

THE LAZY TRAVELLING SALESMAN PROBLEM IN \mathbb{R}^2

PAZ POLAK¹ AND GERSHON WOLANSKY²

Abstract. We study a parameter (σ) dependent relaxation of the Travelling Salesman Problem on \mathbb{R}^2 . The relaxed problem is reduced to the Travelling Salesman Problem as $\sigma \rightarrow 0$. For increasing σ it is also an ordered clustering algorithm for a set of points in \mathbb{R}^2 . A dual formulation is introduced, which reduces the problem to a convex optimization, provided the minimizer is in the domain of convexity of the relaxed functional. It is shown that this last condition is generically satisfied, provided σ is large enough.

Mathematics Subject Classification. 49K30, 49K35, 65K10.

Received February 23, 2005. Revised January 31, 2006.
Published online June 20, 2007.

1. INTRODUCTION

The object of this paper is to investigate an optimization problem which we name “The Lazy Travelling Salesman Problem” (LTSP). As the name suggests, it is closely related to the classical Problem of Travelling Salesman (TSP).

The version of the LTSP we have in mind is as follows: let V be a set of n points $v_1, \dots, v_n \in \mathbb{R}^2$. Let $\sigma > 0$. Given a closed, Jordan curve $\Gamma \subset \mathbb{R}^2$, let $L(\Gamma)$ be the length of this curve and $\text{Dist}(v, \Gamma)$ the distance of a point $v \in \mathbb{R}^2$ to Γ . The object is to find such a closed curve Γ which minimize

$$F_\sigma(\Gamma) := \frac{1}{n} \sum_{i=1}^n \text{dist}^2(v_i, \Gamma) + \sigma L^2(\Gamma). \quad (1)$$

It is easily seen that a minimal curve Γ must be a polygonal, closed curve, of *at most* n vertices (and edges). Given this fact, we can restrict F_σ to the set of closed polygons of (at most) n vertices and represent it as a function of $\vec{u} \in \mathbb{R}^n \times \mathbb{R}^n$, $\vec{u} = (u_1, \dots, u_n)$ where $u_j \in \mathbb{R}^2$, via

$$\overline{\mathbf{G}}_{\sigma,n}(\vec{u}) := \frac{1}{n} \sum_{j=1}^n \min_{1 \leq i \leq n} |u_i - v_j|^2 + \sigma L^2(\vec{u}) \quad (2)$$

Keywords and phrases. Travelling Salesman Problem, Legendre-Fenchel transform.

¹ Weizmann Institute of Science, Rehovot, Israel.

² Department of Mathematics, Technion, Haifa 32000, Israel; gershonw@math.technion.ac.il

where

$$L(\vec{u}) := \sqrt{\left[\sum_1^n |u_{i+1} - u_i| \right]^2}. \tag{3}$$

Here, the closed polygon Γ is represented by an ordered set of the (not necessarily distinct) vertices u_1, \dots, u_n , where we identify u_{n+1} with u_1 .

Note that F_σ is a relaxation of a version of the TSP defined as:

TSP: Find the shortest possible closed path in \mathbb{R}^2 which contains all the points $\{v_1, \dots, v_n\} := V$.

Indeed, TSP is the limit of the LTSP where $\sigma \rightarrow 0$. Intuitively, this follows from the observation that, in this limit, the first term of F_σ must be very small for the minimizer, so the optimal curve Γ corresponding to small $\sigma > 0$ must be very close to any point $v \in V$. Hence the optimal Γ is a polygon of n vertices, where each vertex is very close (in fact, $O(\sigma^{1/2})$ close) to a corresponding $v \in V$. Once the set of vertices $u_1, \dots, u_n \in \mathbb{R}^2$ of Γ is determined, their ordering should minimize the second part of F_σ , namely, the polygonal curve Γ is the solution of TSP corresponding to the points $u_1, \dots, u_n \approx v_1, \dots, v_n$ in the optimal order.

The LTSP can be seen in a broader context of snakes, elastic networks and Kohonen maps.

The snake was proposed in 1987 [13] as an active contour algorithm for edge linking. This is a basic operation in image analysis whose goal is as follows: For a given set of n points representing an edge of an image, find an ordered set (usually large) of m points which approximates a continuous boundary. It is formulated as a nonlinear optimization problem for the cost function (see [2])

$$J(\vec{u}) = - \sum_{i=1}^m \sum_{j=1}^n \phi_\beta (|u_i - v_j|^2) + \sigma \vec{u}^t \mathbf{A} \vec{u} \tag{4}$$

where $\phi_\beta(d) = \exp(-d/2\beta^2)$, A is a symmetric, $2m \times 2m$ Toeplitz matrix, and $\sigma > 0$ a fixed parameter (which may depend on m).

Elastic networks were proposed by Durbin and Willshaw [3], (see also [9]) for the purpose of finding the shorts routes of the TSP. It was also suggested as a tool for image analysis [12]. It's objective, similar to this of the snake, is to order the set of detected points along a continuous curve. However, there is a difference between the two methods, as an edge detector (such as the snake) is designed to disregard spurious edge points and, in addition, to penalize curves of large curvature. The elastic network, on the other hand, is proposed to take into account all the points in the data base and penalize long curves. The cost function for an elastic network is of the form

$$J(\vec{u}) = -2 \frac{\beta^2}{n} \sum_{i=1}^n \log \left(\sum_{j=1}^m \phi_\beta (|u_j - v_i|^2) \right) + m \sigma \vec{u}^t \mathbf{A} \vec{u} \tag{5}$$

where \mathbf{A} is a regularization matrix, e.g.

$$\vec{u}^t \mathbf{A} \vec{u} = \sum_{j=1}^m |u_{j+1} - u_j|^2; \quad u_{n+1} = u_1 \tag{6}$$

and the scaling factor m of the second term of (5) is determined so that it's stationary value is an approximation of the square of the path's length ¹ as $m \rightarrow \infty$. Note also that the first term in (5) tends to the first term of (2), as $\beta \rightarrow 0$.

The Kohonen map [7] is another algorithm of objects linking. Like (4) and (5), it contains an external potential term and a regularization term. A unified approach of snakes, elastic nets and Kohonen maps is described in [1]. For comparison of all these methods and the Burr's modified elastic net (which is very similar to our approach), see [4].

¹The function (6) is an approximation for the square length if $m \gg \Delta/\delta$, where Δ (respectively δ) is the maximal (respectively minimal) distance between two "cities" in V .

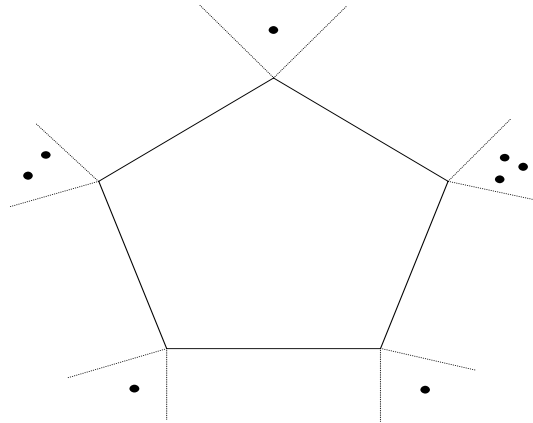


FIGURE 1. Typical solution of the LTSP when σ is large enough: the optimal polygon is convex and each input point belongs to a normal cone on one of the vertices (marked by the dash lines).

The LTSP can also be seen as a particular case of this unified approach. The difference between the LTSP and the elastic net (in the limit $\beta = 0$) is the form of the regularizing term. In LTSP the regularizing term (3) is given by the *exact* length square of the orbit while in elastic nets the regularizing term (6) is the sum of squares of the distances between consecutive points. The relation $m\vec{u}^t \mathbf{A}\vec{u} \approx L(\vec{u})$ is valid only for optimal solutions in the limit of large m .

The main advantage of the LTSP is, therefore, that it represents a *continuous* algorithm for the TSP *without* the need to increase the number of test points (m) above the number of “cities” (n). The disadvantage is that the path length (3) is not a quadratic function as in (6). As such, it is harder to calculate numerically (since it involves the evaluation of the square-root), and turns out to be non-smooth at points where two (or more) of the vertices coincide. This last point, however, makes it also more interesting, as explained below.

The interest we have in the LTSP is not only due to its relation to the TSP. As σ is increased, the optimal polygonal solutions Γ_σ may undergo a “phase change”, where two, or more of its vertices may coincide. The number of vertices k of the optimal polygons may satisfy $k < n$ for sufficiently large σ – see Figure 1.

This leads to an, apparently, new version of a clustering problem, where the given set $V \subset \mathbb{R}^2$ is clustered into k disjoint subsets of “mutually closed” points. Unlike other clustering algorithms, the number of clusters k is not determined *a priori*, but depends on the parameter σ . For a related method, see [5]. Application of clustering induced by TSP in biology can be found in [10].

In this paper we apply a dual formulation for the cost function (3) of the TSP, as well as to the LTSP (2). This dual formulation enables us to calculate the *convexification* of these problems in an efficient² way. This yields an efficient computation of *lower bound* of the optimal solution. In addition, we show that, for sufficiently large σ , the LTSP is convex near the optimal solution. In particular, it is possible to calculate efficiently the optimal solution for the clustering problem associated with the LTSP of sufficiently large σ .

1.1. Layout

In Section 2 we establish some notations and formulate the LTSP. We also study some properties of its minimizers. In Section 3 we consider the TSP as a limit $\sigma = 0$ of the LTSP as formulated in Section 2 and

²That is, in time which is polynomial in the number of “cities” n .

introduce its dual formulation. The dual problem is convex and can be computed efficiently. However, the optimality of this estimate, as well as the validity of the dual formulation, is conditioned on the assumption that the minimizer is in the domain of convexity of the TSP.

In Section 4 we extend the dual formulation to the LTSP. In addition, we introduce the analog of the results of Section 3 for this case.

Finally, in Section 5 we show the optimality of the main results of Section 4 for almost all distributions of points (v_1, \dots, v_n) if the relaxation parameter σ is large enough.

2. FORMULATION OF THE LTSP

Let $\{v_i\}_{i=1}^n$ be a set of points in $\mathbb{C} \equiv \mathbb{R}^2$. We define an order on this set by representing it as an element $\vec{v} \in \mathbb{C}^n$ in the following way

$$\vec{v} = (v_1, \dots, v_n).$$

Remark 2.1. We *always* identify v_{n+1} with v_1 .

Notations. Let S_n be the group of permutations on $\{1, \dots, n\}$, τ any permutation in S_n , τ_c the cyclic permutation $\tau_c(i) = i + 1$ and $\vec{u} = (u_1, \dots, u_n) \in \mathbb{C}^n$,

- (i) $\vec{u}_\tau := (u_{\tau_1}, \dots, u_{\tau_n})$;
- (ii) $\vec{\delta}u := \vec{u}_{\tau_c} - \vec{u} \in \mathbb{C}^n$;
- (iii) for $u, v \in \mathbb{C}$, $\overline{[u, v]} := \cup_{0 \leq t \leq 1} \{tu + (1-t)v\} \subset \mathbb{C}$;
- (iv) the closed polygon $\mathbf{p}_{\vec{\delta}u} := \cup_1^n \overline{[u_{i+1} - u_i]} \subset \mathbb{C}$;
- (v) the length of this path is $\mathbf{L}_n(\vec{u}) := L(\mathbf{p}_{\vec{\delta}u}) := \sum_{i=1}^n |u_{i+1} - u_i|$;
- (vi) $\|\vec{u} - \vec{v}\| := \sqrt{\sum_1^n |u_i - v_i|^2}$ stands for the Euclidean norm in \mathbb{C}^n .

Definition 2.2. Let $\mathbf{G}_{\sigma, n} : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{R}^+$ given by

$$\mathbf{G}_{\sigma, n}(\vec{u}, \vec{v}) := \frac{1}{n} \|\vec{u} - \vec{v}\|^2 + \sigma \mathbf{L}_n^2(\vec{u}).$$

Clearly, \mathbf{L}_n is a coercive, convex function. If we fix $\vec{v} \in \mathbb{C}^n$ then $\mathbf{G}_{\sigma, n}(\cdot, \vec{v})$ is a strictly convex, coercive and continuous function of the first variable. Hence

Lemma 2.3. For every $n \in \mathbb{N}$, $\sigma > 0$ and $\vec{v} \in \mathbb{C}^n$, $\mathbf{G}_{\sigma, n}(\cdot, \vec{v})$ has a unique minimizer.

Another definition we need is:

Definition 2.4. $\mathbf{G}_{\sigma, n}^\tau(\vec{u}, \vec{v}) := \mathbf{G}_{\sigma, n}(\vec{u}, \vec{v}_\tau) \quad \forall \vec{u}, \vec{v} \in \mathbb{C}^n$.

The last definition provides us with new function on \mathbb{C}^{2n} .

Definition 2.5. $\overline{\mathbf{G}}_{\sigma, n}(\vec{u}, \vec{v}) := \min_{\tau \in S_n} \mathbf{G}_{\sigma, n}^\tau(\vec{u}, \vec{v})$.

The observation that a minimizer Γ of the functional F_σ given by (1) is a polygon of at most n vertices implies that $\Gamma = \mathbf{p}_{\vec{\delta}u}$ where \vec{u} is a minimizer of $\overline{\mathbf{G}}_{\sigma, n}(\cdot, \vec{v})$. In particular, some of the coordinates of \vec{u} may coincide, if the number of the vertices of the optimal polygon Γ is smaller than n .

By Lemma 2.3 we get:

Proposition 2.6. For any \vec{v} , $\overline{\mathbf{G}}_{\sigma, n}(\cdot, \vec{v})$ has a (possibly nonunique) global minimizer.

We next investigate properties of the minimizer of $\mathbf{G}_{\sigma, n}^\tau(\cdot, \vec{v})$ for prescribed τ and \vec{v} . By Definition 2.4 we may take τ to be the identity permutation and replace $\mathbf{G}_{\sigma, n}^\tau$ by $\mathbf{G}_{\sigma, n}$. Let \vec{u}^0 be a minimizer of $\mathbf{G}_{\sigma, n}(\vec{u}, \vec{v})$ of the form

$$\vec{u}^0 = (\underbrace{u_1^0, \dots, u_1^0}_{m_1}, \dots, \underbrace{u_k^0, \dots, u_k^0}_{m_k}), \quad \text{with } u_i^0 \neq u_j^0 \text{ for } i \neq j \tag{7}$$

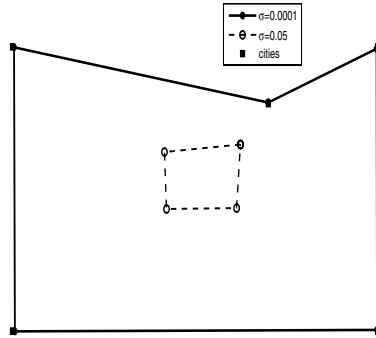


FIGURE 2. The orbit of the LTSP for small σ (solid line) and larger σ (dashed line). Note that the number of vertices reduced from 5 points for small σ to 4 points for larger σ .

with $\sum_{j=1}^k m_j = n$. Define $M_l = m_1 + \dots + m_{l-1}$, ($m_0 = 0$) for $l = 1, \dots, k$. We cluster the set $V := \{v_1, \dots, v_n\}$ into k disjoint subsets as follows:

$$V_l := \{v_{M_l+1}, \dots, v_{M_{l+1}}\}; l = 1, \dots, k.$$

Define also

$$\vec{u}^1 := (u_1^0, \dots, u_k^0) \in \mathbb{C}^k.$$

Lemma 2.7. *A minimizer \vec{u}^0 of $\mathbf{G}_{\sigma,n}$ of the form (7) satisfies, for $l = 1, \dots, k$:*

$$u_l^0 - \frac{1}{m_l} \sum_{j=1}^{m_l} v_{l,j} = \frac{n}{m_l} \sigma \mathbf{L}_n(\vec{u}^0)(e_l - e_{l+1}) \tag{8}$$

where

$$e_l = \frac{u_l^0 - u_{l-1}^0}{|u_l^0 - u_{l-1}^0|}. \tag{9}$$

The geometrical meaning is that the center of mass of the points in V_l lies on the bisection of the normal cone at u_l (see Fig. 1).

Proof. Let

$$\mathbf{h}^\sigma(\vec{u}) = \frac{1}{n} \sum_{l=1}^k \sum_{v \in V_l} |u_l - v|^2 + \sigma \mathbf{L}_k^2(\vec{u}) \tag{10}$$

be a function $\mathbf{h}^\sigma : \mathbb{C}^k \times \mathbb{C}^n \rightarrow \mathbb{R}^+$. Evidently, $\mathbf{h}^\sigma(\vec{u}^1) = \mathbf{G}_{\sigma,n}(\vec{u}^0)$ and \vec{u}^1 is a minimizer of \mathbf{h}^σ on \mathbb{C}^k if and only if \vec{u}^0 is a minimizer of $\mathbf{G}_{\sigma,n}$ on \mathbb{C}^n . Note that, in contrast to $\mathbf{G}_{\sigma,n}$ at \vec{u}^0 , \mathbf{h}^σ is differentiable at \vec{u}^1 . The equality (8) follows by equating its derivative to zero. \square

We immediately get

Corollary 2.8. *The center of mass of a minimizer \vec{u}_0 of $\mathbf{G}_{\sigma,n}$ coincides with the center of mass of \vec{v} , namely*

$$u_{0,cm} := \frac{1}{n} \sum_{i=1}^n u_i^0 = \frac{1}{n} \sum_{i=1}^n v_i := v_{cm}.$$

3. DUAL PROBLEM FOR THE TSP

The inner product in \mathbb{C}^n is defined as

$$\langle \vec{u}, \vec{v} \rangle := \sum_1^n \Re_e(v_i \bar{u}_i),$$

with \bar{u} stands for the complex conjugate of $u \in \mathbb{C}$.

We recall some basic definitions:

Let $f : \mathbb{C}^n \rightarrow \mathbb{R}$. The *subdifferential* of f at \vec{u} is given by

$$\partial_{\vec{u}} f := \{ \vec{s} \in \mathbb{C}^n; f(\vec{y}) \geq f(\vec{u}) + \langle \vec{s}, \vec{y} - \vec{u} \rangle, \forall \vec{y} \in \mathbb{C}^n \}.$$

Below we list some properties of ∂f :

Lemma 3.1.

- (1) If f is a convex function on \mathbb{C}^n then $\partial_{\vec{u}} f \neq \emptyset$ for any $\vec{u} \in \mathbb{C}^n$.
- (2) If f is differentiable at $\vec{u} \in \mathbb{C}^n$ then either $\partial_{\vec{u}} f = \emptyset$ or $\partial_{\vec{u}} f$ is a singleton, given by the gradient $\{\nabla_{\vec{u}} f\}$ of f at \vec{u} .
- (3) If f is homogeneous of degree k then $f(\vec{u}) = k^{-1} \langle \vec{u}, \vec{y} \rangle \quad \forall \vec{y} \in \partial_{\vec{u}} f$.
- (4) If $g(\vec{y}) \geq f(\vec{y}) \quad \forall \vec{y} \in \mathbb{C}^n$ and $g(\vec{u}) = f(\vec{u})$, then $\partial_{\vec{u}} f \subset \partial_{\vec{u}} g$.

Recall the definition of the length function \mathbf{L}_n (notation (v)). The following lemma follows from Lemma 3.1.

Lemma 3.2. For $\vec{u} = (u_1, \dots, u_n) \in \mathbb{C}^n$, set $u_0 = u_n, u_1 = u_{n+1}$. For $i = 1, \dots, n+1$ define $e_i = \frac{u_i - u_{i-1}}{|u_i - u_{i-1}|} \in \mathbb{C}$ if $u_i \neq u_{i-1}$, while $|e_i| \leq 1$, if $u_i = u_{i-1}$. Then

$$\partial_{\vec{u}} \mathbf{L}_n = \{ \vec{s} \in \mathbb{C}^n; \vec{s} = (s_1, \dots, s_n) \text{ where } s_i = e_i - e_{i+1} \}.$$

Let

$$W_0 = \left\{ \vec{u} = (u_1, \dots, u_n) \in \mathbb{C}^n, u_i \in \mathbb{C}, \sum_1^n u_i = 0 \right\}.$$

We obtain, in particular,

Corollary 3.3. If $\vec{s} \in \partial_{\vec{u}} \mathbf{L}_n$, then $\vec{s} \in W_0$. If $u_j \neq u_i$ for $i \neq j$ and $n > 1$, then $\partial_{\vec{u}} \mathbf{L}_n$ is a singleton, that is, the gradient of \mathbf{L}_n exists at \vec{u} and $\partial_{\vec{u}} \mathbf{L}_n = \{\nabla_{\vec{u}} \mathbf{L}_n\}$.

We can identify, therefore, $\vec{s} \in \partial_{\vec{u}} \mathbf{L}_n$ with a closed polygon $\mathbf{p}_{\vec{s}}$ whose vertices are given in terms of the corresponding ordered sequence $\vec{s} = \vec{\delta}e$. Here, $s_i = e_{i+1} - e_i$ where e_i are given, in terms of \vec{u} , by (9) (see notation (iv)).

For any such polygon \mathbf{p} , consider the minimal circle which bound \mathbf{p} (M.B.C.). Let $\rho(\mathbf{p})$ be the radius of this circle.

Definition 3.4. The function $R : \mathbb{C}^n \rightarrow \mathbb{R}^+ \cup \{\infty\}$ is

$$R(\vec{s}) = \begin{cases} \rho(\mathbf{p}_{\vec{s}}); & \text{if } \vec{s} \in W_0 \\ +\infty; & \text{if } \vec{s} \notin W_0. \end{cases}$$

Proposition 3.5. Let $\vec{s} \in \partial_{\vec{u}} \mathbf{L}_n$ where $\vec{u} = \{u_1, \dots, u_n\}$. Assume $u_i \neq u_j$ for some $i, j \in \{1, \dots, n\}$. Then $R(\vec{s}) = 1$.

Proof. First note that $\vec{s} \in W_0$ hence $R(\vec{s}) < \infty$. By definition of $\mathbf{p}_{\vec{s}}$, the vertices of this polygon are given by the vectors e_i , which are contained in the unit disk. Hence the unit circle contains this polygon. We now show that this is the minimal circle. Consider the set $I := \{i_1, \dots, i_m\} \subset \{1, \dots, n\}$, $i_{m+1} \equiv i_1$, such that $|e_{i_k}| = 1$. All we have to show now is that any π -arc of the unit circle contains at least one of the vectors e_{i_k} . Assume the contrary. Then there exists $l \in \mathbb{C}$ for which $\langle l, e_{i_k} \rangle > 0$ for $k = \{1, \dots, m\}$. Hence $\langle u_{i_k} - u_{i_{k-1}}, l \rangle > 0$. It follows that $\sum_{k=1}^m \langle u_{i_{k+1}} - u_{i_k}, l \rangle > 0$. However, $\sum_{k=1}^m (u_{i_{k+1}} - u_{i_k}) = \sum_{k=1}^n (u_{k+1} - u_k) = 0$. \square

Corollary 3.6. *If $\vec{s} \in \partial_{\vec{u}}[\mathbf{L}_n^2/2] = \mathbf{L}_n(\vec{u})\partial_{\vec{u}}\mathbf{L}_n$ then $\mathbf{L}_n(\vec{u}) = R(\vec{s})$.*

Using the 1-homogeneity of L with Lemma 3.1-(3) we also have:

$$\langle \vec{u}, \vec{s} \rangle = \mathbf{L}_n(\vec{u}) \quad \forall \vec{s} \in \partial_{\vec{u}}\mathbf{L}_n. \tag{11}$$

Similarly, $\mathbf{L}_n^2/2$ is 2-homogeneous, hence

$$\langle \vec{u}, \vec{s} \rangle = \mathbf{L}_n^2(\vec{u}) \quad \forall \vec{s} \in \partial_{\vec{u}}\mathbf{L}_n^2/2. \tag{12}$$

Let us recall the definition of the Legendre-Fenchel transform f^* of a function $f : \mathbb{C}^n \rightarrow \mathbb{R}$:

$$f^*(\vec{s}) = \sup \{ \langle \vec{s}, \vec{v} \rangle - f(\vec{v}) ; \vec{v} \in \mathbb{C}^n \}. \tag{13}$$

Recall also the *convex hull* of a function $f : \mathbb{C}^n \rightarrow \mathbb{R}$:

$$f^c(\vec{s}) = \sup_h \{ h(\vec{v}) ; h \text{ convex and } h \leq f \text{ on } \mathbb{C}^n \}. \tag{14}$$

We summarize below the main properties of f^* . Most of the proofs can be found in [6], Chapter X.

Lemma 3.7. *Suppose f is bounded from below by some affine function. Then*

- (1) f^* is a convex function;
- (2) $f^{**} = f^c$;
- (3) $\partial_{\vec{v}}f \neq \emptyset \iff f^c(\vec{v}) = f(\vec{v})$;
- (4) $\vec{s} \in \partial_{\vec{v}}f \implies \vec{v} \in \partial_{\vec{s}}f^*$;
- (5) *Given $\vec{v} \in \mathbb{C}^n$, if there exists \vec{s} such that $f(\vec{v}) + f^*(\vec{s}) = \langle \vec{v}, \vec{s} \rangle$ then $f(\vec{v}) = f^c(\vec{v}) = f^{**}(\vec{v})$. Moreover, $\vec{v} \in \partial_{\vec{s}}f^*$ and $\vec{s} \in \partial_{\vec{v}}f$.*

Lemma 3.8. *The Legendre-Fenchel transform of $\mathbf{L}_n^2/2$ is given by*

$$\left[\frac{1}{2}\mathbf{L}_n^2 \right]^*(\vec{s}) = \left[\frac{1}{2}\mathbf{L}_n^2 \right](\vec{u}) = \frac{1}{2}R^2(\vec{s})$$

for any $\vec{u} \in \partial_{\vec{s}}[R^2/2]$.

Proof. First we note that $\partial_{\vec{u}}\mathbf{L}_n \subset W_0$. By (12), the convexity of \mathbf{L}_n^2 and Lemma 3.7-(3) we obtain, for $\vec{s} \in \partial_{\vec{u}}[\mathbf{L}_n^2/2]$,

$$[\mathbf{L}_n^2/2]^*(\vec{s}) = \langle \vec{s}, \vec{u} \rangle - \mathbf{L}_n^2(\vec{u})/2 = \mathbf{L}_n^2(\vec{u})/2.$$

By Corollary 3.6 we get the desired result for $s \in W_0$.

Next we observe that $\mathbf{L}_n(\vec{w}) = 0$ for any $\vec{w} \in W_0^\perp$. If $\vec{s} \notin W_0$ let $\vec{w} \in W_0^\perp$ for which $\langle \vec{s}, \vec{w} \rangle = \delta > 0$. Then, by (13),

$$\left[\frac{1}{2}\mathbf{L}_n^2 \right]^*(\vec{s}) \geq \langle \vec{s}, \alpha\vec{w} \rangle - \frac{1}{2}\mathbf{L}_n^2(\alpha\vec{w}) = \alpha\langle \vec{s}, \vec{w} \rangle = \alpha\delta$$

for any $\alpha > 0$. Letting $\alpha \rightarrow \infty$ we obtain $[\mathbf{L}_n^2]^*(\vec{s}) = \infty$ for $\vec{s} \notin W_0$. \square

Next define

$$T(\vec{u}) = \min (\mathbf{L}_n^2(\vec{u}_\tau)/2 ; \tau \in S_n).$$

Evidently, T is not, in general, a convex function on \mathbb{C}^n , since a minimum of a family of convex functions is not necessarily convex. However, T^* is defined and convex (by Lem. 3.7 (1)). In Lemma 3.9 below we state that $T^*(\vec{s})$ can be interpreted in terms of the *maximal* radius of the minimal circles which bound all polygons obtained by permutations of the components of \vec{s} .

Lemma 3.9. *The Legendre-Fenchel transform of T is a convex function, given by*

$$T^*(\vec{s}) = \frac{1}{2} \max_{\tau \in S_n} R^2(\vec{s}_\tau).$$

Proof. By definition:

$$T^*(\vec{s}) = \sup_{\vec{u}} \left\{ \langle \vec{s}, \vec{u} \rangle - \min_{\tau} [\mathbf{L}_n^2(\vec{u}_\tau)/2] \right\} = \max_{\tau \in S_n} \sup_{\vec{u}} \left\{ \langle \vec{s}, \vec{u} \rangle - [\mathbf{L}_n^2(\vec{u}_\tau)/2] \right\}.$$

Now we use $\langle \vec{s}, \vec{u} \rangle = \langle \vec{s}_\tau, \vec{u}_\tau \rangle$ for any $\tau \in S_n$ to obtain:

$$\begin{aligned} T^*(\vec{s}) &= \max_{\tau \in S_n} \sup_{\vec{u}} \left\{ \langle \vec{s}_\tau, \vec{u}_\tau \rangle - [\mathbf{L}_n^2(\vec{u}_\tau)/2] \right\} = \max_{\tau \in S_n} \sup_{\vec{u}} \left\{ \langle \vec{s}_\tau, \vec{u} \rangle - [\mathbf{L}_n^2(\vec{u})/2] \right\} \\ &= \max_{\tau} \left[\frac{1}{2} \mathbf{L}_n^2 \right]^* (\vec{s}_\tau) \end{aligned}$$

and the result follows from Lemma 3.8. □

We turn now to the dual representation of the TSP. Let T^c be the convex hull of T . Note that $T^c(\vec{u}) \leq T(\vec{u})$ for any $\vec{u} \in \mathbb{R}^n$. Define:

$$D_0 = \{ \vec{u} \in \mathbb{C}^n ; T^c(\vec{u}) = T(\vec{u}) \}. \tag{15}$$

Next, define

$$\mathbf{H}(\vec{s}, \vec{v}) = \langle \vec{s}, \vec{v} \rangle - T^*(\vec{s}). \tag{16}$$

Evidently, \mathbf{H} is a concave and coercive function of the first argument. Hence, for any fixed $\vec{v} \in \mathbb{C}^n$, a maximizer $\vec{s} \in W_0$ of \mathbf{H} exists.

Definition 3.10. Given $\vec{v} \in \mathbb{C}^n$, a permutation $\tau \in S_n$ is *ordered well* for \vec{v} if $T(\vec{v}) = \mathbf{L}_n^2(\vec{v}_\tau)/2$. The vector $\vec{v} \in \mathbb{C}^n$ is *ordered* if $T(\vec{v}) = \mathbf{L}_n^2(\vec{v})/2$.

Similarly, τ is a *cyclic order* for $\vec{s} \in W_0$ if $T^*(\vec{s}) = R^2(\vec{s}_\tau)/2$, and $\vec{s} \in W_0$ is **-ordered* if $R^2(\vec{s})/2 = T^*(\vec{s})$.

From Lemma 3.1-(4), the convexity of \mathbf{L}_n and Definition 3.10 we obtain

Lemma 3.11. *If \vec{v} is ordered then $\partial_{\vec{v}}T \subset \partial_{\vec{v}}\mathbf{L}_n^2/2$.*

Recall that $\partial_{\vec{v}}T$ can be an empty set (even though $\partial_{\vec{v}}\mathbf{L}_n$ is never empty). By Lemmas 3.11 and 3.7 we obtain

Corollary 3.12. *$\vec{v} \in D_0$ if and only if $\partial_{\vec{v}}T \neq \emptyset$. In this case, the following claims are equivalent:*

- (i) $\vec{s} \in \partial_{\vec{v}}T$;
- (ii) $\vec{v} \in \partial_{\vec{s}}T^*$;
- (iii) $T(\vec{v}) + T^*(\vec{s}) = \langle \vec{s}, \vec{v} \rangle$;
- (iv) \vec{s} is a maximizer of $\mathbf{H}(\cdot, \vec{v})$ on W_0 ;
- (v) $\mathbf{H}(\vec{s}, \vec{v}) = T(\vec{v})$.

The main theorems of this section are:

Theorem 3.13. *If $\vec{v} \in D_0$ is ordered then there exists $\vec{s} \in \partial_{\vec{v}}T$ which is *-ordered.*

An interesting conclusion is the following:

Theorem 3.14. *Let $\vec{v} = (v_1, \dots, v_n) \in D_0$ satisfy $v_i \neq v_j$ for $i \neq j$. Then there is a unique maximizer \vec{s}_0 of $\mathbf{H}(\cdot, \vec{v})$ (16). Moreover, if τ is ordered well for \vec{v} then it is also a cyclic order for \vec{s}_0 .*

The next theorem provides us with a natural lower bound for the optimal solution of the TSP:

Theorem 3.15. *Given $\vec{v} \in \mathbb{C}^n$, then for any $\vec{s} \in W_0$:*

$$T(\vec{v}) \geq \frac{\langle \vec{s}, \vec{v} \rangle^2}{4T^*(\vec{s})}. \tag{17}$$

Moreover, this is a sharp estimate if $\vec{v} \in D_0$.

Proof of Theorem 3.13. By definition of the Legendre-Fenchel transform, Lemma 3.7-(2), the convexity of \mathbf{L}_n^2 and Lemma 3.8, for any $\vec{s} \in W_0$:

$$\langle \vec{s}t, \vec{v} \rangle - R^2(\vec{s}t)/2 \leq [\mathbf{L}_n^2/2]^{**}(\vec{v}) = \mathbf{L}_n^2(\vec{v})/2.$$

Since $T^* \geq R^2/2$ by definition, it follows that

$$\langle \vec{s}t, \vec{v} \rangle - R^2(\vec{s}t)/2 \geq \mathbf{H}(\vec{s}t, \vec{v}), \tag{18}$$

so

$$\mathbf{L}_n^2(\vec{v})/2 \geq \mathbf{H}(\vec{s}t, \vec{v}) \tag{19}$$

holds for any $\vec{s}t \in W_0$. On the other hand, $\vec{s} \in \partial_{\vec{v}}T$ implies, via Corollary 3.12, that the right hand side of (19) is maximized at $\vec{s}t = \vec{s}$, and $\mathbf{H}(\vec{s}, \vec{v}) = T^{**}(\vec{v})$. Since $\vec{v} \in D_0$ then $T^{**}(\vec{v}) \equiv T^c(\vec{v}) = T(\vec{v})$. Since \vec{v} is ordered, $\mathbf{L}_n^2(\vec{v})/2 = T(\vec{v})$. As a result, there is an equality in (19) for $\vec{s}t = \vec{s}$. In particular, it follows that there must be an equality in (18) for $\vec{s}t = \vec{s}$ as well, so

$$\langle \vec{s}, \vec{v} \rangle - R^2(\vec{s})/2 = \mathbf{H}(\vec{s}, \vec{v}),$$

so $T^*(\vec{s}) = R^2(\vec{s})/2$ by (16), which implies that \vec{s} is $*$ -ordered by definition. □

Proof of Theorem 3.14. By definition, $T^{**} \leq T \leq \mathbf{L}_n^2/2$. Assume first \vec{v} is ordered. Then, by assumptions, $T^{**}(\vec{v}) = T(\vec{v}) = \mathbf{L}_n^2(\vec{v})/2$. By Lemma 3.1-(4) it follows that $\partial_{\vec{v}}T^{**} \subset \partial_{\vec{v}}T \subset \partial_{\vec{v}}\mathbf{L}_n^2/2$. However, by Lemma 3.7 and convexity of T^{**} it follows that $\partial_{\vec{v}}T^{**} \neq \emptyset$. In addition, by Corollary 3.3, $\partial_{\vec{v}}\mathbf{L}_n^2/2 = \{\nabla_{\vec{v}}\mathbf{L}_n^2/2\}$ is a singleton. Hence $\partial_{\vec{v}}T^{**}$ is singleton as well. It follows that the maximizer $\vec{s}_0 \in \partial_{\vec{v}}T^{**}$ of $\mathbf{H}(\cdot, \vec{v})$ is singleton as well.

Now, if \vec{v} is not ordered, let τ be an ordered well permutation for \vec{v} , so \vec{v}_τ is ordered by definition. Since $\mathbf{H}(\vec{s}, \vec{v}_\tau) = \mathbf{H}(\vec{s}_\tau, \vec{v})$, it follows that \vec{s}_0 is a maximizer of $\mathbf{H}(\cdot, \vec{v}_\tau)$ if and only if $\vec{s}_{0,\tau}$ is a maximizer of $\mathbf{H}(\cdot, \vec{v})$. By Theorem 3.13 it follows that $\vec{s}_{0,\tau}$ is $*$ -ordered, hence τ is a cyclic order. □

Proof of Theorem 3.15. First, note that

$$T(\vec{v}) \geq T^{**}(\vec{v}) \geq \mathbf{H}(\vec{s}, \vec{v})$$

for any $\vec{s} \in W_0$. So

$$T(\vec{v}) \geq \max_{\lambda \in \mathbb{R}} \mathbf{H}(\lambda \vec{s}, \vec{v}) = \max_{\lambda \in \mathbb{R}} [\lambda \langle \vec{v}, \vec{s} \rangle - \lambda^2 T^*(\vec{s})] = \frac{\langle \vec{s}, \vec{v} \rangle^2}{4T^*(\vec{s})}.$$

Finally, we already know that the equality is achieved for $\vec{s} \in \partial_{\vec{v}}T$, if $\vec{v} \in D_0$. □

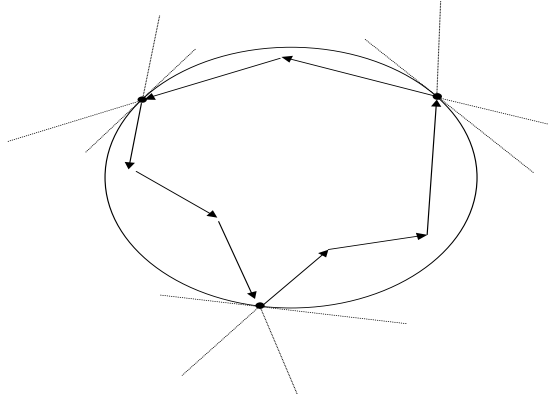


FIGURE 3. The minimal circle of the polygon in the figure is increased when reordering the edges and getting a convex polygon.

The significance of Theorems 3.14 and 3.15 is originated from the fact that the computation of T^* , in contrast to the computation of T , is efficient.

To compute $T^*(\vec{s})$ it is necessary to compute the radius of the minimal bounding circle (M.B.C) $R(\vec{s}_\tau)$ for all the polygons obtained by permutations $\tau \in S_n$ of the edges (v_1, \dots, v_n) . It is known (see [8]) that, for a given n -gon in the plane, the radius of its M.B.C. can be computed by $O(n \log n)$ calculations. Recently, it was improved to $O(n)$ [11]. So, the apparent difficulty in the computation of T^* seems to be originated from the necessity to perform all permutations, which is, of course, an exponentially large set.

However, the computational cost of T^* is only linear! This is the result of the following:

Proposition 3.16. *Given $\vec{s} = (s_1, \dots, s_n) \in W_0$, there exists a unique convex n -gon whose edges are composed of the coordinates s_j of \vec{s} . The maximal bounding circle corresponding to this polygon is the maximal possible one within all rearrangements of the coordinates of \vec{s} .*

There is a simple way to construct this convex polygon. Let $\vec{s} = (|s_1|e^{i\theta_1}, \dots, |s_n|e^{i\theta_n})$. Then all we have to do is just to permute the coordinates $i \rightarrow \tau(i)$ so that $0 \leq \theta_{\tau(i)} \leq \theta_{\tau(i+1)} < 2\pi, i = 1, \dots, n - 1$. The polygon corresponding to \vec{s}_τ is the required convex one.

It should be emphasized, however, that the above convex polygon is not, in general, the only one corresponding to a maximal M.B.R (see Fig. 3 below). Nevertheless, since $v_i \neq v_j$ for all components of \vec{v} , it follows that the maximizer \vec{s} of $\mathbf{H}(\cdot, \vec{v})$ is the gradient of $\mathbf{L}_n^2/2$ at the ordered well rearrangement of \vec{v} , and all the vertices of the corresponding polygon are on a circle by Lemma 3.2.

Proof. ³ Let \mathbf{p} be the optimal polygon (of maximal radius of the minimal bounding circle C). Let $p_1, \dots, p_k, k \leq n$, be the vertices of \mathbf{p} on C . Let D be the closed disk bounded by C . Evidently, D is also the minimal disk which contains the points p_1, \dots, p_k . Moreover, if we add a point $p_{k+1} \notin D$ then the minimal disk containing p_1, \dots, p_{k+1} is of radius larger than D .

For any $i \in \{1, \dots, k - 1\}$, let V_i be the set of indices j such that

$$p_{i+1} - p_i = \sum_{j \in V_i} s_j.$$

Evidently, $\cup_{i=1}^k V_i = \{1, \dots, n\}$ and $V_i \cap V_j = \emptyset$ for $i \neq j$. We note two facts:

- (i) A permutation within each of the sets V_i preserves the points p_1, \dots, p_k . It follows that such permutations do not decrease the radius of the bounding circle C . Since C is the maximal bounding circle

³We are thankful to Prof. A. Leizerowitz for helping us with this proof.

by assumption, then any such permutation leaves C invariant. In particular, we can arrange the indices within each V_i such that the corresponding directions of $s_j, j \in V_i$ are in cyclic order without changing C .

- (ii) If the directions of $s_j, j \in V_i$ are so ordered, then the argument of the first vector $s \in V_{i+1}$ is larger than the argument of the last vector $s' \in V_i$. For otherwise we have that $p_i + s \notin D$, hence C is not the maximal bounding circle, contradicting the assumption.

From the above we see that the optimal arrangement of s_i can be made in cyclic order. It is easy to see that such an arrangement yields, necessarily, a convex polygon (in fact, the only possible one). \square

4. DUAL REPRESENTATION OF THE LTSP

We turn now to the LTSP. First, recall the definition of ϵ – Moreau-Yosida regularization [6], Chapter X of a convex function f on, say, \mathbb{C}^n :

$$f_\epsilon(\vec{v}) := \inf_{\vec{u}} \left[\frac{1}{2\epsilon} \|\vec{u} - \vec{v}\|^2 + f(\vec{u}) \right].$$

Below we list some of the properties of f_ϵ :

Lemma 4.1.

- (i) *If f is convex then f_ϵ is convex and differentiable with a continuous derivative.*
- (ii) *Let $\vec{u}(\vec{v})$ given by $f_\epsilon(\vec{u}) = \frac{1}{2\epsilon} \|\vec{u}(\vec{v}) - \vec{v}\|^2 + f(\vec{u}(\vec{v}))$. Then $\nabla f_\epsilon(\vec{v}) = \epsilon^{-1}(\vec{v} - \vec{u}) \in \partial_{\vec{u}} f$.*
- (iii) *The Legendre-Fenchel transform of f_ϵ is $f_\epsilon^*(\vec{s}) = \frac{\epsilon}{2} \|\vec{s}\|^2 + f^*(\vec{s})$.*

Recall now Definitions 2.2 and 2.4. For convenience we divide $\mathbf{G}_{\sigma,n}^\tau$ by 2σ :

$$(2\sigma)^{-1} \mathbf{G}_{\sigma,n}(\vec{u}, \vec{v}) := \frac{1}{2\sigma n} \|\vec{u} - \vec{v}\|^2 + \frac{1}{2} \mathbf{L}_n^2(\vec{u}).$$

Then

$$\Xi^\sigma(\vec{v}) := (2\sigma)^{-1} \inf_{\vec{u} \in \mathbb{C}^n} \mathbf{G}_{\sigma,n}(\vec{u}, \vec{v})$$

is, in fact, the Moreau-Yosida regularization of $\mathbf{L}_n^2(\cdot)/2$ with respect to the parameter $\epsilon = n\sigma$. From Lemmas 3.8 and 4.1-(iii) we have:

$$\Xi^{\sigma,*}(\vec{s}) = \frac{n\sigma}{2} \|\vec{s}\|^2 + \frac{1}{2} R^2(\vec{s}). \tag{20}$$

Likewise, by Definition 2.4,

$$\Xi_\tau^\sigma(\vec{v}) := (2\sigma)^{-1} \inf_{\vec{u} \in \mathbb{C}^n} \mathbf{G}_{\sigma,n}^\tau(\vec{u}, \vec{v}); \quad \Xi_\tau^{\sigma,*}(\vec{s}) = \sup_{\vec{v} \in \mathbb{C}^n} [\langle \vec{s}, \vec{v} \rangle - \Xi_\tau^\sigma(\vec{v})], \tag{21}$$

and

$$\bar{\Xi}^\sigma(\vec{v}) := \min_{\tau \in S_n} \Xi_\tau^\sigma(\vec{v}); \quad \bar{\Xi}^{\sigma,*}(\vec{s}) = \sup_{\vec{v} \in \mathbb{C}^n} [\langle \vec{s}, \vec{v} \rangle - \bar{\Xi}^\sigma(\vec{v})] = \max_{\tau \in S_n} \Xi_\tau^{\sigma,*}(\vec{s}). \tag{22}$$

Note that $\bar{\Xi}^\sigma(\vec{v})$ is also the Moreau-Yosida regularization of T , given by

$$\bar{\Xi}^\sigma(\vec{v}) = (2\sigma)^{-1} \min_{\vec{u} \in \mathbb{C}^n} \bar{\mathbf{G}}_{\sigma,n}(\vec{u}, \vec{v}). \tag{23}$$

By Lemma 3.9, (22) and (20):

Lemma 4.2. $\bar{\Xi}^{\sigma,*}(\vec{s}) = \frac{n\sigma}{2} \|\vec{s}\|^2 + T^*(\vec{s})$ is a strictly convex function on \mathbb{C}^n .

The following definition is analogous to (part of) Definition 3.10:

Definition 4.3. $\vec{v} \in \mathbb{C}^n$ is called σ -ordered if $\Xi^\sigma(\vec{v}) = \overline{\Xi}^\sigma(\vec{v})$. That is, $\Xi^\sigma(\vec{v}) \leq \Xi_\tau^\sigma(\vec{v})$ for all $\tau \in S_n$.

We now describe the dual version of the LTSP. Let

$$D_\sigma = \left\{ \vec{u} \in \mathbb{C}^n; \overline{\Xi}^{\sigma,c}(\vec{u}) = \overline{\Xi}^\sigma(\vec{u}) \right\}.$$

and

$$\mathbf{H}^\sigma(\vec{s}, \vec{v}) = \langle \vec{s}, \vec{v} \rangle - \left[\frac{n\sigma}{2} |\vec{s}|^2 + T^*(\vec{s}) \right]. \tag{24}$$

Note that this functional is strictly concave with respect to $\vec{s} \in \mathbb{C}^n$ for any $\vec{v} \in \mathbb{C}^n$. Moreover, $\mathbf{H}^\sigma > -\infty$ if $\vec{s} \in W_0$. It follows that the problem

$$\max_{\vec{s} \in W_0} \mathbf{H}^\sigma(\vec{s}, \vec{v}) = \overline{\Xi}^{\sigma,**}(\vec{v}) \tag{25}$$

admits a *unique* solution $\vec{s}_0 \in W_0$ for any $\vec{v} \in \mathbb{C}^n$. The analog of Theorem 3.13 is

Theorem 4.4. Assume $\vec{v} \in D_\sigma$ is σ -ordered, then the unique maximizer \vec{s} of $\mathbf{H}^\sigma(\cdot, \vec{v})$ is $*$ -ordered, and the (unique) minimizer of $\overline{\mathbf{G}}_{\sigma,n}(\cdot, \vec{v})$ is given by

$$(u_1, \dots, u_n) = \vec{u} := \vec{v} - \sigma n \vec{s}. \tag{26}$$

Proof. Following the line of the proof of Theorem 3.13 we have

$$\langle \vec{s}l, \vec{v} \rangle - \Xi^{\sigma,*}(\vec{s}l) \leq \Xi^{\sigma,**}(\vec{v}) = \Xi^\sigma(\vec{v}) = \overline{\Xi}^\sigma(\vec{v}) = \overline{\Xi}^{\sigma,**}(\vec{v}) \tag{27}$$

where the first equality follows from convexity of Ξ^σ , the second from the σ -order of \vec{v} and the third one from the assumption $\vec{v} \in D_\sigma$. In addition, it follows from (22)

$$\langle \vec{v}, \vec{s}l \rangle - \Xi^{\sigma,*}(\vec{s}l) \geq \mathbf{H}^\sigma(\vec{s}l, \vec{v}).$$

Hence,

$$\mathbf{H}^\sigma(\vec{s}l, \vec{v}) \leq \overline{\Xi}^{\sigma,**}(\vec{v}).$$

Again, we know that the equality above holds for an unique \vec{s} . Hence, the equality holds for (27) where $\vec{s}l = \vec{s}$ as well. In particular

$$\langle \vec{s}, \vec{v} \rangle - \Xi^{\sigma,*}(\vec{s}) = \langle \vec{s}, \vec{v} \rangle - \left(\frac{n\sigma}{2} \|\vec{s}\|^2 + \frac{1}{2} R^2(\vec{s}) \right) = \overline{\Xi}^{\sigma,**}(\vec{v}) = \langle \vec{s}, \vec{v} \rangle - \left(\frac{n\sigma}{2} \|\vec{s}\|^2 + T^*(\vec{s}) \right)$$

where the first equality follows from (20), the second from the equality in (27) with $\vec{s}l = \vec{s}$ and the third equality from Lemma 4.2. Thus, $\frac{1}{2} R^2(\vec{s}) = T^*(\vec{s})$ and \vec{s} is $*$ -ordered.

Now, Lemma 4.1-(ii), with $f = \mathbf{L}_n^2/2$, $f_\varepsilon = \Xi^\sigma$ and $\varepsilon = n\sigma$ implies that

$$\vec{v} = \sigma n \nabla_{\vec{v}} \overline{\Xi}^\sigma + \vec{u}, \tag{28}$$

where \vec{u} is the minimizer of $\overline{\mathbf{G}}_{\sigma,n}(\cdot, \vec{v})$ (see (23)). But $\vec{s} = \nabla_{\vec{v}} \Xi^\sigma = \nabla_{\vec{v}} \overline{\Xi}^\sigma$, so our claim follows. □

Evidently, the condition that \vec{v} is σ -ordered is not known *a priori*. However, if $\vec{v} \in \mathbb{C}^n$ is given, and $\tau \in S_n$ a permutation such that \vec{v}_τ is σ -ordered, then, by definition,

$$\sup_{\vec{s}l \in \mathbb{C}^n} \mathbf{H}(\vec{s}l, \vec{v}) = \sup_{\vec{s}l \in \mathbb{C}^n} \mathbf{H}(\vec{s}l, \vec{v}_\tau)$$

and the maximizer of the right hand side is given by \vec{s}_τ , if \vec{s} is the maximizer of the left hand side. We obtain the analog of Theorem 3.14:

Theorem 4.5. *Let $\vec{v} = (v_1, \dots, v_n) \in D_\sigma$. Let \vec{s} be the (unique) maximizer of the concave function $\mathbf{H}^\sigma(\vec{s}, \vec{v})$ (24). Then there exists a cyclic order τ on \vec{s} such that the minimizer of $\overline{\mathbf{G}}_{\sigma,n}$ is given by (26) where \vec{v} and \vec{s} are replaced by \vec{v}_τ and \vec{s}_τ .*

The analog of Theorem 3.15 also holds for a lower estimate on the minimum value of the LTSP. It is proved in, essentially, the same way:

Theorem 4.6. *Given $\vec{v} \in \mathbb{C}^n$, then for any $\vec{s} \in W_0$:*

$$\overline{\Xi}^\sigma(\vec{v}) \geq \frac{\langle \vec{s}, \vec{v} \rangle^2}{4T^*(\vec{s}) + 2\sigma n \|\vec{s}\|^2}.$$

Moreover, this is a sharp estimate if $\vec{v} \in D_\sigma$.

5. OPTIMALITY CONDITIONS

In this section we investigate sufficient conditions for $\vec{v} \in \mathbb{C}^n$ to be in D_σ . The first result is:

Lemma 5.1. *For $\sigma > 0$, let \vec{u} be a minimizer of $\overline{\mathbf{G}}_{\sigma,n}(\cdot, \vec{v})$. Suppose its components are the vertices of a convex polygon in \mathbb{C} . Then $\vec{v} \in D_\sigma$.*

Proof. From Lemma 3.7-(5), it is enough to prove the existence of such \vec{s} for which

$$\overline{\Xi}^\sigma(\vec{v}) + \overline{\Xi}^{\sigma,*}(\vec{s}) = \langle \vec{v}, \vec{s} \rangle. \tag{29}$$

We may assume, without limitation to generality, that \vec{v} is σ -ordered. Indeed, recall that $\vec{v} \in D_\sigma$ if and only if $\vec{v}_\tau \in D_\sigma$ for all $\tau \in S_n$, and we can always permute the components of \vec{v} to make it σ -ordered. Then $\overline{\Xi}^\sigma(\vec{v}) = \overline{\Xi}^\sigma(\vec{v})$. Since $\overline{\Xi}^\sigma$ is both convex and continuously differentiable (see Lem. 4.1-i), then

$$\overline{\Xi}^\sigma(\vec{v}) + \overline{\Xi}^{\sigma,*}(\vec{s}) = \langle \vec{v}, \vec{s} \rangle \tag{30}$$

holds for $\vec{s} = \nabla_{\vec{v}} \overline{\Xi}$. However, Lemma 4.1-ii with $f = \mathbf{L}_n^2/2$ and $\varepsilon = \sigma n$ implies, in addition, that

$$\vec{s} = \nabla_{\vec{v}} \overline{\Xi}^\sigma = \frac{\vec{v} - \vec{u}}{n\sigma} \in \partial_{\vec{u}}(\mathbf{L}_n^2/2) \tag{31}$$

where \vec{u} is the minimizer of $\mathbf{G}_{\sigma,n}(\cdot, \vec{v})$. We claim that \vec{s} is $*$ -ordered, hence $\overline{\Xi}^{\sigma,*}(\vec{s}) = \overline{\Xi}^{\sigma,*}(\vec{s})$ by Definition 3.10, (20) and Lemma 4.2, so (29) follows from (30).

The claim that, indeed, \vec{s} is $*$ -ordered follows from the assumption that the components of \vec{u} constitute the vertices of a convex polygon, (31), Lemma 2.7 and Proposition 3.16. □

If we return to the TSP (Sect. 3), the analogous result holds also for D_0 with respect to $T = T(\vec{v})$, namely:

Claim. *If the components of $\vec{v} \in \mathbb{C}^n$ constitute the vertices of a convex polygon in \mathbb{C} , then $\vec{v} \in D_0$.*

Evidently, the case where all points v_i constitute vertices of a convex polygon is a trivial one for the Travelling Salesman Problem. So, the above sufficient condition for the validity of the dual formulation in Theorems 3.13 and 3.14, and the optimality of the estimate of Theorem 3.15 is not very useful.

However, in the case of the Lazy Travelling Salesman, the result of Lemma 5.1 is applied for the following:

Theorem 5.2. *Let $\vec{v} = (v_1, \dots, v_n) \in \mathbb{C}^n$ such that no point $v_i \in \mathbb{C}$ is in the center of mass of all other points. Then for σ sufficiently large, $\vec{v} \in D_\sigma$.*

In particular, we conclude that Theorems 4.4 and 4.5 can be applied, and Theorem 4.6 is optimal for almost all distributions of the points v_j , provided σ is large enough.

The proof of Theorem 5.2 is obtained from Lemma 5.1 and the following:

Proposition 5.3. *Assume that the set $\{v_i\}$ satisfies*

$$|v_j - v_{c.m}|^2 > \frac{1}{n\sigma} \sum_{i=1}^n |v_i - v_{c.m}|^2, \quad \forall j = 1, \dots, n \tag{32}$$

where

$$v_{c.m} = \frac{1}{n} \sum_{i=1}^n v_i.$$

If \vec{u}^0 is a minimizer of $\overline{\mathbf{G}}_{\sigma,n}(\cdot, \vec{v})$, then u_i^0 are the vertices of a convex polygon.

Proof. Since \vec{u}^0 is a minimizer it follows that

$$\overline{\mathbf{G}}_{\sigma,n}(\vec{u}^0, \vec{v}) \leq \overline{\mathbf{G}}_{\sigma,n}(\vec{u}_{cm}, \vec{v}) = \frac{1}{n} \sum_{i=1}^n |v_i - v_{c.m}|^2.$$

where $\vec{u}_{cm} := (v_{cm}, \dots, v_{cm})$.

In particular, there exists a permutation τ (which we can take, with no loss of generality, as the identity $\tau = I$) such that

$$\min_{\vec{u} \in \mathbb{C}^n} \overline{\mathbf{G}}_{\sigma,n}(\vec{u}, \vec{v}) = \mathbf{G}_{\sigma,n}^I(\vec{u}^0, \vec{v}) \leq \frac{1}{n} |\vec{v} - \vec{u}_{c.m}|^2. \tag{33}$$

In particular

$$\sigma \mathbf{L}^2(\vec{u}^0) \leq \frac{1}{n} |\vec{v} - \vec{u}_{c.m}|^2$$

and

$$r^2(\vec{u}^0) := \max_{1 \leq i < j \leq n} |u_i^0 - u_j^0|^2 \leq \frac{1}{n\sigma} |\vec{v} - \vec{u}_{c.m}|^2 =: M_1. \tag{34}$$

By Corollary 2.8 and (34) we obtain that

$$|u_i^0 - v_{c.m}|^2 \leq r^2(\vec{u}^0) \leq \frac{1}{n\sigma} |\vec{v} - \vec{u}_{c.m}|^2, \quad \forall i = 1, \dots, n. \tag{35}$$

By (35) and the assumption of the proposition we obtain that there exists a disk $B \subset \mathbb{C}$ which contains all the points u_i^0 , $i = 1, \dots, n$, and non of the points v_j , $j = 1, \dots, n$.

Now, let $\mathbf{C}^0 \subset \mathbb{C}$ be the convex hull of $\{u_1^0, \dots, u_n^0\}$. Evidently,

$$\mathbf{L}_n(\vec{u}^0) \geq L(\mathbf{C}^0) \tag{36}$$

where $L(\mathbf{C}^0)$ is the perimeter of \mathbf{C}^0 . Let w_i be the closest point to v_i in \mathbf{C}^0 , that is, $w_i \in \mathbf{C}^0$ satisfies

$$|v_i - w_i| \leq |v_i - w| \quad \forall w \in \mathbf{C}^0. \tag{37}$$

Let $\vec{w} := (w_1, \dots, w_n)$. Then

$$|\vec{v} - \vec{w}| \leq |\vec{v} - \vec{u}^0|, \text{ so } |\vec{v}_\tau - \vec{w}_\tau| \leq |\vec{v}_\tau - \vec{u}_\tau^0| \tag{38}$$

for any permutation τ of $\{1, \dots, n\}$. So let τ be such a permutation which arrange $w_j := r_j e^{i\theta_j}$, $\theta_j \in [0, 2\pi)$ in a cyclical order $\theta_{\tau(j)} \leq \theta_{\tau(j+1)}$, $j = 1, \dots, n-1$. Thus

$$\mathbf{L}_n(\vec{w}_\tau) \leq L(\mathbf{C}^0), \quad (39)$$

so, from (36, 38, 39) we obtain

$$\mathbf{G}_{\sigma,n}^\tau(\vec{w}, \vec{v}) \equiv \mathbf{G}_{\sigma,n}^I(\vec{w}_\tau, \vec{v}_\tau) \leq \mathbf{G}_{\sigma,n}^I(\vec{u}^0, \vec{v}) \equiv \overline{\mathbf{G}}_{\sigma,n}(\vec{u}^0, \vec{v}). \quad (40)$$

Since \vec{u}^0 is a minimizer of $\overline{\mathbf{G}}_{\sigma,n}(\cdot, \vec{v})$ by assumption, it follows that there is an equality in (40). This implies an equality in (36). But, since \mathbf{C}^0 is a convex set, it follows that $\vec{w} = \vec{u}^0$. As a result, the points u_i^0 are on the boundary of a convex set (since, evidently, w_i are). \square

REFERENCES

- [1] A.J. Abrantes and J.S. Marques, Unified approach to snakes, elastic nets and Kohonen maps, in *Proceeding ICASSP IEEE International Conference on Acoustics Speech Signal Process* (1995) 3427–3430.
- [2] L. Cohen, On active contour models and balloons. *CVGIP, Image Underst.* **52** (1991) 211–218.
- [3] R. Durbin and D. Willshaw, An analogue approach to the travelling salesman problem using an elastic net method. *Nature* **326** (1987) 681–691.
- [4] S.J. Gilson and R.I. Damper, An empirical comparison of neural techniques for edge linking of images. *Neural Comput. Appl.* **6** (1997) 64–78 (Historical Archive).
- [5] E. Gurewitz, K. Rose and G.C. Fox, Constrained clustering as an optimization method. *IEEE Trans. Pattern Anal. Machine Intelligence* **15** (1993) 785–794.
- [6] J.B. Hiriart-Urruty and C. Lemaréchal, *Convex Analysis and Minimization Algorithms II, Grundlehren der Mathematischen Wissenschaften* **306**, Chap. 10. Springer-Verlag (1993).
- [7] T. Kohonen, *Self-Organizing Maps, Springer Series in Information Sciences* **30**. Springer-Verlag (1997).
- [8] S. Skyum, A simple algorithm for computing the smallest enclosing circle. *Process. Lett.* **37** (1991) 121–125.
- [9] R. Szeliski, R. Durbin and A. Yuille, An analysis of the elastic net approach to the travelling salesman problem. *Neural Comput.* **1** (1989) 348–358.
- [10] D. Tsafir, I. Tsafir, L. Ein-Dor, O. Zuk, D.A. Notterman and E. Domany, Sorting points into neighborhoods (spin): data analysis and visualization by ordering distance matrices. *Bioinformatics* **21** (2005) 2301–2308.
- [11] E. Welzl, Smallest enclosing disks (balls) and ellipsoids, in *New Results and New Trends in Computer Science*, H. Maurer Ed., *Lect. Notes Comput. Sci.* (1991) 359–370.
- [12] C. Williams, *Combining deformable models and neural networks for hand-printed digit recognition*. Ph.D. thesis, University of Toronto (1994).
- [13] A. Witkin, M. Kass and D. Terzopoulos, Snakes: Active contour models. *First International Conference on Computer Vision* (1987).