internal vertices (cf. Figure 7, left-hand side). For $f \in M$ we denote by E_f the set of pairs $\{v, w\}$ such that G contains a path from v to w whose set of internal vertices is f. We will pick an $e_f \in E_f$ for each $f \in M$ such that $(V, \{e_f : f \in M\})$ has as few connected components as possible.

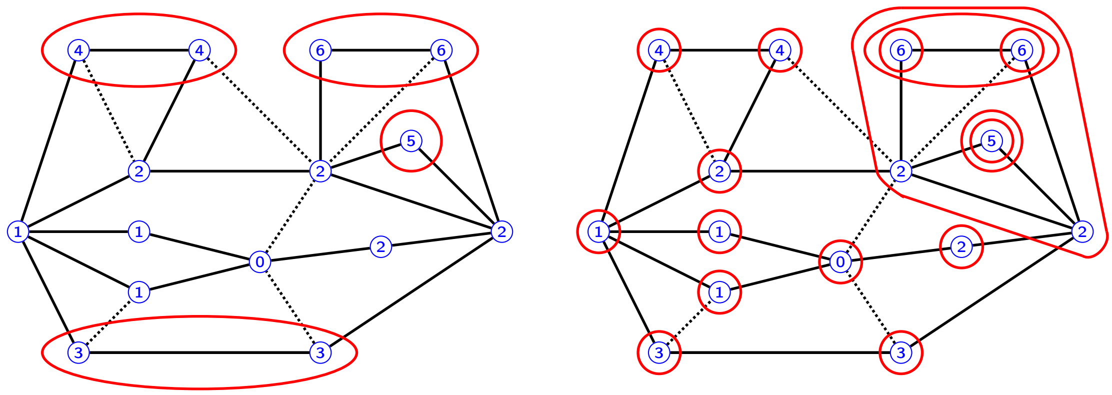
Equivalenty, we pick an $e_f \in E_f$ for each $f \in M$ such that the rank of $\{e_f : f \in M\}$ in the graphic matroid is maximum. Denote this maximum by μ. By Rado's [1942] Theorem,

$$\mu \;=\; \min\{r(\textstyle\bigcup_{i \in I} E_i) + |M \setminus I| : I \subseteq M\}. \tag{10}$$

(In the example of Figure 7, $|M| = 4$, $\mu = 3$ and $I = \{5, 6\}$ attains the minimum.) The maximum can be found by an algorithm for matroid intersection. Sebő and Vygen [2012] found an $O(|V||E|)$-time algorithm.

We then replace the short ear with internal vertices f by a path with internal vertices f and endpoints e_f, for each $f \in M$. This may change the set of trivial ears, but the ear-decomposition remains nice. Let V_M denote the union of the sets in M (i.e., the set of internal vertices of short ears). We have:

Figure 7. *Left: Optimizing the ear-decomposition of Figure 6. The elements of M are the red sets. Here only ear 4 needed to be replaced. Right: the cuts in the proof of Theorem 18.*



Theorem 17. *Let G be a 2-connected graph with a nice ear-decomposition. Then we can compute in polynomial time another nice ear-decomposition of G such that (V, F), where F contains all edges of short ears, has $|V| - |V_M| - \mu$ connected components.* □

We also get another lower bound:

Theorem 18. *Let G be a 2-connected graph with a nice ear-decomposition with $|M|$ short ears. Then*

$$L_\mu \;:=\; n - 1 + |M| - \mu \;\leq\; \mathrm{LP}(G).$$

Proof. Let I be a set attaining the minimum in (10). Let $\mathcal{U}$ be the partition of V such that for each $f \in I$ there is a $U \in \mathcal{U}$ with $f \subseteq U$ and $e_f \subseteq U$ for all $e_f \in E_f$, and $|\mathcal{U}| = |V| - |V_I| - r(\bigcup_{i \in I} E_i)$. By Rado's Theorem (cf. (10)) we have $|\mathcal{U}| = |V| - |V_I| - \mu + |M \setminus I|$.

Consider the family of sets $\mathcal{U} \cup I \mathbin{\dot\cup} \{\{v\} : v \in V_I\}$, taking singletons in I twice. Summing over the inequalities $x(\delta(U)) \geq 2$ for these sets U (unless $U = V$) completes the proof because no edge is contained in more than two of these at least $n - 1 + |M| - \mu$ cuts. $\square$

See Figure 7 (right-hand side) for an illustration. The set F in Theorem 17 consists of the black edges in Figure 8, left-hand side.

5.4 The $\frac{7}{5}$-approximation algorithm

Now we can explain the $\frac{7}{5}$-approximation algorithm for the Graphic TSP by Sebő and Vygen [2012]. Let $\Lambda^G := \frac{2}{3}L_\mu + \frac{1}{3}L_\varphi$. Note that Λ^G is a lower bound on the optimum and in fact on $\mathrm{LP}(G)$, and $\Lambda^G \geq n - 1$ (cf. Corollary 14 and Theorem 18). We first show:

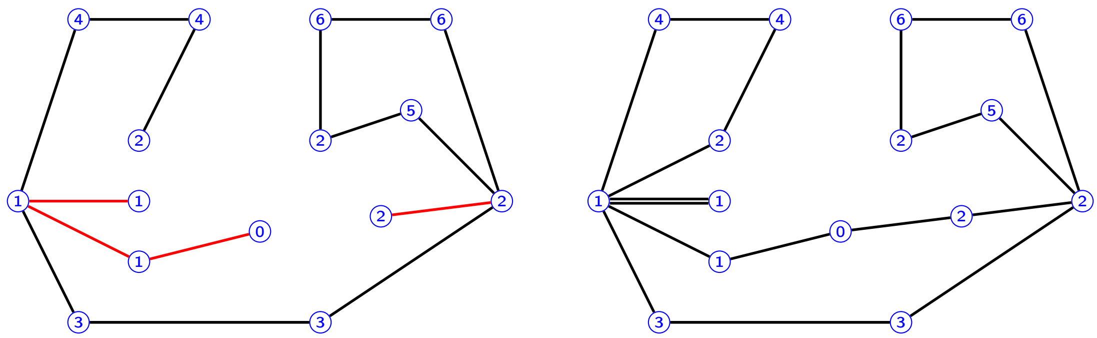
Lemma 19. *Let G be a 2-vertex-connected graph with a nice ear-decomposition that has no trivial ears and for which the union of all short ears have $|V| - |V_M| - \mu$ connected components. Then one can compute a tour in G with at most $\frac{7}{5}\Lambda$ edges in $O(n^3)$ time.*

Proof. If $\pi \leq \frac{\Lambda}{10}$, apply Theorem 15 and obtain a tour with at most $\frac{4}{3}(n-1) + \frac{2}{3}\pi \leq \frac{7}{5}\Lambda$ edges.

Otherwise take all $|V_\pi| + \pi$ edges of pendant ears, where V_π denotes the internal vertices of pendant ears. Add at most $n - |V_\pi| - \mu - 1$ edges of $G[V \setminus V_\pi]$ to obtain a connected spanning subgraph. Let $T \subseteq V \setminus V_\pi$ be the set of vertices with odd degree in this subgraph. Then add a minimum T-join in $G[V \setminus V_\pi]$; by Theorem 13 it has at most $\frac{1}{2}(n - |V_\pi| - 1 + \varphi - \varphi_\pi)$ edges. Summing up, our tour has at most $\pi + n - \mu - 1 + \frac{1}{2}(n - |V_\pi| - 1 + \varphi - \varphi_\pi)$ edges. Observing $|V_\pi| \geq 4\pi - 2|M| - \varphi_\pi$, this is at most $L_\mu + \frac{1}{2}L_\varphi - \pi \leq \frac{7}{5}\Lambda$. $\square$

Figure 8 illustrates the second part of this proof for the ear-decomposition in Figure 7 (left-hand side).

Figure 8. *Left: the edges of pendant ears after optimizing short ears (black) and a minimal set of edges of non-pendant ears (red) to make a connected spanning subgraph. Right: For correcting parities we need three more edges; a possible resulting tour is shown.*



The overall algorithm is now easily described:

1. Compute a nice ear-decomposition of G (Lemma 16).
2. Optimize the short ears (Theorem 17).
3. Delete all trivial ears.
4. Apply Lemma 19 to each block of the remaining graph.

It is not difficult to show that the sum of the lower bounds Λ for all blocks equals the lower bound Λ for G. This implies that we have a $\frac{7}{5}$-approximation algorithm.

6 The path version and connected T-joins

What if we do not require the walk to be closed? Then we look for a (Hamiltonian) path in the metric closure. We assume that the endpoints are given (otherwise we can try all pairs, or take a tour and delete one edge) and distinct: the path must begin in s and end in t (where $s, t \in V$ and $s \neq t$). For any of the problems studied in this paper, this variant is called the *s-t-path version* or simply the *path version*.

6.1 Asymmetric path version

Obviously, any ρ-approximation for the s-t-path version implies a ρ-approximation algorithm for the ASYMMETRIC TSP itself: just guess any edge (t, s) in an optimum solution (fix s and try all $n-1$ possibilities for t). Feige and Singh [2007] showed that the opposite also holds approximately: any ρ-approximation for the ASYMMETRIC TSP implies a $(2+\epsilon)\rho$-approximation algorithm for the path version.

6.2 Undirected path version and connected T-joins

In the undirected case, if we ask for a walk from s to t, by Section 1.3 this is equivalent to ask for a connected spanning multi-subgraph in which s and t have odd degree and all other vertices have even degree. It is natural to generalize this further to prescribe arbitrary parities: the CONNECTED T-JOIN PROBLEM asks for a set F such that (V, F) is a connected graph and F is a T-join. We call such a set F simply a *connected T-join*, or a *T-tour*. Again, F may contain pairs of parallel edges.

For $T = \emptyset$ this is the SYMMETRIC TSP, and for $T = \{s, t\}$ this is its s-t-path version. Again we may consider the GRAPHIC special case, where all edges of F must be copies of edges of the input graph.

Christofides' [1976] algorithm also works for the CONNECTED T-JOIN PROBLEM: take a minimum spanning tree (V, S) and add a minimum-weight T_S-join, where T_S is now the set of vertices whose degree in (V, S) has the wrong parity (so S is a $(T_S \triangle T)$-join).

However, this generalization of Christofides' algorithm is only a $\frac{5}{3}$-approximation algorithm (Hoogeveen [1991], Sebő and Vygen [2012]). To see this, let (V, S) be a minimum spanning tree. Let J be a minimum weight T_S-join, and J^* an optimum solution (a minimum weight connected T-join). Then $S \,\dot\cup\, J^*$ is a T_S-join. Since both S and J^* are connected, each contains a T_S-join; so $S \,\dot\cup\, J^*$ can be partitioned into three T_S-joins. Hence $3c(S \,\dot\cup\, J) \leq 3c(S) + 3c(J) \leq 3c(S) + c(S \,\dot\cup\, J^*) = 4c(S) + c(J^*) \leq 5c(J^*)$. The bound is tight even for $|T| = 2$ (Hoogeveen [1991]), as the graphs $(\{0, \ldots, 3k\}, \{\{i, i+1\} : 0 \leq i < 3k\} \cup \{\{3i, 3i+3\} : 0 \leq i < k\})$ show.

Sebő and Vygen [2012] showed that their techniques (outlined in Section 5 above) also lead to a $\frac{3}{2}$-approximation algorithm for the GRAPHIC CONNECTED T-JOIN PROBLEM. Previously, there were only algorithms for the special case $|T| = 2$; here the best was the 1.578-approximation algorithm of An, Kleinberg and Shmoys [2012].

6.3 Best-of-many Christofides' algorithm

An, Kleinberg and Shmoys [2012] also found a 1.619-approximation algorithm for the path version ($|T|=2$) with general weights. This was the first improvement of Christofides' algorithm that is not restricted to the graphic special case. This algorithm was generalized by Cheriyan, Friggstad and Gao [2012]; they obtain an approximation ratio of 1.625 for $|T|\geq 4$. Then Sebő [2012] obtained an $\frac{8}{5}$-approximation algorithm for arbitrary T and general weights.

All these three papers analyze essentially the same algorithm, which An, Kleinberg and Shmoys [2012] called *best-of-many Christofides*: it computes an optimum solution to a natural LP relaxation (see (11) below) and writes it as convex combination of spanning trees (plus a nonnegative vector). For each of these spanning trees, S, we again compute a minimum weight T_S-join J, where T_S is the set of vertices of S whose degree has the wrong parity, and output the best of these T-tours $S\,\dot\cup\,J$.

Following Sebő and Vygen [2012] and Sebő [2012], we consider the LP relaxation

$$\begin{array}{rrcll} \min & c(x) & & & \\ \text{subject to} & x(\delta(U)) & \geq & 2 & (\emptyset\neq U\subset V,\ |U\cap T| \text{ even}) \\ & x(\delta(\mathcal{W})) & \geq & |\mathcal{W}|-1 & (\mathcal{W} \text{ partition of } V) \\ & x_e & \geq & 0 & (e\in E) \end{array} \tag{11}$$

Here $\delta(\mathcal{W})$ denotes the set of edges with endpoints in different classes of the partition $\mathcal{W}$. For an optimum solution x (in fact for every feasible solution) we can write $x\geq\sum_{S\in\mathcal{S}}p_S\chi^S$, where again $\mathcal{S}$ denotes the set of edge sets of spanning trees, χ^S denotes the incidence vector of S, and $p_S\geq 0$ for all $S\in\mathcal{S}$ and $\sum_{S\in\mathcal{S}}p_S=1$. (An, Kleinberg and Shmoys [2012] and Cheriyan, Friggstad and Gao [2012] work in the metric closure and use a stronger LP in order to obtain $x=\sum_{S\in\mathcal{S}}p_S\chi^S$, but this is not necessary.)

By Carathéodory's theorem we can assume that $p_S>0$ for less than n^2 spanning trees (V,S). An optimum LP solution x, such spanning trees, and such numbers p_S can be computed in polynomial time, as can be shown with the ellipsoid method. For each $S\in\mathcal{S}$ with $p_S>0$, the algorithm computes a minimum weight T_S-join J and consider the T-tour $S\,\dot\cup\,J$; we output the best of these. Its cost is $\min_{S\in\mathcal{S}:\,p_S>0}(c(S)+\min\{c(J): J \text{ is a } T_S\text{-join}\})\leq\sum_{S\in\mathcal{S}}p_S(c(S)+\min\{c(J): J \text{ is a } T_S\text{-join}\})\leq c(x)+\sum_{S\in\mathcal{S}}p_Sc(y^S)$, for any set of vectors $(y^S)_{S\in\mathcal{S}}$ such that y^S is in the T_S-join polyhedron (cf. (4)). The difficulty in the analysis lies in finding an appropriate set of vectors $(y^S)_{S\in\mathcal{S}}$.

6.4 Analysis

Let $\mathcal{Q}:=\{Q=\delta(U):\emptyset\neq U\subset V,\ x(Q)<2\}$. An, Kleinberg and Shmoys [2012] proposed to choose $y^S:=(1-2\beta)\chi^S+\beta x+r^S$, where $\beta\leq\frac{1}{2}$ and r^S is a nonnegative vector satisfying

$$r^S(Q)\ \geq\ 4\beta-1-\beta x(Q) \tag{12}$$

for all $S\in\mathcal{S}$ and all $Q\in\mathcal{Q}$ with $|Q\cap S|\geq 2$.

Then for each $S\in\mathcal{S}$ and each T_S-cut Q we have $y^S(Q)\geq 1$. Indeed, if $Q\notin\mathcal{Q}$, then $y^S(Q)\geq(1-2\beta)|S\cap Q|+\beta x(Q)\geq 1-2\beta+2\beta=1$. If $Q\in\mathcal{Q}$, then Q is not only a T_S-cut but also a T-cut, so $|Q\cap S|$ is even and hence at least two, and we have $y^S(Q)=(1-2\beta)|Q\cap S|+\beta x(Q)+r^S(Q)\geq 2-4\beta+\beta x(Q)+r^S(Q)\geq 1$.

So y^S is in the T_S-join polyhedron for all $S\in\mathcal{S}$. Moreover, $\sum_{S\in\mathcal{S}}p_Sc(y^S)\leq(1-2\beta)\sum_{S\in\mathcal{S}}p_Sc(S)+\beta c(x)+\sum_{S\in\mathcal{S}}p_Sc(r^S)\leq(1-\beta)c(x)+\sum_{S\in\mathcal{S}}p_Sc(r^S)$.

An, Kleinberg and Shmoys [2012] chose $\beta=1/\sqrt{5}$ and found a vector r such that $r^S=r$ for all $S\in\mathcal{S}$ satisfies (12) and $c(r)\leq(7\sqrt{5}-15)/10$, yielding the approximation ratio $(1+\sqrt{5})/2$ (the golden ratio).

Sebő [2012] improved this by letting $v^Q := \sum_{S\in\mathcal{S}:|Q\cap S|=1} p_S \chi^{Q\cap S}$ for $Q \in \mathcal{Q}$, and

$$r^S \;:=\; \sum_{Q\in\mathcal{Q}:|Q\cap S|\geq 2} \max\left\{0, \frac{4\beta-1-\beta x(Q)}{2-x(Q)}\right\} v^Q.$$

Note that $v^Q(Q) = \sum_{S\in\mathcal{S}:|Q\cap S|=1} p_S \geq 2-\sum_{S\in\mathcal{S}} p_S|Q\cap S| \geq 2-x(Q)$. To show (12), simply observe that for $S\in\mathcal{S}$ and $Q\in\mathcal{Q}$ with $|Q\cap S|=2$ we have $r^S(Q) \geq \frac{4\beta-1-\beta x(Q)}{2-x(Q)} v^Q(Q)$.

We now bound the cost. Note that $\sum_{S\in\mathcal{S}:|Q\cap S|\geq 2} p_S \leq x(Q)-1$ for $Q\in\mathcal{Q}$. Using this, we obtain the bound $\sum_{S\in\mathcal{S}} p_S c(r^S) \leq \sum_{Q\in\mathcal{Q}}(x(Q)-1)\max\left\{0, \frac{4\beta-1-\beta x(Q)}{2-x(Q)}\right\} c(v^Q) \leq \sum_{Q\in\mathcal{Q}} \frac{1}{9} c(v^Q) = \frac{1}{9}\sum_{S\in\mathcal{S}} p_S \sum_{Q\in\mathcal{Q}:|Q\cap S|=1} c(Q\cap S) \leq \frac{1}{9}\sum_{S\in\mathcal{S}} p_S c(S\setminus J_S)$, where J_S denotes the unique subset of S that is a T_S-join.

Let $\beta = \frac{4}{9}$. If $\sum_{S\in\mathcal{S}} p_S c(S\setminus J_S) \leq \frac{2}{5}c(x)$, we have $\sum_{S\in\mathcal{S}} p_S c(y^S) \leq (1-\frac{4}{9}+\frac{2}{45})c(x) = \frac{3}{5}c(x)$. Otherwise $\sum_{S\in\mathcal{S}} p_S c(J_S) \leq \frac{3}{5}c(x)$, and noting that χ^{J_S} is in the T_S-join polyhedron, we get the same bound. This gives Sebő's [2012] $\frac{8}{5}$-approximation algorithm.

7 Further results

We briefly mention some other related results. However, it is impossible to mention all important results here.

7.1 Inapproximability

Lampis [2012] proved that no $\frac{185}{184}$-approximation algorithm exists for the Symmetric TSP unless *P* = *NP*. Papadimitriou and Vempala [2006] proved that no $\frac{118}{117}$-approximation algorithm exists for the Asymmetric TSP unless *P* = *NP*.

7.2 Integrality Ratios

We have seen in Section 2.2 that the integrality ratio of (2) is at least $\frac{4}{3}$ even for graphic metrics. The integrality ratio of (2) is conjectured to be exactly $\frac{4}{3}$ even for general metrics, but this so-called *TSP-$\frac{4}{3}$-conjecture* is open; we only know Wolsey's [1980] upper bound $\frac{3}{2}$ in general (cf. Section 2.5) and the upper bound $\frac{7}{5}$ for graphic metrics by Sebő and Vygen [2012] (cf. Section 5).

The *TSP-$\frac{4}{3}$-conjecture* is supported by computational verification for $n \leq 12$ (Boyd and Elliott-Magwood [2007]) and by theoretical work of Goemans [1995], who proved that for any instance with ratio greater than $\frac{4}{3}$ even the LP that arises from (7) by adding many classes of inequalities that are valid for the graphical traveling salesman polyhedron does not have an integral optimum solution.

Schalekamp, Williamson and van Zuylen [2012] showed that the worst ratio of an optimum perfect 2-matching over (2) is $\frac{10}{9}$, as conjectured by Boyd and Carr [2011].

For (6) and (7), the integrality ratio is between $\frac{6}{5}$ and $\frac{3}{2}$ (Alexander, Boyd and Elliott-Magwood [2006]). Carr and Ravi [1998] conjectured it to be $\frac{4}{3}$. The ratio restricted to unit weights is between $\frac{9}{8}$ and $\frac{4}{3}$ (Sebő and Vygen [2012]).

The integrality ratio of the asymmetric subtour LP is at least 2 (shown by Charikar, Goemans and Karloff [2006], disproving a conjecture of Carr and Vempala [2004]) and at most $2 + 8\ln n/\ln\ln n$ (Asadpour et al. [2010]). The same holds for the path version (Friggstad, Gupta and Singh [2012]).

7.3 Further special cases

An even more special case than the GRAPHIC TSP is the 1-2-TSP, in which $c(v,w) \in \{1,2\}$ for all $v,w \in V$. To see that this is essentially a special case of the GRAPHIC TSP (up to an additive constant of 1), add a vertex x and consider the graph $(V \cup \{x\}, \{\{v,x\} : v \in V\} \cup \{\{v,w\} : v,w \in V, u \neq v, c(v,w) = 1\})$ The 1-2-TSP has an $\frac{8}{7}$-approximation algorithm (Berman and Karpinski [2006]) but no $\frac{744}{743}$-approximation algorithm (Engebretsen and Karpinksi [2006]). The integrality ratio of (2) for the 1-2-TSP is between $\frac{10}{9}$ and $\frac{19}{15}$ (Qian et al. [2012]).

In the special case of the GRAPHIC TSP where the instance is a k-regular graph (with k large), Vishnoi [2012] showed how to find a tour of length at most $(1+\sqrt{64/\ln k})n$ in polynomial time.

7.4 Geometric instances and planar graphs

Arora [1998] found an approximation scheme for geometric instances. Here, each city is associated with a point in $\mathbb{R}^d$, and the distances are ℓ_p-distances. This case is also *NP*-hard, for any fixed $d \geq 2$ and any p. The most prominent case $d = p = 2$ is called the EUCLIDEAN TSP (see also Mitchell [1999]). Rao and Smith [1998] improved the running time: for every fixed $\epsilon > 0$ they have a $(1+\epsilon)$-approximation algorithm that runs in $O(n \log n)$ time. However, the constants involved are still quite large for reasonable values of ϵ, and thus the practical value seems to be limited. Bartal, Gottlieb and Krauthgamer [2012] found a randomized approximation scheme for metric spaces with bounded doubling dimension.

For planar graphs with nonnegative edge weights, Klein [2008] found an approximation scheme that has linear running time for every fixed $\epsilon > 0$. An approximation scheme exists even for bounded genus graphs (Demaine, Hajiaghayi and Mohar [2010]).

Interestingly, it is not known whether the decision version of the EUCLIDEAN TSP belongs to *NP*.

7.5 Polyhedral Descriptions

Many classes of facets of the TSP polytope have been discovered, but a complete description is out of reach. Recently, Fiorini et al. [2012] proved that every polyhedron that projects to the TSP polytope (i.e., any *extended formulation*) has $2^{\Omega(n^{1/4})}$ facets. It may not be surprising that the TSP has no compact extended formulation, but this was not known before, and this result is unconditional (i.e., it does not assume *P* $\neq$ *NP*). The proof reveals an interesting connection to communication complexity.

7.6 The 2ECSS problem

The integral solutions to (6) are the 2-edge-connected spanning subgraphs (2ECSS). Monma, Munson and Pulleyblank [1990] showed that the smallest 2ECSS can be smaller than the shortest tour by up to a factor $\frac{4}{3}$; this also follows directly from applying Theorem 11 to each block of a smallest 2ECSS. The bound is tight as Figure 2 shows.

Sebő and Vygen [2012] observed that the techniques of Section 5 directly imply a $\frac{4}{3}$-approximation algorithm for the (unweighted) 2ECSS problem. Indeed, if $\pi \geq \frac{1}{6}\mathrm{LP}(G)$, the second case of the proof of Lemma 19 yields a tour (and hence a 2ECSS) of length $\frac{4}{3}\mathrm{LP}(G)$. Otherwise one can simply take all k nontrivial ears: we get $n-1+k$ edges, and this is at most $\frac{5}{4}L_\varphi + \frac{\pi}{2}$ since $n-1 \geq 4k - 2\pi - \varphi(G)$.

Better approximation ratios have been claimed, but no complete proof has been published.

Fernandes [1998] proved that the problem is *MAXSNP*-hard. For the weighted case, Khuller and Vishkin [1994] found a 2-approximation algorithm, which is still the best known.

Sebő and Vygen [2012] also showed the following: if there is a ρ-approximation algorithm for the unweighted 2ECSS problem, then there is a $\frac{2}{3}(\rho+1)$-approximation algorithm for the Graphic TSP.

8 Open problems

We conclude this survey by listing some open research problems that we consider important. Almost all of these problems have been formulated earlier, and indeed most of them are very natural. None of them seems to be easy. However, given the remarkable progress that has been made during the last few years, one may hope that we will see some solutions soon.

1. Improve Christofides' algorithm: find a ρ-approximation algorithm for the Symmetric TSP for some $\rho < \frac{3}{2}$.

2. Find a constant-factor approximation algorithm for the Asymmetric TSP, or at least the special case in which a strongly connected digraph (V, E) is given and $c(v, w) = 1$ if $(v, w) \in E$ and $c(v, w) = \infty$ otherwise (one might call this the Digraphic TSP).

3. Determine the integrality ratio of the subtour relaxation (2) of the Symmetric TSP.

4. Prove a better bound on the integrality ratio for another (polynomial-time solvable) LP relaxation of the Symmetric TSP.

5. Solve the LPs (2), (6), (7), and (8) by combinatorial algorithms.

6. Answer the question whether the integrality ratio of the directed subtour relaxation (8) is bounded by a constant.

7. How good is the best-of-many Christofides' algorithm (cf. Section 6.3) really; i.e., what is the worst case? The answer can of course be different for $T = \emptyset$ (the Symmetric TSP) and for general T.

8. Improve the lower bounds on the approximability substantially.

9. Find a $\frac{4}{3}$-approximation algorithm for the Graphic TSP.

10. Find a $\frac{3}{2}$-approximation algorithm for the Connected T-Join Problem with arbitrary nonnegative weights, at least in the special case $|T| = 2$.

11. Improve on the 2-approximation algorithm for the weighted 2ECSS problem. Note that Wolsey's analysis (Section 2.5) shows that Christofides' algorithm is also a $\frac{3}{2}$-approximation algorithm for the variant of the 2ECSS problem where doubling edges is allowed. However, in contrast to the unweighted special case, allowing to double edges really changes the problem.

Acknowledgement

Thanks to Corinna Gottschalk, Swati Gupta, Satoru Iwata, Volker Kaibel, Marcin Mucha, R. Ravi, András Sebő, David Shmoys, and Ola Svensson for careful reading and useful remarks.

Bibliography

Alexander, A., Boyd, S., and Elliott-Magwood, P. [2006]: On the integrality gap of the 2-edge connected subgraph problem. Technical Report TR-2006-04, SITE, University of Ottawa 2006

An, H.-C., Kleinberg, R., and Shmoys, D.B. [2012]: Improving Christofides' algorithm for the s-t path TSP. Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC 2012), 875–886

Applegate, D.L., Bixby, R., Chvátal, V., and Cook, W.J. [2006]: The Traveling Salesman Problem: A Computational Study. Princeton University Press 2006

Arora, S. [1998]: Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. Journal of the ACM 45 (1998), 753–782

Asadpour, A., Goemans, M.X., Mądry, A., Oveis Gharan, S., and Saberi, A. [2010]: An $O(\log n/\log\log n)$-approximation algorithm for the asymmetric traveling salesman problem. Proceedings of the 21st Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2010), 379–389

Bartal, Y., Gottlieb, L.-A., and Krauthgamer, R. [2012]: The traveling salesman problem: low-dimensionality implies a polynomial time approximation scheme. Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC 2012), 663–672

Berman, P., and Karpinski, M. [2006]: 8/7-approximation algorithm for (1,2)-TSP. Proceedings of the 17th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2006), 641–648

Bläser, M. [2003]: A new approximation algorithm for the asymmetric TSP with triangle inequality. Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2003), 638–645

Boyd, S., and Carr, R. [2011]: Finding low cost TSP and 2-matching solutions using certain half-integer subtour vertices. Discrete Optimization 8 (2011), 525–539

Boyd, S., and Elliott-Magwood, P. [2007]: Structure of the extreme points of the subtour elimination polytope of the STSP. Technical Report TR-2007-09, SITE, University of Ottawa, Ottawa, 2007

Boyd, S., Sitters, R., van der Ster, S., and Stougie, L. [2011]: TSP on cubic and subcubic graphs. In: Integer Programming and Combinatorial Optimization; Proceedings of the 15th International IPCO Conference; LNCS 6655 (O. Günlük, G.J. Woeginger, eds.), Springer, Berlin 2011, pp. 65–77

Carr, R., and Ravi, R. [1998]: A new bound for the 2-edge connected subgraph problem. In: Integer Programming and Combinatorial Optimization; Proceedings of the 6th International IPCO Conference; LNCS 1412 (R.E. Bixby, E.A. Boyd, R.Z. Ríos-Mercado, eds.), Springer, Berlin 1998; pp. 112-125

Carr, R., and Vempala, S. [2004]: On the Held-Karp relxation for the asymmetric and symmetric traveling salesman problems. Mathematical Programming A 100 (2004), 569–587

Charikar, M., Goemans, M.X., and Karloff, H. [2006]: On the integrality ratio for the asymmetric traveling salesman problem. Mathematics of Operations Research 31 (2006), 245–252

Chekuri, C., Vondrák, J., and Zenklusen, R. [2010]: Dependent randomized rounding via exchange properties of combinatorial structures. Proceedings of the 51st Annual IEEE Symposium on Foundations of Computer Science (FOCS 2010), 575–584

Cheriyan, J., Friggstad, Z., and Gao, Z. [2012]: Approximating minimum-cost connected T-joins. Approximation, Randomization, and Combinatorial Optimization; Proceedings of the 16th APPROX Workshop; LNCS 7408 (A. Gupta, K. Jansen, J. Rolim, R. Servedio, eds.), Springer, Berlin 2012, pp. 110–121

Cheriyan, J., Sebő, A., and Szigeti, Z. [2001]: Improving on the 1.5-approximation of a smallest 2-edge connected spanning subgraph. SIAM Journal on Discrete Mathematics 14 (2001), 170–180

Christofides, N. [1976]: Worst-case analysis of a new heuristic for the traveling salesman problem. Technical Report 388, Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh 1976

Cook, W.J. [2012]: In Pursuit of the Traveling Salesman: Mathematics at the Limits of Computation. Princeton University Press 2012

Cornuéjols, G., Fonlupt, J., and Naddef, D. [1985]: The traveling salesman problem on a graph and some related integer polyhedra. Mathematical Programming 33 (1985), 1–27

Correa, J.R., Larré, O., and Soto, J.A. [2012]: TSP tours in cubic graphs: beyond 4/3. In: Algorithms – ESA 2012; LNCS 7501 (L. Epstein, P. Ferragina, eds.), Springer, Berlin 2012, pp. 790–801

Dantzig, G.B., Fulkerson, D.R., and Johnson, S.M. [1954]: Solution of a large scale traveling salesman problem. Operations Research 2 (1954), 393–410

Demaine, E.D., Hajiaghayi, M., and Mohar, B. [2010]: Approximation algorithms via contraction decomposition. Combinatorica 30 (2010), 533–552

Edmonds, J. [1965]: The Chinese postman's problem. Bulletin of the Operations Research Society of America 13 (1965), B-73

Edmonds, J. [1970]: Submodular functions, matroids and certain polyhedra. In: Combinatorial Structures and Their Applications; Proceedings of the Calgary International Conference on Combinatorial Structures and Their Applications 1969 (R. Guy, H. Hanani, N. Sauer, J. Schönheim, eds.), Gordon and Breach, New York 1970, pp. 69–87

Edmonds, J., and Johnson, E.L. [1973]: Matching, Euler tours and the Chinese postman. Mathematical Programming 5 (1973), 88–124

Engebretsen, L., and Karpinski, M. [2006]: TSP with bounded metrics. Journal of Computer and System Sciences 72 (2006), 509–546

Euler, L. [1736]: Solutio problematis ad geometriam situs pertinentis. Commentarii Academiae Petropolitanae 8 (1736), 128–140

Feige, U., and Singh, M. [2007]: Improved approximation algorithms for traveling salesperson tours and paths in directed graphs. Proceedings of the 10th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems; LNCS 4627 (M. Charikar, K. Jansen, O. Reingold, J.D.P. Rolim, eds.), Springer, Berlin 2007, pp. 104–118

Fernandes, C.G. [1998]: A better approximation ratio for the minimum size k-edge-connected spanning subgraph problem. Journal of Algorithms 28 (1998), 105–124

Fiorini, S., Massar, S., Pokutta, S., Tiwary, H.R., and de Wolf, R. [2012]: Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC 2012), 95–106

Frank, A. [1993]: Conservative weightings and ear-decompositions of graphs. Combinatorica 13 (1993), 65–81

Frieze, A.M., Galbiati, G., and Maffioli, F. [1982]: On the worst-case performance of some algorithms for the asymmetric traveling salesman problem. Networks 12 (1982), 23–39

Friggstad, Z., Gupta, A., and Singh. M. [2012]: An improved integrality gap for asymmetric TSP paths. Accepted for IPCO 2013

Gamarnik D., Lewenstein M., and Sviridenko M. [2005]: An improved upper bound for the TSP in cubic 3-edge-connected graphs. Operations Research Letters, 33 (2005), 467–474

Garey, M.R., Johnson, D.S., and Tarjan, R.E. [1976]: The planar Hamiltonian circuit problem is NP-complete. SIAM Journal on Computing 5 (1976), 704–714

Goemans, M.X. [1995]: Worst-case comparison of valid inequalities for the TSP. Mathematical Programming, 69 (1995), 335–349

Goemans, M.X. [2006]: Minimum bounded-degree spanning trees. Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS 2006), 273–282

Goemans, M.X., and Bertsimas, D.J. [1993]: Survivable networks, linear programming relaxations and the parsimonious property. Mathematical Programming 60 (1993), 145–166

Held, M., and Karp, R.M. [1970]: The traveling-salesman problem and minimum spanning trees. Operations Research 18 (1970), 1138–1162

Hierholzer, C. [1873]: Über die Möglichkeit, einen Linienzug ohne Wiederholung und ohne Unterbrechung zu umfahren. Mathematische Annalen 6 (1873), 30–32

Hoffman, A.J. [1960]: Some recent applications of the theory of linear inequalities to extremal combinatorial analysis. In: Combinatorial Analysis (R.E. Bellman, M. Hall, eds.), AMS, Providence 1960, pp. 113–128

Hoogeveen, J.A. [1991]: Analysis of Christofides' heuristic: some paths are more difficult than cycles. Operations Research Letters 10 (1991), 291–295

Kaplan, H., Lewenstein, M., Shafrir, N., and Sviridenko, M. [2005]: Approximation algorithms for asymmetric TSP by decomposing directed regular multigraphs. Journal of the ACM 52 (2005), 602–626

Karger, D.R., and Stein, C. [1996]: A new approach to the minimum cut problem. Journal of the ACM 43 (1996), 601–640

Karp, R.M. [1972]: Reducibility among combinatorial problems. In: Complexity of Computer Computations (R.E. Miller, J.W. Thatcher, eds.), Plenum Press, New York 1972, pp. 85–103

Khuller, S., and Vishkin, U. [1994]: Biconnectivity approximations and graph carvings. Journal of the ACM 41 (1994), 214–235

Klein, P.N. [2008]: A linear-time approximation scheme for TSP in undirected planar graphs with edge-weights. SIAM Journal on Computing 37 (2008), 1926–1952

Korte, B., and Vygen, J. [2012]: Combinatorial Optimization: Theory and Algorithms. Springer, Berlin, Fifth Edition 2012

Lampis, M. [2012]: Improved inapproximability for TSP. Approximation, Randomization, and Combinatorial Optimization; Proceedings of the 16th APPROX Workshop; LNCS 7408 (A. Gupta, K. Jansen, J. Rolim, R. Servedio, eds.), Springer, Berlin 2012, pp. 243–253

Lovász, L. [1976]: On some connectivity properties of Eulerian graphs. Acta Mathematica Academiae Scientiarum Hungaricae 28 (1976), 129–138

Mitchell, J. [1999]: Guillotine subdivisions approximate polygonal subdivisions: a simple polynomial-time approximation scheme for geometric TSP, k-MST, and related problems. SIAM Journal on Computing 28 (1999), 1298–1309

Mömke, T., and Svensson, O. [2011]: Approximating graphic TSP by matchings. Proceedings of the 52nd Annual Symposium on Foundations of Computer Science (FOCS 2011), 560–569

Monma, C.L., Munson, B.S., and Pulleyblank, W.R. [1990]: Minimum-weight two-connected spanning networks. Mathematical Programming 46 (1990), 153–171

Mucha, M. [2012]: $\frac{13}{9}$-approximation for graphic TSP. Proceedings of the 29th International Symposium on Theoretical Aspects of Computer Science (2012), 30–41

Oveis Gharan, S., Saberi, A., and Singh, M. [2011]: A randomized rounding approach to the traveling salesman problem. Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2011), 550–559

Papadimitriou, C.H., and Vempala, S. [2006]: On the approximability of the traveling salesman problem. Combinatorica 26 (2006), 101–120

Papadimitriou, C.H., and Yannakakis, M. [1993]: The traveling salesman problem with distances one and two. Mathematics of Operations Research 18 (1993), 1–12

Qian, J., Schalekamp, F., Williamson, D.P., and van Zuylen, A. [2012]: On the integrality gap of the subtour LP for the 1,2-TSP. In: LATIN 2012; Theoretical Informatics; LNCS 7256 (D. Fernández-Baca, ed.), Springer, Berlin 2012, pp. 606–617

Rado, R. [1942]: A theorem on independence relations. Quarterly Journal of Mathematics 13 (1942), 83–89

Rao, S.B., and Smith, W.D. [1998]: Approximating geometric graphs via "spanners" and "banyans". Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC 1998), 540–550

Sahni, S., and Gonzalez, T. [1976]: P-complete approximation problems. Journal of the ACM 23 (1976), 555–565

Schalekamp, F., Williamson, D.P., and van Zuylen, A. [2012]: A proof of the Boyd-Carr conjecture. Proceedings of the 23rd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA 2012), pp. 1477–1486

Schrijver, A. [2003]: Combinatorial Optimization: Polyhedra and Efficiency. Springer, Berlin 2003

Sebő, A. [2012]: Eight fifth approximation for TSP paths. arXiv:1209.3523; accepted for IPCO 2013

Sebő, A., and Vygen, J. [2012]: Shorter tours by nicer ears: 7/5-approximation for graphic TSP, 3/2 for the path version, and 4/3 for two-edge-connected subgraphs. arXiv:1201.1870

Vishnoi, N.K. [2012]: A permanent approach to the traveling salesman problem, Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science (FOCS 2012), 76–80

Whitney, H. [1932]: Non-separable and planar graphs. Transactions of the American Mathematical Society 34 (1932), 339–362

Wolsey, L.A. [1980]: Heuristic analysis, linear programming and branch and bound. Mathematical Programming Study 13 (1980), 121–134