Journal of Global Optimization **21:** 139–155, 2001.

Gradient-constrained minimum networks. I. Fundamentals ⋆

M. BRAZIL[1], J.H. RUBINSTEIN[2], D.A. THOMAS[1], J.F. WENG[1] and N.C. WORMALD[2]
[1]*ARC Special Research Centre for Ultra-Broadband Information Networks, Department of Electrical and Electronic Engineering, The University of Melbourne, Victoria 3010, Australia*
[2]*Department of Mathematical Sciences, The University of Melbourne, Victoria 3010, Australia*

Abstract. In three-dimensional space an embedded network is called gradient-constrained if the absolute gradient of any differentiable point on the edges in the network is no more than a given value m. A *gradient-constrained minimum Steiner tree T* is a minimum gradient-constrained network interconnecting a given set of points. In this paper we investigate some of the fundamental properties of these minimum networks. We first introduce a new metric, the *gradient metric*, which incorporates a new definition of distance for edges with gradient greater than m. We then discuss the variational argument in the gradient metric, and use it to prove that the degree of Steiner points in T is either three or four. If the edges in T are labelled to indicate whether the gradients between their endpoints are greater than, less than, or equal to m, then we show that, up to symmetry, there are only five possible labellings for degree 3 Steiner points in T. Moreover, we prove that all four edges incident with a degree 4 Steiner point in T must have gradient m if m is less than 0.38. Finally, we use the variational argument to locate the Steiner points in T in terms of the positions of the neighbouring vertices.

Key words: Gradient constraint; Steiner trees; Minimum networks.

1. Introduction

The Euclidean Steiner tree problem asks for a shortest network T interconnecting a given point set N in Euclidean space. Such a network is necessarily a tree, and may include additional nodes (not in N) to minimise the length of the tree. The tree T is called a Euclidean minimum Steiner tree on N. The given points N are called *terminals* while the additional points are called *Steiner points* (Hwang et al., 1992). The degree of a Steiner point is assumed to be no less than 3, since otherwise there would be no advantage in adding it.

One application of the Steiner tree problem in three-dimensional space is to the underground mining industry. Given a number of ore deposits, whose locations are known, the infrastructure costs of an underground mine can be minimised by finding the shortest network of tunnels interconnecting these deposits to a given access point. In practice, however, an important constraint due to haulage needs to be imposed, namely that the slope of a tunnel cannot be very steep. The typical

⋆ Supported by a grant from the Australian Research Council.

maximum gradient m for tunnels is about 1:7. Another possible application of minimum Steiner trees is to the design of major road networks. Again, the gradient of a road cannot be very large. In this paper we study the properties of any shortest network T with such a gradient constraint, interconnecting a given set of points. We refer to such a network as a *gradient-constrained minimum Steiner tree*. We assume the maximum allowed gradient m is strictly greater than 0, otherwise T cannot exist unless all terminals lie on a horizontal plane. In the latter case, the problem becomes the classical Euclidean Steiner tree problem in a plane.

Let x_p, y_p, z_p denote the Cartesian coordinates of a point p in three-dimensional space. Assume that the z-axis is vertical. Then, by the gradient of an edge pq we mean the absolute value of the slope from p to q, which is denoted by $g(pq)$. That is,

$$g(pq) = \frac{|z_q - z_p|}{\sqrt{(x_q - x_p)^2 + (y_q - y_p)^2}}.$$

Suppose pq is an edge of T embedded in Euclidean space. If $g(pq) \leqslant m$, then pq is a straight line joining p and q, and is referred to as a *straight edge*. However, if $g(pq) > m$, then pq cannot be represented as a straight line without violating the gradient constraint, but it can be represented by a zigzag line joining p and q with each segment having gradient m. Such edges are referred to as *bent edges*.

The most widely studied Euclidean Steiner tree problem is the planar problem, which has been shown to be NP-hard (Garey et al., 1977). The three-dimensional Euclidean Steiner tree problem is known to be considerably more difficult than the planar version. Some of the basic properties of such trees have been studied in (Gilbert & Pollack, 1968) and (Smith, 1992). The gradient-constrained Steiner tree problem, although likely to be more applicable, is even more complicated, and to date has received very little study. The only previously published paper directly addressing this problem is (Brazil et al., 1998), in which the terminals N are assumed to all lie on a single vertical plane.

Despite the high levels of complexity involved in solving the Steiner problem, it is nevertheless important to study the properties of exact minimum Steiner trees. In the planar case this has lead to the exact algorithm of Warme et al. (2000), which is remarkably efficient for large, randomly generated instances. In higher dimensional cases, such properties have proved crucial in development of approximation algorithms. For example, Smith's approximation algorithm for d-dimensional Steiner trees (Smith, 1992) uses the fact that all angles at Steiner points in exact minimum Steiner trees are $2\pi/3$.

In this present series of papers we will conduct a rigorous study of the most important properties of gradient-constrained minimum Steiner trees, beginning, in this paper, with the fundamental properties of their Steiner points. The main tool used here is the variational argument (Rubinstein and Thomas, 1991), which has been proved to be very powerful in the study of the Euclidean Steiner tree problem. In Section 2, we introduce a new metric, the *gradient metric*, and discuss the variational argument in the gradient metric. We prove that all Steiner points lie in the convex hull of the terminals in the gradient metric. Section 3 is the main part of this paper. We prove that the degree of Steiner points is either three or four. If the gradient of an edge is less than, equal to, or greater than m, then the edge is labelled by 'f' (meaning flat), 'm' (meaning maximum), or 'b' (meaning bent), respectively. The set of labels around a Steiner point is referred to as the *labelling* of that point. A labelling that can be achieved in a minimum Steiner tree is referred to as a *feasibly optimal* labelling. We prove that, up to symmetry, there are only five feasibly optimal labellings for the edges incident with a degree 3 Steiner point. Moreover, we prove that all four edges incident with a degree 4 Steiner point must have label 'm' if the maximum gradient m is less than 0.38. In Section 4, using the variational argument, we discuss how to locate geometrically any Steiner point in T (for small m), in terms of the positions of its adjacent nodes. The final section is a brief discussion of the extent to which the results in this paper generalise to higher dimensions.

2. The gradient metric and the variational argument

The lengths of edges in a gradient-constrained tree can be measured in a special metric, called the *gradient metric*. Suppose o is the origin and $p = (x_p, y_p, z_p)$ is a point in space. Define the *vertical metric* of the line op to be $|op|_v = cz_p$ where c is a given constant. Then the gradient metric can be defined in terms of the Euclidean and vertical metrics. Suppose m is the maximum gradient allowed in gradient-constrained trees. The length of op in the gradient metric is defined to be

$$|op|_g = \begin{cases} |op| = \sqrt{x_p^2 + y_p^2 + z_p^2}, & \text{if } g(op) \leqslant m; \\ |op|_v = (\sqrt{1+m^{-2}})|z_p|, & \text{if } g(op) \geqslant m. \end{cases}$$

It is easily checked that this defines a metric. Note that $|op| \leqslant |op|_g$, and the gradient metric is convex though it is not strictly convex.

Let T be a gradient-constrained minimum Steiner tree. We can now assume that all edges of T are straight lines whose lengths are given by the gradient metric. Let $|T|_g$ be the sum of the lengths of all edges in T. An edge pq in T is called an f-edge, m-edge or b-edge if pq is labelled 'f' ($g(pq) < m$), 'm' ($g(pq) = m$) or 'b' $g(pq) > m$), respectively. The label of an edge can be thought of as indicating which metric is 'active' for that edge, with an 'm' label indicating that both metrics hold simultaneously.

The variational argument in the Steiner tree problem is as follows: for a minimum tree T, the directional derivative of $|T|_g$ is greater than or equal to zero when its Steiner points are perturbed in any directions. Note that under an arbitrarily small perturbation the only edges which can change labelling are m-edges. If no m-edges change their labelling under a given perturbation, then it is easily checked that the variation is reversible, and hence the directional derivative is strictly zero.

Suppose $e = sa$ is an edge in T, and s is a Steiner point which is perturbed to s' in direction $\mathbf{u}$. Let $\dot{e}_{\mathbf{u}}$ (or simply $\dot{e}$ if $\mathbf{u}$ is known) denote the directional derivative of the length of e. It is easy to show the following lemma and corollary:

LEMMA 1. *(i) If e is an f-edge, then* $\dot{e}_{\mathbf{u}} = -\cos(\angle ass')$.

(ii) If e is a b-edge, then $\dot{e}_{\mathbf{u}} = -\cos(\angle zss')\sqrt{1+m^{-2}}$ *where* z *is a point on the vertical line through* s *such that* $\angle asz \leqslant \pi/2$.

(iii) If e *is an m-edge, then* $\dot{e}_{\mathbf{u}}$ *is equal to either* $-\cos(\angle ass')$ *or* $-\cos(\angle ass')\sqrt{1+m^2}$, *depending on whether* $g(s'a) \leqslant m$ *or* $g(s'a) > m$.

COROLLARY 1. *(i) When* s *moves horizontally, the length of* e *does not change if* e *is a* b*-edge or if* e *is an* m*-edge and becomes a* b*-edge in the move.*

(ii) When s *moves vertically to the same side (or the opposite side) of the horizontal plane through* s *as* a, e *becomes shorter (or, respectively, longer) regardless of the gradient of* e.

REMARK 1. *From the lemma it is clear that, as in the Euclidean metric, the directional derivative of* e *in the gradient metric is determined only by the direction of* $\mathbf{u}$ *and is independent of the length of* sa.

Because an f-edge is still an f-edge under a small perturbation, Lemma 1 gives the following easy corollary:

LEMMA 2. *Any* f*-edge of a Steiner point* s *meets other edges incident to* s *at an angle no less than* $\pi/2$.

We end this section with a theorem involving the convexity of the gradient metric. Suppose T is a minimum Steiner tree on a point set N in space. If there is no gradient constraint, i.e., if T is a Euclidean Steiner tree, then T has the following properties (Hwang et al., 1992):

(S1) any Steiner point s lies in the plane $\mathcal{P}_{\triangle}$ containing its three adjacent vertices; and

(S2) consequently all Steiner points of T lie in the convex hull, hull(N), of N

The first statement (S1) does not generally carry over to gradient-constrained minimum Steiner trees. (S1) is true if s has degree 3 and its incident edges are all f-edges. Similarly, (S1) is true if the adjacent nodes to s all lie in a vertical plane (Brazil et al. 1998). However, in general, s may not lie in the plane $\mathcal{P}_{\triangle}$, as we will see in the geometric characterisations of s in Section 4.

On the other hand, we can show that the second statement (S2) still holds in the gradient-constrained metric. In the Euclidean metric, there are several equivalent definitions of hull(N), one being the minimal set for which N is in hull(N) and, if p and q are two points in hull(N), then all points on the line segment joining p and q are in hull(N). Now we define the hull in the gradient-constrained metric in a similar way: the hull of N in the gradient-constrained metric, denoted by $\mathrm{hull}_g(N)$, is the minimal set such that N is in $\mathrm{hull}_g(N)$ and, if p and q are two points in

$\mathrm{hull}_g(N)$, then all points in all shortest paths joining p and q are in $\mathrm{hull}_g(N)$. It is easily checked that $\mathrm{hull}_g(N)$ is the union of hull(N) and all shortest paths between points in hull(N). Note that no straight line segment on the boundary of $\mathrm{hull}_g(N)$ has gradient greater than m.

THEOREM 1. *If N is the terminal set of the gradient-constrained minimum Steiner tree T then all Steiner points of T lie in* $\mathrm{hull}_g(N)$.

Proof. Let p, q be points lying outside or on the boundary of $\mathrm{hull}_g(N)$. Define p',q' to be the respective projections of p, q (i.e., the nearest points in the Euclidean metric) onto the boundary of $\mathrm{hull}_g(N)$.

We first show that $|pq|_g \geqslant |p'q'|_g$, with strict inequality if one of p, q lies on the boundary of $\mathrm{hull}_g(N)$ and the other is outside $\mathrm{hull}_g(N)$. We can assume, without loss of generality that q lies outside $\mathrm{hull}_g(N)$. By the definition of p' and q', we have $\angle pp'q' \geqslant \pi/2$ and $\angle p'q'q \geqslant \pi/2$. So if $g(p'q') \leqslant m$ then

$$|pq|_g \geqslant |pq| \geqslant |p'q'| = |p'q'|_g.$$

where the second inequality is strict if p lies on $\mathrm{hull}_g(N)$. On the other hand, if $g(p'q') > m$ then we can assume, without loss of generality, that $z_{p'} > z_{q'}$. Note that any point p'' near and below p' such that $g(p'p'') \geqslant m$, must lie in $\mathrm{hull}_g(N)$. Hence $z_p \geqslant z_{p'}$, and similarly $z_{q'} > z_q$, which implies $|pq|_g > |p'q'|_g$ by the definition of the gradient metric.

Now suppose, contrary to the theorem, there is a Steiner point s lying outside hull$g(N)$. Then there exists a subtree T_1 of T containing s and at least one point on the boundary of $\mathrm{hull}_g(N)$, but no points from the interior of $\mathrm{hull}_g(N)$. It follows from the above argument that projecting T_1 onto the boundary of $\mathrm{hull}_g(N)$ decreases the length of T, giving the required contradiction. □

3. Feasibly optimal labellings at Steiner points

A Steiner point in a gradient-constrained tree is called *optimal* if its perturbation cannot shorten the tree. Let s be a Steiner point in a gradient-constrained Steiner tree T. Our aim in this section is to classify all possible sets of labellings of edges incident with s, and to determine which of these edges can lie on the same side of the horizontal plane through s. Throughout this section we will assume that $m \leqslant 1$. Some further restrictions on m will also be required for some of the results in Subsection 3.2.

First, by the triangle inequality, a basic property of optimal Steiner points is:

LEMMA 3. *If a Steiner point s is optimal with respect to its adjacent points $a, b, c, \ldots$, then s is also optimal with respect to any points $a', b', c', \ldots$ that lie on $sa, sb, sc, \ldots$, respectively.*

By this lemma an edge incident with s is often assumed to have a particular length, e.g. 1 or $\sqrt{1+m^2}$, without explanation in the proofs below.

3.1. PROJECTIONS OF ANGLES, AND DEGREE 3 STEINER POINTS

An important application of the variational argument is that of splitting a small angle in a non-optimal tree. If we perturb a Steiner point s of T in direction $\mathbf{u}$, denote the directional derivative of $|T|_g$ by $\dot{T}_{\mathbf{u}}$, or just $\dot{T}$. Now suppose s is a Steiner point joining a, b and c in a Euclidean Steiner tree $T = sa \cup sb \cup sc$. If $\angle acb \geqslant 2\pi/3$, then s collapses into c. On the other hand, if $\angle acb < 2\pi/3$ but $s = c$, then T cannot be minimum by the following variational argument: Let s_1 be a point on the bisector $\mathbf{u}$ of $\angle asb$ and close to s, then $\angle ass_1 = \angle bss_1 < \pi/3$ and $\dot{T} = 1 - \cos(\angle ass_1) - \cos(\angle bss_1) < 0$. In general, we refer to this process as *splitting* the angle at s in the subtree $sa \cup sb$ along a vector $\mathbf{u}$. Throughout this section, we apply this technique to the gradient-constrained Steiner minimum tree T.

We begin by considering Steiner points with incident b-edges. For any point p, denote the horizontal plane through p by $\mathcal{H}_p$. For simplicity, a point or an edge will be said to be above (or below) p if it is above (or below) $\mathcal{H}_p$. We also employ the convention of saying that two edges incident with s lie on the same side of $\mathcal{H}_s$ if they lie in the same closed half-space determined by $\mathcal{H}_s$; i.e., this includes the possibility that one or both edges lie on $\mathcal{H}_s$. The two edges are said to be on different sides of $\mathcal{H}_s$ only if their interiors lie in different open half-spaces (determined by $\mathcal{H}_s$).

LEMMA 4. *If sa is a b-edge in T, then no other edge incident with s lies on the same side of $\mathcal{H}_s$ as a. Moreover, all other edges incident with s are m-edges.*

Proof. Let sb be another edge of T incident with s. Let $\mathbf{u}$ be a vector with gradient m, perturbing s to s', such that s' lies on the same side of $\mathcal{H}_s$ as a and the angle $\theta = \angle s'sb$ is as small as possible. Applying the variational argument to T by splitting the angle $\angle asb$ at s along $\mathbf{u}$, we note that $\dot{T} = \dot{sb}_{\mathbf{u}}$. If sb is a b-edge lying on the same side of $\mathcal{H}_s$ as a then $\dot{T} < 0$ by Lemma 1(ii). If, on the other hand, sb is an m-edge or f-edge on the same side of $\mathcal{H}_s$ as a, or sb is an f-edge on the opposite side of $\mathcal{H}_s$ to a, then $\theta < \pi/2$ (since $m \leqslant 1$ in the latter case). So, again $\dot{T} < 0$ by Lemma 1. In each case we have a contradiction to the minimality of T.

Finally, if sb is a b-edge lying on the opposite side of $\mathcal{H}_s$ to a, then s has only two incident edges by the above argument and s is not a Steiner point. □

THEOREM 2. *Let s be a Steiner point of T, and let $\mathcal{P}$ be a horizontal or vertical plane through s not containing any edges of T incident to s. Then, the edges of T incident to s do not all lie on one side of $\mathcal{P}$.*

Proof. If $\mathcal{P}$ is horizontal, then the theorem follows by Corollary 1(ii). So, suppose, contrary to the theorem, that all edges of s lie on one side of a vertical plane $\mathcal{P}$ through s. Move s perpendicularly to $\mathcal{P}$ and into the half-space where the edges of s lie. By such a move, any edge of s, say sa, becomes shorter by Lemma 1(i) if it is an f-edge, or sa does not change its length and becomes a b-edge by Corollary 1(i) if it is an m-edge or a b-edge. Therefore, either T is shortened, or the perturbed

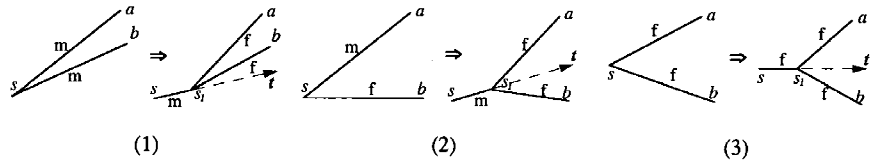

Figure 1. Splitting an angle.

tree has the same length as T and has two b-edges lying on the same side of $\mathcal{H}_s$ (since s has at least three incident edges). In the latter case T is again not minimal by Lemma 4. Therefore not all edges of s lie on one side of $\mathcal{P}$ if T is minimal. □

Lemma 4 characterises Steiner points with an incident b-edge; so for the remainder of this section we focus on cases where no b-edge exists. Suppose sa, sb are two edges incident with s, and not labelled 'b'. As a convention, if p is a point in T, then we will denote its projection on $\mathcal{H}_s$ by p'. To find the properties of s, we will study the projection $\angle a'sb'$ instead of $\angle asb$ itself, in order to take advantage of properties of angles on a plane. Note that sa, sb may both be m-edges, or both be f-edges, or one may be an m-edge and the other an f-edge, and in each case they either lie on the same side of $\mathcal{H}_s$, or on two sides of $\mathcal{H}_s$. Hence, there are six different ways of classifying a and b with respect to s. For each classification the lower bound of $\angle a'sb'$ is denoted by $\gamma_{(\mathrm{mm1})}$, $\gamma_{(\mathrm{ff1})}$, $\gamma_{(\mathrm{mf1})}$, $\gamma_{(mm2)}$, $\gamma_{(\mathrm{ff2})}$, or $\gamma_{(\mathrm{mf2})}$. Below is a list of these lower bounds.

Labels	Lying on	Projection of angle
m,m	1 side of $\mathcal{H}_s$	$\geqslant \gamma_{(\mathrm{mm1})} = 2\arccos\frac{1-m^2}{2} > \frac{2\pi}{3}$
f,f	1 side of $\mathcal{H}_s$	$\geqslant \gamma_{(\mathrm{ff1})} = \frac{2\pi}{3}$
m,f	1 side of $\mathcal{H}_s$	$> \gamma_{(mf1)} = 2\arccos\frac{1}{1+\sqrt{1+m^2}} > \frac{2\pi}{3}$
m,m	2 sides of $\mathcal{H}_s$	$\geqslant \gamma_{(\mathrm{mm2})} = 0$
f,f	2 sides of $\mathcal{H}_s$	$> \gamma_{(\mathrm{ff2})} = \arccos\frac{-1+m^2}{2} \geqslant \frac{\pi}{2}$
m,f	2 sides of $\mathcal{H}_s$	$> \gamma_{(\mathrm{mf2})} = \arccos\frac{-1+3m^2+m^4}{4-m^2-m^4} \geqslant 0$

The bound $\gamma_{(\mathrm{mm2})}$ is trivial. The other bounds will be proved separately in several lemmas. Let $\angle(\mathbf{v}_1, \mathbf{v}_2)$ denote the angle between two vectors $\mathbf{v}_1$, $\mathbf{v}_2$.

LEMMA 5. *If sa, sb are m-edges lying on the same side of $\mathcal{H}_s$ then $\angle a'sb' \geqslant \gamma_{(mm1)} \geqslant 2\pi/3$.*

Proof. Let 2α be the angle between sa' and sb'. Without loss of generality let $s = (0,0,0)$, $a = (\cos\alpha, \sin\alpha, m)$, $b = (\cos\alpha, -\sin\alpha, m)$. Suppose we split $sa \cup sb$ by moving s to s_1 in direction $\mathbf{t} = (1,0,m)$. Under such a move ss_1 is an m-edge and both sa and sb become f-edges (Fig. 1(1)). Since T is optimal we obtain, by the variational argument,

$$\begin{aligned}\dot{T} &= 1 - \cos\angle(\vec{sa}, \mathbf{t}) - \cos\angle(\vec{sb}, \mathbf{t})\\ &= 1 - 2\frac{\cos\alpha + m^2}{1 + m^2} \geqslant 0.\end{aligned}$$

This inequality gives the required lower bound for 2α. □

LEMMA 6. *Suppose sa is an m-edge in T with $s = (0, 0, 0)$, $a = (\cos\alpha, \sin\alpha, m)$, $0 \leqslant \alpha < \pi/2$. Suppose the projection of a vector* **t** *meets the x-axis at angle θ so that a move of s in direction* **t** *preserves the gradient of sa. Then*

$$\mathbf{t} = (\cos\theta, \sin\theta, m\cos(\alpha - \theta)).$$

Proof. By the assumption let $s_1 = r(\cos\theta, \sin\theta, h)$ be a point on **t**, where r satisfies $rh < m$. Since

$$g(as_1) = \frac{m - rh}{\sqrt{(r\cos\theta - \cos\alpha)^2 + (r\sin\theta - \sin\alpha)^2}} = m,$$

we have

$$h = \left(1 - \sqrt{(r\cos\theta - \cos\alpha)^2 + (r\sin\theta - \sin\alpha)^2}\right) m/r$$

and

$$\lim_{r\to 0} h = m\cos(\alpha - \theta).$$

The lemma is proved. □

REMARK 2. *In the proofs that make use of this lemma, some results can be improved by choosing a different parameter θ. However, to make the proofs simple, we always choose $\theta = 0$ and* $\mathbf{t} = (1, 0, m\cos\alpha)$.

LEMMA 7. *Suppose sa is an m-edge and sb is an f-edge in T.*

(i) If both edges lie on the same side of $\mathcal{H}_s$, then $\angle a'sb' > \gamma_{(\mathrm{mf1})} \geqslant 2\pi/3$.

(ii) If sa, sb lie on two sides of $\mathcal{H}_s$, then $\angle a'sb' > \gamma_{(\mathrm{mf2})} \geqslant 0$.

Proof. As in Lemma 5, without loss of generality let $s = (0, 0, 0)$ and $a = (\cos\alpha, \sin\alpha, m)$, where $2\alpha = \angle a'sb'$. So $b = (\cos\alpha, -\sin\alpha, h)$, for some h. In each case we split $\angle asb$ in a direction close to the bisector, and show this contradicts $\dot{T} \geqslant 0$ for 2α below the given bound.

(i) First note that $\alpha \leqslant \pi/2$ by the minimality of T. Split $sa \cup sb$ by moving s in direction **t** = (1,0,m), which results in sa becoming an f-edge (Fig. 1(2)). For a fixed α, $\dot{T}$ is maximised under this split when $\angle(\vec{sb}, \mathbf{t})$ is maximum, which clearly occurs when $h = 0$. So, we can assume that $b = (\cos\alpha, -\sin\alpha, 0)$. Since T is optimal,

$$\dot{T} = 1 - \cos\angle(\vec{sa}, \mathbf{t}) - \cos\angle(\vec{sb}, \mathbf{t})$$
$$= 1 - \frac{\cos\alpha + m^2}{1 + m^2} - \frac{\cos\alpha}{\sqrt{1 + m^2}} \geqslant 0.$$

The bound $\gamma_{(\mathrm{mf1})}$ is easily derived from this inequality.

(ii) In this case, split $\angle asb$ in the direction $\mathbf{t} = (1, 0, m\cos\alpha)$, preserving the gradient of sa (by Lemma 6). Here $\angle(\vec{sb}, \mathbf{t})$ is maximised for a given α when $h \to -m$. So, setting $b = (\cos\alpha, -\sin\alpha, -m)$, we obtain $\gamma_{(\mathrm{mf2})}$ by a similar argument to that above. Note that the expression for $\gamma_{(\mathrm{mf2})}$ has been simplified by using the identity $\cos 2x = 2\cos^2 x - 1$. □

LEMMA 8. *Suppose sa, sb are two f-edges in T.*

(i) If sa, sb lie on two sides of $\mathcal{H}_s$, then $\angle a'sb' > \gamma_{(\mathrm{ff2})} \geqslant \pi/2$.

(ii) If both edges lie on the same side of $\mathcal{H}_s$, then $\angle a'sb' \geqslant \gamma_{(\mathrm{ff1})} = 2\pi/3$.

Proof. Again, let $2\alpha = \angle a'sb'$; then we can assume, for some h_1, h_2 satisfying $0 \leqslant h_1, h_2 < m$, that $s = (0, 0, 0)$, $a = (\cos\alpha \sin\alpha, h_1)$ and $b = (\cos\alpha, -\sin\alpha, -h_2)$ in Case (i) or $b = (\cos\alpha, -\sin\alpha, h_2)$ in Case (ii).

(i) Let $a^* = (\cos\alpha, \sin\alpha, m)$, $b^* = (\cos\alpha, -\sin\alpha, -m)$ and $s_1 = (1, 0, 0)$. Then $\angle a^*ss_1 \geqslant \angle ass_1$, $\angle b^*ss_1 \geqslant \angle bss_1$. Hence, for a given α we need only consider the extreme case where $h_1 = h_2 = m$. Thus, by splitting s in the direction $\vec{ss_1}$, we have $\dot{T} = 1 - 2\cos\alpha/\sqrt{1 + m^2} \geqslant 0$ and obtain the required $\gamma_{(\mathrm{ff2})}$.

(ii) First note, by a similar argument to that used in Case (i) (using the same splitting direction), that $\alpha > \pi/4$. Hence $\angle asb \leqslant \angle a'sb'$. Let $\mathbf{t}$ be the vector bisecting $\angle asb$. If $g(\mathbf{t}) \leqslant m$ then $\angle asb \geqslant 2\pi/3$ by the same variational argument used in the Euclidean case, and the result follows.

So assume, on the other hand, that $g(\mathbf{t}) > m$. Let $s_1 = (1, 0, m)$. We will split $\angle asb$ along $\vec{ss_1}$ (Fig. 1(3)). Consider the plane $\mathcal{P}$ containing s_1 and the y-axis. By the assumption, either a or b lies above $\mathcal{P}$. Hence, without loss of generality, we can assume that a lies above $\mathcal{P}$. It follows that $a^* = (\cos\alpha, \sin\alpha, m)$ satisfies $\angle a^*ss_1 \geqslant \angle ass_1$.

Now there are two possibilities for b. Either $\angle bss_1 \leqslant b^*ss_1$, where $b^* = (\cos\alpha, -\sin\alpha, m)$, in which case $2\alpha > 2\pi/3$ by the proof of Lemma 5. Otherwise $\angle bss_1 \leqslant b_0ss_1$, where $b_0 = (\cos\alpha, -\sin\alpha, 0)$, in which case we again have $2\alpha > 2\pi/3$ by the proof of Lemma 7(i). This concludes the proof. □

LEMMA 9. *If s is a Steiner point in T then there are at most two incident edges lying strictly above (or below) $\mathcal{H}_s$. Consequently the degree of any Steiner point is either three or four.*

Proof. If s has an incident b-edge, then the lemma is valid by Lemma 4. If, on the other hand, there are more than two incident straight edges (each an f-edge or m-edge) on the same side of $\mathcal{H}_s$, then among their projections two of them meet at an angle of no more than $2\pi/3$. This contradicts either Lemma 5, Lemma 7(i) or Lemma 8(ii). □

Suppose s is a degree 3 Steiner point in T with two incident edges sa, sb lying on one side of $\mathcal{H}_s$, and the third sc lying on the other side of $\mathcal{H}_s$. (This includes the possibility that all three edges lie on $\mathcal{H}_s$.) Let g_a, g_b, g_c denote the respective labels of these edges. Then we say the *labelling* of this degree 3 Steiner point is $(g_a g_b/g_c)$. By symmetry, we can assume a and b both lie on or above $\mathcal{H}_s$.

THEOREM 3. *If s is a degree 3 Steiner point in T, then up to symmetry there are five feasibly optimal labellings: (ff/f), (ff/m), (fm/m), (mm/m) and (mm/b).*

Proof. By Lemma 4 and Lemma 9, the only possible labellings of s, other than those listed in the statement of the theorem, are (mm/f) and (mf/f). So, suppose, contrary to the theorem, there is only one edge, the f-edge sc, lying below s, and there is an m-edge, say sa, lying above s. Because $g(sa) > g(sc)$, sa shrinks strictly faster than sc stretches when s moves vertically upwards. Since the third edge, sb, lies on the same side of $\mathcal{H}_s$ as as, and $\dot{sb} \leqslant 0$ under this move, we have $\dot{T} < 0$, contradicting the minimality of T. □

3.2. LABELLINGS FOR DEGREE 4 STEINER POINTS.

We now use the angle projections obtained in the previous lemmas to study degree 4 Steiner points. Suppose s is a degree 4 Steiner point in T with two edges sa, sb lying on one side of $\mathcal{H}_s$, and the other two edges sc, sd lying on the other side of $\mathcal{H}_s$. Let g_a, g_b, g_c, g_d be the respective labels of these edges. Then the labelling of s is denoted $(g_a g_b/g_c g_d)$. Again we assume a and b both lie on or above $\mathcal{H}_s$.

By Lemma 4 and Lemma 9, none of the edges incident with s is a b-edge. Therefore, up to symmetry the possible labellings of s are (ff/ff), (mf/ff), (mm/ff), (mf/mf), (mm/mf) and (mm/mm). Two of these labellings can immediately be shown not to be feasibly optimal.

LEMMA 10. *If s is a degree 4 Steiner point in T, then the labelling of s is not (mm/ff) or (ff/ff).*

Proof. First assume s has labelling (mm/ff). As in the proof of the previous theorem, if s moves vertically upwards then the m-edges decrease in length strictly faster than the f-edges increase in length. Hence $\dot{T} < 0$, contradicting the minimality of T.

If, on the other hand, s has labelling (ff/ff), project the four edges incident with s onto the plane $\mathcal{H}_s$. By Lemma 8, the angle between each adjacent pair of these projected edges is strictly greater than $\pi/2$; hence the sum of the angles around s is more than 2π, again giving a contradiction. □

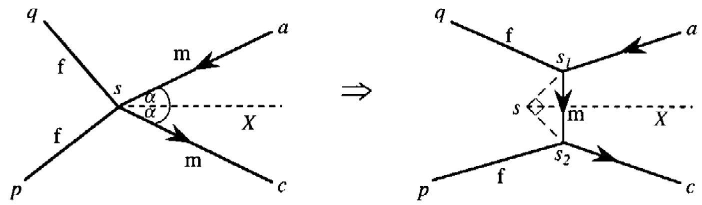

Figure 2. Splitting an (mf/mf) Steiner point. This is a plan view. The arrowheads on edges indicate direction of slope, and point downwards. The angles indicated are those between projected edges.

In the remainder of this section, we show that no degree 4 Steiner point is optimal for small m unless its labelling is (mm/mm). Let $\sum_{proj}(\angle)$ denote the sum of the four angles, made by sa', sb', sc', sd', the projections of the four edges at s. As in the proof of the previous lemma, our strategy is to show that $\sum_{proj}(\angle) > 2\pi$ for all other labellings. Up to symmetry there are two patterns of the projections sa', sb', sc', sd', distinguishing between whether adjacent edges have been projected from the same or opposite sides of $\mathcal{H}_s$. There are two patterns of the cyclic order of a', b', c', d' around s. Let *Pattern A* be $(a'b'c'd')$ and *Pattern B* be $(a'c'b'd')$.

LEMMA 11. *The labelling (mf/ff) is not feasibly optimal if m < 0.82.*

Proof. Without loss of generality assume sa is an m-edge. In Pattern A

$$\begin{aligned}\sum_{proj}(\angle) &= \angle a'sb' + \angle b'sc' + \angle c'sd' + \angle d'sa' \\ &\geqslant \gamma_{(\mathrm{mf1})} + \gamma_{(\mathrm{ff2})} + \frac{2\pi}{3} + \gamma_{(\mathrm{mf1})} \\ &> 2\pi \quad \text{if } m < 0.99325.\end{aligned}$$

In Pattern B,

$$\begin{aligned}\sum_{proj}(\angle) &= \angle a'sc' + \angle c'sb' + \angle b'sd' + \angle d'sa' \\ &\geqslant 2(\gamma_{(\mathrm{mf2})} + \gamma_{(\mathrm{ff2})}) \\ &> 2\pi \quad \text{if } m < 0.82016.\end{aligned}$$

□

In the following two lemmas, instead of splitting the angle at s we split the point s itself into two Steiner points s_1, s_2 so that $g(s_1 s_2) = m$.

LEMMA 12. *The labelling (mf/mf) is not feasibly optimal if m < 0.514.*

Proof. Without loss of generality assume sa and sc are m-edges. Consider the projections of the four edges on $\mathcal{H}_s$. In the cyclic sequence of edges around s, sa'

and sc' may or may not be adjacent. In the latter case

$$\sum_{proj}(\angle) = 2(\gamma_{(\mathrm{mf1})} + \gamma_{(\mathrm{mf2})}) > 2\pi,$$

if $m < 0.94622$.

So, we can now assume sa' and sc' are adjacent in the cyclic sequence of edges around s. Let p and q be the other two adjacent vertices to s, such that the cyclic sequence of projected vertices in $\mathcal{H}_s$ around s is $(a'c'p'q')$. Note that pa and qa are both f-edges, one above and one below $\mathcal{H}_s$. To prove the theorem, it is convenient to consider two subcases:

Case (i): Let $\angle a'sc' = 2\alpha < \pi/2$. Without loss of generality, we can assume $s = (0, 0, 0)$, $a = (\cos\alpha, \sin\alpha, m)$ and $c = (\cos\alpha, -\sin\alpha, -m)$. We split s into two Steiner points s_1 and s_2, both equidistant from s, along vectors $\mathbf{u}_1 = (1,1,m)$ and $\mathbf{u}_2 = (1,-1,-m)$. The edges in the new subtree after splitting s are s_1a, s_1s_2, s_2c, s_1q and s_2p (Fig. 2). Note that $g(s_1s_2) = m$ and $g(s_1a) = g(s_2c) \geqslant m$. The projection of the path as_1s_2c onto $\mathcal{H}_s$ is shorter than the path $a'sc'$ so it follows that the path $\mathrm{as_1s_2c}$ has the same length as asc originally (unless $\alpha = 0$, in which case the instantaneous change in length of the path is 0). Hence $\dot{T}$ equals the change in lengths of the two edges sp and sq, that is,

$$\dot{T} = -\cos\angle(qss_1) - \cos\angle(pss_2) \geqslant 0,$$

since T is minimum. Let $\theta_1 = \angle(qss_1)$ and let $\theta_2 = \angle(pss_2)$.

We wish to maximise $\dot{T}$, for a fixed projected angle $\angle p'sq'$. This means maximising θ_1 and θ_2, which obviously means forcing sp and sq to their extremal positions at slope m, with p above $\mathcal{H}_s$ and q below. Furthermore, since $\angle p'sq' > \pi/2$, and since cos is a convex function over the range of θ_1 and θ_2, it follows that the maximum of $\dot{T}$ is achieved when $\theta_1 = \theta_2 \geqslant \pi/2$. Now, consider the angle $\angle q'ss_1'$ in this extremal configuration. Since $g(qs) = m$, $g(s_1s) = m/\sqrt{2}$ and $\theta_1 \geqslant \pi/2$, it is easy to compute that the projected angle $\angle q'ss_1' \geqslant \arccos(m^2/\sqrt{2})$. Similarly, $\angle p'ss_2' \geqslant \arccos(m^2/\sqrt{2})$.

Finally, since $\angle s_1'ss_2' = \pi/2$ and $\angle p'sq' \geqslant \arccos((-1+m^2)/2)$ (by Lemma 8(i)), we have (for T minimum),

$$\begin{aligned} 2\pi &= \angle p'sq' + \angle q'ss_1' + \angle p'ss_2' + \angle s_1'ss_2' \\ &\geqslant \arccos((-1+m^2)/2) + 2\arccos(m^2/\sqrt{2}) + \pi/2, \end{aligned}$$

a contradiction when $m < 0.51444$.

Case (ii): Let $\angle a'sc' = 2\alpha > \pi/2$. The argument is the same as in the previous case, but now we let s_1 and s_2 move along the lines sa and sc. The decrease in length is even more than in Case (i). □

REMARK 3. *The variation in the above proof is far from obvious. It was found with the help of a computer program developed to construct gradient-constrained*

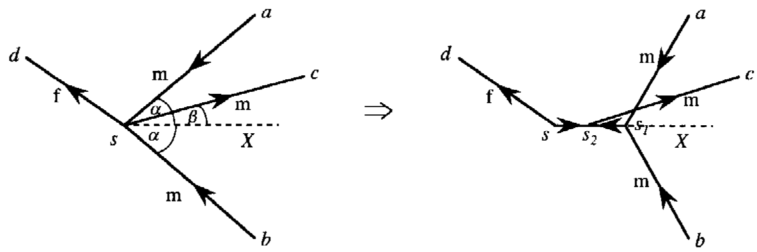

Figure 3. Splitting an (mm/mf) Steiner point. This is a plan view. The arrowheads on edges indicate direction of slope, and point downwards. The angles indicated are those between projected edges.

Steiner minimum trees. The program uses a damped Newton method which employs some of the theory developed in this paper. When using the program to attempt to construct stable Steiner points of degree 4 with labelling (mf/mf), we noticed that promising candidates for small values of m *tended to split in a manner similar to that outlined above. It should also be noted that for larger values of* m *(such as* $m \approx 0.95$*) the computer program found degree 4 points labelled (mf/mf) that appear to be stable. This suggests that, although the upper bound given in the lemma can almost certainly be improved, the labelling (mf/mf) nevertheless appears to be feasibly optimal for some larger values of* $m \leqslant 1$.

LEMMA 13. *The labelling (mm/mf) is not feasibly optimal if m < 0.38.*

Proof. In Pattern A,

$$\sum_{proj}(\angle) = \gamma_{(\mathrm{mm1})} + \gamma_{(\mathrm{mf2})} + \gamma_{(\mathrm{mf1})} > 2\pi$$

if $m < 0.92007$.

In Pattern B, without loss of generality, suppose $a = (\cos\alpha, \sin\alpha, m)$, $b = (\cos\alpha, -\sin\alpha, m)$ and $c = (\cos\beta, \sin\beta, -m), 0 \leqslant \beta \leqslant \alpha$. Since $\angle a'sd' > \gamma_{(\mathrm{mf2})}$, $\angle b'sd' > \gamma_{(\mathrm{mf2})}$ and $\angle a'sd' + \angle b'/sd' + 2\alpha = 2\pi$, we have

$$\alpha < \pi - \gamma_{(\mathrm{mf2})} = \pi - \arccos\frac{-2 + 3m^2 + m^4}{4 - m^2 - m^4} = \phi. \tag{1}$$

We can now assume that $\alpha < \pi/2$, since it follows from the above inequality for all $m < 0.74736$.

As in the previous lemma we split s into two Steiner points s_1, s_2, this time so that the new edges interconnecting s with a, b, c are s_1a, s_1b, s_1s_2 s_2c and s_2s (Fig 3). (The edge ds remains unchanged, but s is no longer a Steiner point after the split). We perform the split of s to s_1 and s_2 along vectors $(1,0,m\cos\alpha)$ and $(1,0,-m\cos\beta)$ respectively. By Lemma 6 this guarantees that $g(s_1a) = g(s_1b) = g(s_2c) = m$. Furthermore, let the ratio of lengths of s_1s and s_2s be given by

$$\rho = \frac{|s_1s|}{|s_2s|} = \frac{(1 + \cos\beta)}{1 - \cos\alpha)}.$$

Then an easy computation shows that $g(s_1 s_2) = m$.

Since the paths bsc and bs_1s_2c have the same length, we obtain, for the above variation,

$$\begin{aligned}\dot{T} &= 1 - \rho \cos \angle(ass_1) \\ &= 1 - \frac{(1+\cos\beta)(1+m^2)\cos\alpha}{(1-\cos\alpha)\sqrt{1+m^2}\sqrt{1+m^2\cos^2\alpha}} = f(m,\beta,\alpha).\end{aligned}$$

We require that $f(m,\beta,\alpha) \geqslant 0$, by the minimality of T. It is easy to see that $f(m,\beta,\alpha) \leqslant f(m,\alpha,\alpha) \leqslant (m,\phi,\phi)$ by Inequality (2). An analysis of $f(m,\phi,\phi)$ as a function of m shows that $f(m,\phi,\phi) < 0$ for all $m \leqslant 0.38091$, concluding the proof. □

Together, the previous four Lemmas prove the following theorem.

THEOREM 4. *If s is a degree 4 Steiner point in a gradient-constrained minimum Steiner tree T and if $m < 0.38$, then the labelling of s is (mm/mm).*

4. Geometric locations of Steiner points

By Theorems 3 and 4, there are six possible labellings at a Steiner point when $m < 0.38$. In this section we show how to use the variational argument to locate the Steiner point s in the gradient-constrained minimum Steiner tree T for these labellings, in terms of the adjacent vertices of T. Suppose the edges incident with s are sa, sb, sc (or sa, sb, sc, sd if s is of degree 4), where s has labelling $(g_a g_b/g_c)$ (or $(g_a g_b/g_c g_d)$ if s is of degree 4). For any point p, let $\mathcal{C}_p$, denote the cone generated by rotating a line through p with gradient m about the vertical line through p. So, $\mathcal{C}_p$ is a right vertical cone whose vertex is p.

We now consider each of the labellings in turn. If the labelling at s is (ff/f), then it is clear that the location of s is the same as in the unconstrained three dimensional Steiner problem. If the labelling is (mm/m), then s is determined by equations $g(sa) = g(sb) = g(sc) = m$. (If the equations have no solution, then there is no tree with labelling (mm/m).) Furthermore, we require that $\angle a'sb' \geqslant \gamma_{(\text{mm1})}$, by Lemma 5. Using these observations, it is easy to prove the following theorem.

THEOREM 5. *(i) If s has labelling (ff/f), then s is the point on the plane through a, b, c such that $\angle asb = \angle bsc = \angle csa = 2\pi/3$.*

(ii) If s has labelling (mm/m), then s is a point at which the three cones $\mathcal{C}_a$, $\mathcal{C}_b$ and $\mathcal{C}_c$ all intersect, and $\angle asb \geqslant$ arccos$(-1+m^2)/2 \geqslant \pi/2$.

For the remaining labellings, we employ the variational argument. Since any vector in space can be decomposed into a sum of three non-coplanar components, to prove s can be a Steiner point of a minimum tree T, by the variational argument, we need only to show $\dot{T} \geqslant 0$ for perturbations of s in three non-coplanar directions.

THEOREM 6. *If s has labelling (mm/b), then s is a point at the intersection of $\mathcal{C}_a$, $\mathcal{C}_b$ and the vertical plane through b, c.*

Proof. By Lemma 1 it is easy to see that $\dot{T} \geqslant 0$ when s moves either up or down, or in any horizontal direction. This proves the theorem. □

THEOREM 7. *Suppose s has labelling (ff/m), and let $\mathbf{v}$ be a horizontal vector tangent to $\mathcal{C}_c$ at s. Then s is a point on $\mathcal{C}_c$ such that $\angle(\mathbf{v}, \vec{sa}) = \angle(-\mathbf{v}, \vec{sb})$ and* $\cos\angle(\vec{cs}, \vec{sa}) + \cos\angle(\vec{cs}, \vec{sb}) = 1$.

Proof. Let $\mathcal{P}$ be the tangent plane of $\mathcal{C}_a$ at s. Because s has to lie on $\mathcal{C}_a$, we require that $\dot{T} = 0$ in two linearly independent directions on $\mathcal{P}$. Choosing $\vec{cs}$ as one direction and $\mathbf{v}$ as the other, the statement of the theorem easily follows. □

The final two theorems require the following lemma.

LEMMA 14. *Let p and q be points in three dimensional space such that $g(pq) > m$. Then $\mathcal{C}_p \cap \mathcal{C}_q$ is an ellipse.*

Proof. Suppose $\mathcal{C}_p$ is cut into two sections by a plane $\mathcal{P}$ (not containing p) whose slope is less than m. As is well known, the intersection $\mathcal{C}_p \cap \mathcal{P}$ is an ellipse. The proof follows from the symmetry of an ellipse. Make a copy of the section of $\mathcal{C}_p$, containing p, rotate it through an angle π about a horizontal line, and translate the rotated copy so that the two ellipses coincide. If the vertex of the rotated copy is at p^*, then this proves that $\mathcal{C}_p \cap \mathcal{C}_{p^*}$ is an ellipse. By varying the slope and position of $\mathcal{P}$, p^* can be any point satisfying $g(pp^*) > m$. □

THEOREM 8. *If s has labelling (fm/m), then s is a point lying on $\mathcal{C}_b \cap \mathcal{C}_c$, which is an ellipse, such that sa is perpendicular to this ellipse.*

Proof. Since sb and sc have gradient m and lie on two sides of $\mathcal{H}_s$, we have $g(bc) \geqslant m$. First suppose $g(bc) = m$, i.e., b, s and c are collinear. Because $\angle a'sb' > \gamma_{(\mathrm{mf1})} > 2\pi/3$, it follows that $\angle asb > \pi/2$. Therefore $\angle asc < \pi/2$, which contradicts Lemma 2.

Hence $g(bc) > m$, and $\mathcal{C}_b \cap \mathcal{C}_c$ is an ellipse (by Lemma 14). Now let $\mathbf{t}$ be a vector tangent to the ellipse $\mathcal{C}_b \cap \mathcal{C}_c$ at s. Note that $|sb| + |sc|$ is constant as s moves around the ellipse $\mathcal{C}_b \cap \mathcal{C}_c$. Since T is minimum, and since perturbing s in direction $\mathbf{t}$ leaves the labelling at s unchanged, it follows that $\dot{T}_{\mathbf{t}} = -\cos\angle(\vec{sa}, \mathbf{t}) = 0$. Hence sa is perpendicular to the ellipse. □

Finally, we consider the case where s has degree 4. If, in this case, we can partition the edges incident with s into two pairs, such that each pair lies in a vertical plane through s and has the property that the projections of the two edges onto $\mathcal{H}_s$ meet at an angle of π at s, then we say that the edges of s are *bi-vertically coplanar*. Note that if two m-edges incident with s lie on opposite sides of $\mathcal{H}_s$ and lie in the same vertical plane then those edges are collinear.

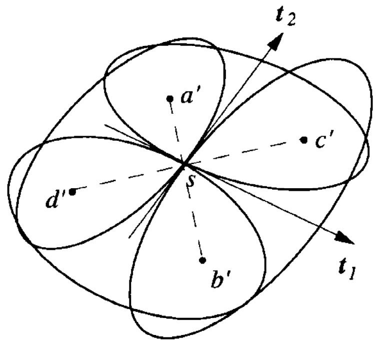

Figure 4. Projections of four intersecting cones.

THEOREM 9. *If s has labelling (mm/mm), then s is a point at the intersection of the four cones $\mathcal{C}_a, \mathcal{C}_b, \mathcal{C}_c, \mathcal{C}_d$, and furthermore the edges incident with s are bi-vertically coplanar.*

Proof. The first property is immediate; it remains to show that the edges incident with s are bi-vertically coplanar. First, assume no two edges incident with s are collinear. This implies that $g(ac) > m$ and $g(bd) > m$. Let $\mathcal{E}_{ac}, \mathcal{E}_{bd}$ represent the ellipses $\mathcal{C}_a \cap \mathcal{C}_c$ and $\mathcal{C}_b \cap \mathcal{C}_d$, respectively. By the labelling of s, $\mathcal{E}_{ac} \cap \mathcal{E}_{bd}$ is not empty. However, if $\mathcal{E}_{ac}$ intersects the region bounded by $\mathcal{C}_b \cup \mathcal{C}_d$ at any point not on $\mathcal{E}_{bd}$, then there exists a degree 4 Steiner point with an incident b-edge, contradicting Theorem 3.14. Hence, $\mathcal{E}_{ac}$ and $\mathcal{E}_{bd}$ are tangent. It follows that the projections of $\mathcal{E}_{ac}, \mathcal{E}_{bd}$ on $\mathcal{H}_s$, which are also ellipses, have a common horizontal tangent vector $\mathbf{t}_1$. Figure 4 depicts these projections and the tangent, where a', b', c', d' are the projections of a, b, c, d, respectively. It is not hard to see that a', c' are the foci of the projection of $\mathcal{E}_{ac}$, and b', d' are the foci of the projection of $\mathcal{E}_{bd}$ Hence,

$$\angle(\vec{a's}, \mathbf{t}_1) = \angle(\vec{c's}, -\mathbf{t}_1), \angle(\vec{b's}, \mathbf{t}_1) = \angle(\vec{d's}, -\mathbf{t}_1).$$

Similarly, the projections of other two ellipses $\mathcal{C}_a \cap \mathcal{C}_d$ and $\mathcal{C}_b \cap \mathcal{C}_c$ also have a common tangent vector $\mathbf{t}_2$. Consequently

$$\angle(\vec{a's}, \mathbf{t}_2) = \angle(\vec{d's}, -\mathbf{t}_2), \angle(\vec{b's}, \mathbf{t}_2) = \angle(\vec{c's}, -\mathbf{t}_2).$$

Combining the above four equations we conclude that a', s', b' are collinear and c', s, d' are collinear. This means that a, s and b lie in a vertical plane, and c, s and d also lie in a vertical plane. Therefore the configuration is bi-vertically coplanar.

Next suppose that two of the edges, say sa and sd, are collinear. If sb, sc are not collinear, then the ellipses $\mathcal{E}_{ac}, \mathcal{E}_{bd}$ exist, and a contradiction to the minimality of T can be obtained by a similar argument to that above. Hence, sb, sc are also collinear, and again the configuration is bi-vertically coplanar. □

5. Generalisations to higher dimensions

Although three-dimensional space is a natural context for these networks, in terms of their most obvious applications, it is worth noting that some of the results in this paper generalise to higher dimensions.

Let $p = (x_1, x_2, \ldots, x_{d-1}, z)$ be a point in d-dimensional space, and let the maximum gradient m be given. If we think of the z-axis as being a vertical axis, then we can define the length of the edge op in the gradient metric as follows:

$$|op|_g = \begin{cases} |op| = \sqrt{x_1^2 + x_2^2 + \cdots + x_{d-1}^2 + z^2}, & \text{if } g(op) \leqslant m; \\ |op|_v = (\sqrt{1 + m^{-2}})|z| & \text{if } g(op) \geqslant m. \end{cases}$$

If in Sections 2 and 3 we replace the word 'plane' by 'hyperplane', and define a horizontal hyperplane to be a hyperplane orthogonal to the z-axis, then all results in Section 2 still hold in d-dimensional space, and Lemmas 3 to 8 also hold. In particular, the lower bounds in the table in Section 3.1 also apply in d-dimensional space. However, because these are bounds on a hyperplane rather than a plane, they are much less useful, and the arguments in Lemma 9, Theorem 3 and Section 3.2 no longer apply. Indeed, even the question of determining the maximum degree of Steiner points in a d-dimensional gradient-constrained minimum Steiner tree is currently open.

Another interesting open question is to determine the nature of Steiner points in a gradient-constrained minimum Steiner tree (in three dimensions) if $m > 1$. For example, it seems likely that Theorem 3 is still true in this case, but the methods employed in this paper do not appear to be enough to easily prove such a result.

References

Brazil, M., Thomas, D.A. and Weng, J.F. (1998), Gradient-Constrained Minimal Steiner Trees, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* 40, 23–38.

Garey, M.R., Graham, R.L. and Johnson, D.S. (1977), The Complexity of Computing Steiner Minimal Trees, *SIAM J. Applied Mathematics* 32, 835–859.

Gilbert, E.N. and Pollak, H.O. (1968), Steiner Minimal Lees, *SIAM J. Applied Mathematics* 16, 1–29.

Hwang, F.K., Richard, D.S. and Winter, P. (1992), *The Steiner Tree Problem; Annals of Discrete Mathematics 53*, Elsvier Science Publishers B.V., Amsterdam.

Rubinstein, J.H. and Thomas, D.A. (1991), A Variational Approach to the Steiner Network Problem, *Annals of Operations Research* 33, 481–499.

Smith, W.D. (1992), How To Find Minimal Trees in Euclidean d-space, *Algorithmica* 7, 137–177.

Warme, D.M., Winter, P. and Zachariasen, M. (2000), Exact Algorithms for Plane Steiner Tree Problems: A Computational Study. In Du, D.-Z., Smith, J.M. and Rubinstein, J.H. *Advances in Steiner Trees*, Kluwer Academic Publishers, Boston, 81–116.