J. Eur. Math. Soc. 28, 2449–2472 (2026) DOI 10.4171/JEMS/1512

José A. Gálvez · Pablo Mira

Linearity of homogeneous solutions to degenerate elliptic equations in dimension three

Received July 27, 2022; revised July 19, 2023

Abstract. Given a linear elliptic equation $\sum a_{ij} u_{ij} = 0$ in $\mathbb{R}^3$, it is a classical problem to determine if its 1-homogeneous solutions u are linear. The answer is negative in general, by a construction of Martinez-Maure. In contrast, the answer is affirmative in the uniformly elliptic case, by a theorem of Han, Nadirashvili and Yuan, and it is a known open problem to determine the degenerate ellipticity condition on (a_{ij}) under which this theorem still holds. In this paper we solve this problem. We prove the linearity of u under the following degenerate ellipticity condition for (a_{ij}), which is sharp by Martinez-Maure's example: if $\mathcal{K}$ denotes the ratio between the largest and smallest eigenvalues of (a_{ij}), we assume $\mathcal{K}|_{\mathcal{O}}$ lies in L^1_{loc} for some connected open set $\mathcal{O} \subset \mathbb{S}^2$ that intersects any configuration of four disjoint closed geodesic arcs of length π in $\mathbb{S}^2$. Our results also give the sharpest possible version under which an old conjecture by Alexandrov, Koutroufiotis and Nirenberg (disproved by Martinez-Maure's example) holds.

Keywords: degenerate elliptic equation, homogeneous function, Alexandrov conjecture, saddle function.

1. Introduction

Let $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ be a (positively) 1-homogeneous solution to the linear equation

$$\sum_{i,j=1}^{3} a_{ij} u_{ij} = 0, \quad a_{ij} \in L^\infty(\mathbb{R}^3), \tag{1.1}$$

in $\mathbb{R}^3$, i.e., $u(\rho x) = \rho u(x)$ for all $\rho > 0$, $x \in \mathbb{R}^3$. Assume that (1.1) is elliptic, i.e.,

$$(a_{ij}(x)) \text{ is positive definite} \tag{1.2}$$

for every $x \in \mathbb{R}^3$. Note that the a_{ij} are not continuous. *Must then u be a linear function?*

José A. Gálvez: Departamento de Geometría y Topología, Instituto de Matemáticas IMAG, Universidad de Granada, 18071 Granada, Spain; jagalvez@ugr.es

Pablo Mira: Departamento de Matemática Aplicada y Estadística, Universidad Politécnica de Cartagena, 30203 Cartagena, Spain; pablo.mira@upct.es

Mathematics Subject Classification 2020: 35J70 (primary); 35R05, 53A05, 53C42 (secondary).

This is a classical question motivated by global surface theory. Using an equivalent formulation, Alexandrov [1] proved in 1939 that the answer is affirmative if u is real analytic, and conjectured that an affirmative answer should also hold in the general case [2, p. 352]. The validity of this conjecture remained elusive for a long time, until Martinez-Maure [11] constructed in 2001 a striking C^2 counterexample to it. Specifically, he proved the existence of a non-linear function $h \in C^2(\mathbb{S}^2)$ such that the *hedgehog* $\psi(\nu) := \nabla h(\nu) + h(\nu)\nu : \mathbb{S}^2 \to \mathbb{R}^3$ has negative curvature at its regular points. The 1-homogeneous extension u of h to $\mathbb{R}^3$ gives a counterexample to Alexandrov's conjecture (see Figure 1.1).

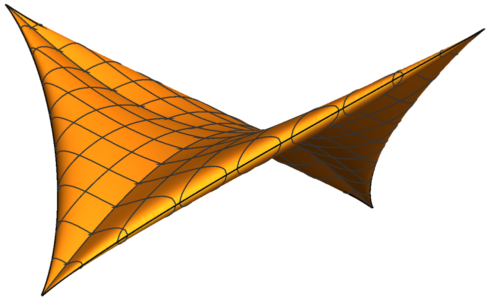

Fig. 1.1. Martinez-Maure's hegdehog $\nabla u(\mathbb{S}^2)$, where u solves (1.1)–(1.2). The preimage in $\mathbb{S}^2$ of each of the four *horns* of the example is a geodesic semicircle.

In contrast, in 2003 Han, Nadirashvili and Yuan [6] proved that the Alexandrov conjecture holds in the uniformly elliptic case. This solved an open problem by Safonov [20], who had proved that there exist α-homogeneous solutions to linear, uniformly elliptic equations in $\mathbb{R}^3$ for $\alpha \in (0,1)$ (see [19]). Specifically, if $0 < \lambda(x) \leq \Lambda(x)$ are the smallest and largest eigenvalues of $(a_{ij}(x))$, and we denote $\mathcal{K}(x) := \Lambda(x)/\lambda(x) \geq 1$, Han, Nadirashvili and Yuan imposed the condition

$$\mathcal{K} \in L^\infty(\mathbb{R}^3) \tag{1.3}$$

and proved the following remarkable result:

Theorem 1.1 ([6]). *Any 1-homogeneous solution $u \in W^{2,2}_{\text{loc}}(\mathbb{R}^3)$ to (1.1)–(1.3) is linear.*

An alternative proof of Theorem 1.1 was obtained in 2016 by Guan, Wang and Zhang [5], again under very weak regularity assumptions on u. For that, they treated the problem directly as a uniformly elliptic equation in $\mathbb{S}^2$, and gave an elegant argument using the Bers–Nirenberg unique continuation theorem. A different approach to Theorem 1.1 via Poincaré–Hopf index theory was given by the present authors and Tassi [4]. The problem of the linearity of homogeneous solutions to (1.1)–(1.2) is discussed in detail in the book [13] by Nadirashvili, Tkachev and Vlăduţ.

The uniform ellipticity assumption (1.3) in Theorem 1.1 cannot be weakened to *plain* ellipticity (1.2), by Martinez-Maure's example. A known natural open problem proposed by Guan, Wang and Zhang [5, Remark 8] is to establish what (non-trivial) degenerate ellipticity conditions on the coefficients a_{ij} are sufficient for Theorem 1.1 to hold, even when u is smooth.

In this paper we give an answer to this problem. We now explain our main results.

Let $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ be a 1-homogeneous solution to a linear equation (1.1). By homogeneity, u also satisfies (1.1) for the coefficients $a_{ij} = a_{ij}(x/|x|)$. For this reason, our hypotheses on (a_{ij}) will be directly viewed at points $x \in \mathbb{S}^2$. Instead of (1.3), we will just assume the considerably weaker condition that

$$\mathcal{K}|_{\mathcal{O}} \in L^1_{\rm loc}(\mathcal{O}) \tag{1.4}$$

holds for *some* connected open set $\mathcal{O} \subset \mathbb{S}^2$ with the property that it intersects any configuration of four disjoint *geodesic semicircles* (i.e. closed geodesic arcs of length π) in $\mathbb{S}^2$. Such a set $\mathcal{O}$ can be quite small. For instance, $\mathcal{O}$ can be chosen as any connected open set of $\mathbb{S}^2$ that contains an arbitrarily thin collar along a geodesic, $C_\gamma := \{x \in \mathbb{S}^2 : \langle x, \nu_0\rangle \in (0,\varepsilon)\}$ for some $\nu_0 \in \mathbb{S}^2$, $\varepsilon > 0$.

We prove the following result.

Theorem 1.2. *Any* 1*-homogeneous solution* $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ *to* (1.1), (1.2), (1.4) *is linear.*

The four semicircles condition imposed on $\mathcal{O}$ is sharp. Indeed, Martinez-Maure's example in [11] yields a 1-homogeneous function $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ such that D^2u is indefinite whenever it is non-zero, and so that $\{x \in \mathbb{S}^2 : D^2u(x) = 0\}$ agrees exactly with a certain configuration $\Gamma \subset \mathbb{S}^2$ of four disjoint geodesic semicircles. By the indefinite nature of D^2u, we can view u as a solution to some elliptic equation (1.1)–(1.2), and the related function $\mathcal{K}$ associated to the coefficients (a_{ij}) of this equation lies in $L^1_{\rm loc}(\mathcal{O})$ for any open set $\mathcal{O} \subset \mathbb{S}^2$ disjoint from Γ.

We can actually prove a more general version of Theorem 1.2, which holds under degenerate ellipticity conditions:

Theorem 1.3. *Let* $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ *be a* 1*-homogeneous solution to* (1.1). *Assume*

$$\begin{cases} \text{(i)} \ (a_{ij}(x)) \text{ is semi-positive definite for all } x \in \mathbb{R}^3, \\ \text{(ii)} \ (a_{ij}(x)) \text{ restricted to the plane } x^\perp \text{ is non-zero, for all } x \in \mathbb{R}^3 \setminus \{0\}, \\ \text{(iii)} \ (a_{ij}) \text{ is positive definite a.e. on } \mathcal{O}, \text{ and } \mathcal{K}|_{\mathcal{O}} \in L^1_{\rm loc}(\mathcal{O}), \end{cases} \tag{1.5}$$

where, in (iii), $\mathcal{O} \subset \mathbb{S}^2$ *is some connected open set that intersects any configuration of four disjoint geodesic semicircles. Then* u *is linear.*

Note that (i) extends (1.2) to the degenerate elliptic setting, and (ii) is needed in that general context to ensure that (1.1) is non-trivial when restricted to 1-homogeneous functions.

The proof of Theorem 1.3 is a blend of geometric and analytic arguments, and is presented in Section 2. The idea, following Alexandrov [1], is to show that $\nabla u(\mathbb{S}^2)$ reduces to a point, by analyzing the support planes in $\mathbb{R}^3$ of this compact set. In the uniformly elliptic case, Han, Nadirashvili and Yuan [6] used this idea and the maximum principle to show that $\nabla u(\mathbb{S}^2)$ is a point. In our situation given by (1.5), we will use instead the Stoilow factorization for planar mappings of finite distortion [3, 7]. However, the main difficulty of the proof is that we are not assuming that $\mathcal{K} \in L^1(\mathbb{S}^2)$, but only that its restriction to the possibly quite small set $\mathcal{O} \subset \mathbb{S}^2$ lies in $L^1_{\rm loc}$. In order to deal with this general situation, we will use an idea of Pogorelov [17]. He claimed a proof of Alexandrov's conjecture, something that is incorrect by the example in [11]. Pogorelov's argument was based on the deep idea of controlling the connected components in which some suitable planes of $\mathbb{R}^3$ divide the saddle graph Σ in $\mathbb{R}^3$ given by $z = u(x, y, 1)$. However, this is a delicate question, and the short argument presented in [17] has several errors in the way these connected components are handled (one of them was pointed out in [15]). Our proof of Theorem 1.3 springs from Pogorelov's brilliant idea, but we give a different, subtler argument that yields full control of the connected components mentioned above.

The term *Alexandrov conjecture* is often used in the literature in reference to a more general statement, in which (1.1) is allowed to be degenerate elliptic; see e.g. [11, 13, 15]. This conjecture admits several equivalent formulations, one of which is the following one, proposed in 1973 by Koutroufiotis and Nirenberg [8]:

The Alexandrov–Koutroufiotis–Nirenberg conjecture. *Any C^2 function v in $\mathbb{S}^2$ that satisfies* $\det(\nabla^2_{\mathbb{S}^2} v) \leq 0$ *at every point must be linear, i.e.,* $\nabla^2_{\mathbb{S}^2} v = 0$.

Here, the spherical Hessian $\nabla^2_{\mathbb{S}^2} v$ is defined by $\nabla^2_{\mathbb{S}^2} v(q) = (v_{ij}(q) + v(q)\delta_{ij})$, where v_{ij} are covariant derivatives with respect to a local orthonormal frame in $\mathbb{S}^2$ (see e.g. [5]). We say that $v \in C^2(\mathbb{S}^2)$ is a *saddle function* on $\mathbb{S}^2$ if it satisfies $\det(\nabla^2_{\mathbb{S}^2} v) \leq 0$. The conjecture is then that saddle functions on $\mathbb{S}^2$ are linear.

The support function h of Martinez-Maure's hedgehog in [11] gives a C^2 counterexample to this conjecture. Panina's construction in [15] provides C^∞ counterexamples, which are actually linear in large open regions of $\mathbb{S}^2$. Based on these results, Nadirashvili, Tkachev and Vladut proposed in [13, Conjecture 1.6.1] a *lopped version* of the conjecture, which can be rephrased as follows: *any C^2 saddle function on $\mathbb{S}^2$ is linear in some open set*.

This beautiful conjecture is open if v is at least of class C^3, but in the general C^2 category, one should reformulate it slightly. Indeed, Martinez-Maure's saddle function $h \in C^2(\mathbb{S}^2)$ is such that $\{q \in \mathbb{S}^2 : \nabla^2_{\mathbb{S}^2} h(q) = 0\}$ is the union of four disjoint geodesic semicircles; in particular, h is not linear on any open set of $\mathbb{S}^2$. Thus, the best possible *lopped* conjecture that can hold in the general C^2 case is that any saddle function $v \in C^2(\mathbb{S}^2)$ always satisfies $\nabla^2_{\mathbb{S}^2} v = 0$ along four disjoint geodesic semicircles. We will prove this exact result as a part of our proof of Theorem 1.3; see Section 3.

Theorem 1.4. *Let $v \in C^2(\mathbb{S}^2)$ satisfy* $\det(\nabla^2_{\mathbb{S}^2} v) \leq 0$. *Then* $\nabla^2_{\mathbb{S}^2} v = 0$ *along four disjoint geodesic semicircles of* $\mathbb{S}^2$.

Theorem 1.4 gives then the sharpest possible version for which the conjecture by Alexandrov, Koutroufiotis and Nirenberg is true, i.e., the sharpest possible *linearity* theorem for saddle C^2 functions in $\mathbb{S}^2$. We should note that Panina [16] claimed a very general statement that would have Theorem 1.4 as a particular case. However, the very short argument given in [16] is not correct; for instance, it relies on Pogorelov's incorrect study of the connected components problem. In Theorem 3.1 we will give an alternative formulation of Theorem 1.4, in the context of the Weingarten inequality $(\kappa_1 - c)(\kappa_2 - c) \leq 0$ for ovaloids of $\mathbb{R}^3$.

The Alexandrov conjecture has been linked by Mooney [12] to the existence of Lipschitz minimizers to functionals $\int F(\nabla u)\, dx$ in $\mathbb{R}^3$, with F strictly convex, that are C^1 except at a finite number of points. It has also been linked in [6,13,14] to the classification of 2-homogeneous solutions to elliptic Hessian equations $F(D^2u) = 0$ in $\mathbb{R}^3$. In particular, our results here might be of interest regarding the following conjecture in the book by Nadirashvili, Tkachev and Vlăduţ [13, Conjecture 1.6.3]: *a 2-homogeneous smooth solution u to a degenerate elliptic Hessian equation $F(D^2u) = 0$ in $\mathbb{R}^3$ must be a quadratic polynomial.*

2. Proof of Theorem 1.3

Let $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ be a 1-homogeneous solution to (1.1), where (1.5) holds. We will assume throughout the proof that u is not linear, i.e. D^2u is not identically zero on $\mathbb{R}^3$, and reach a contradiction. We will split the proof into several steps.

Step 1: Connection with quasiregular mappings. In this step we relate the conditions in (1.5) to the theory of planar mappings with finite distortion, in order to apply the Stoilow factorization by Iwaniec–Šverák [7] to our context.

Consider arbitrary Euclidean coordinates (x, y, z) in $\mathbb{R}^3$ centered at the origin, and define $h \in C^2(\mathbb{R}^2)$ by

$$h(x, y) := u(x, y, 1). \tag{2.1}$$

Note that $u(x, y, z) = zh(x/z, y/z)$ for all $z > 0$, by homogeneity. Then (see [6])

$$\nabla u(x, y, 1) = (h_x, h_y, h - xh_x - yh_y) \tag{2.2}$$

and

$$D^2u(x, y, 1) = \begin{pmatrix} h_{xx} & h_{xy} & -xh_{xx} - yh_{xy} \\ * & h_{yy} & -xh_{xy} - yh_{yy} \\ * & * & x^2h_{xx} + 2xyh_{xy} + y^2h_{yy} \end{pmatrix}. \tag{2.3}$$

From this and the invariance of (1.1) by Euclidean isometries we see that the restriction of (1.1) to points of the form $(x, y, 1)$ turns into a linear PDE for h,

$$A_{11}h_{xx} + 2A_{12}h_{xy} + A_{22}h_{yy} = 0. \tag{2.4}$$

Specifically, if we denote $\mathcal{A} := (a_{ij}(x, y, 1))$ and $\mathcal{M} := (A_{ij}(x, y))$, by (2.3), the coefficients of (2.4) are given for $i, j \in \{1, 2\}$ by

$$A_{ij} = w_i \cdot \mathcal{A} \cdot w_j^T, \tag{2.5}$$

where $w_1 := (1, 0, -x)$ and $w_2 := (0, 1, -y)$. In other words, the bilinear form defined by $\mathcal{M}$ is the restriction of the one given by $\mathcal{A}$ to the plane of $\mathbb{R}^3$ orthogonal to $(x, y, 1)$. By (i) and (ii) of (1.5), the matrix $\mathcal{M}$ is semi-positive definite and non-zero for all (x, y). This clearly implies by (2.4) that, for any (x, y),

$$h_{xx}h_{yy} - h_{xy}^2 \leq 0. \tag{2.6}$$

The converse of this property also holds, i.e., if $h(x, y)$ satisfies (2.6), it solves a degenerate elliptic equation (2.4) in $\mathbb{R}^2$, for suitable coefficients A_{ij}; see e.g. [18] for a similar argument in the elliptic case. Hence, if for *any* Euclidean linear coordinate system (x, y, z), the function $h(x, y)$ given by (2.1) satisfies (2.6), then u solves a linear equation (1.1) whose coefficients a_{ij} satisfy (i), (ii) of (1.5).

Consider the smallest and largest eigenvalues $\lambda \leq \Lambda$ among the three eigenvalues of $\mathcal{A}$ at $(x, y, 1)$, and let $\lambda_1 \leq \lambda_2$ denote the eigenvalues of $\mathcal{M}$. By (2.5), we have

$$0 \leq \lambda \leq \lambda_1 \leq \lambda_2 \leq \Lambda. \tag{2.7}$$

Choose next a point $\nu_0 \in \mathcal{O} \subset \mathbb{S}^2$ with positive z-coordinate, and express it as

$$\nu_0 = \frac{1}{\sqrt{1 + x_0^2 + y_0^2}}(x_0, y_0, 1). \tag{2.8}$$

Since (a_{ij}) is positive definite a.e. on $\mathcal{O}$ by (iii) of (1.5), the matrix $\mathcal{M}$ is positive definite a.e. around (x_0, y_0), by (2.7). Dividing by $A_{11} + A_{22}$, we can rewrite (2.4) as

$$2h_{w\bar{w}} + \mu h_{ww} + \bar{\mu} h_{\bar{w}\bar{w}} = 0 \tag{2.9}$$

around $w_0 := x_0 + iy_0$, where $w = x + iy$ and

$$\mu = \frac{A_{11} - A_{22} + 2iA_{12}}{A_{11} + A_{22}}. \tag{2.10}$$

Thus,

$$|\mu| = \frac{K_\mu - 1}{K_\mu + 1} < 1, \quad \text{where} \quad K_\mu := \frac{\lambda_2}{\lambda_1} \geq 1. \tag{2.11}$$

If we now write $f := h_w$, then by (2.9) and (2.11) we have

$$|f_{\bar{w}}| \leq |\mu|\,|f_w|, \quad |\mu| < 1 \text{ a.e. around } w_0. \tag{2.12}$$

Let us control next the dilatation quotient of f. If we denote

$$J(w, f) := |f_w|^2 - |f_{\bar{w}}|^2 \geq 0, \quad |Df(w)| := |f_w| + |f_{\bar{w}}|,$$

the dilatation quotient of f is given for any $w \in \mathbb{C}$ with $J(w, f) \neq 0$ by

$$K(w, f) = \frac{|Df(w)|^2}{J(w, f)} \geq 1.$$

At the points where $|Df(w)| = J(w, f) = 0$, we define $K(w, f) := 1$. Thus, $K(w, f)$ is defined a.e. around w_0, and by (2.11) and (2.12), at points with $J(w, f) \neq 0$ we have

$$K(w, f) \leq \frac{(|f_w| + |\mu|\,|f_w|)^2}{|f_w|^2 - |\mu|^2|f_w|^2} = \frac{(1 + |\mu|)^2}{1 - |\mu|^2} = K_\mu. \tag{2.13}$$

Hence, it follows from (2.7), (2.13) and our initial hypothesis $\mathcal{K}|_{\mathcal{O}} \in L^1_{\text{loc}}(\mathcal{O})$ (see (1.5) (iii)) that $K(w, f) \in L^1$ in a neighborhood of the point $w_0 = x_0 + iy_0 \in \mathbb{C}$. To see this, recall that by definition, $\mathcal{K} = \Lambda/\lambda$. Thus, we are in the conditions of the Iwaniec–Šverák theorem for degenerate elliptic quasiregular mappings ([7], see also [3]), which provides a Stoilow factorization for f in a neighborhood of w_0. This implies that, around w_0, f is either constant or an open mapping. We summarize this in the following assertion for later use:

Assertion 2.1. *If* $\nu_0 = \frac{1}{\sqrt{1+x_0^2+y_0^2}}(x_0, y_0, 1)$ *lies in* $\mathcal{O} \subset \mathbb{S}^2$*, then* ∇h *is either an open mapping or constant around* (x_0, y_0).

Step 2: Gradient mappings and support planes. In Steps 2 through 9 of the proof of Theorem 1.3, we will let $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ be a 1-homogeneous solution to a linear equation (1.1), and only assume that the coefficients a_{ij} of (1.1) satisfy the degenerate ellipticity conditions (i), (ii) of (1.5). That is, we will not use condition (iii) of (1.5).

By homogeneity, $D^2u(x)$ always has a trivial zero eigenvalue corresponding to the radial direction, for any $x \in \mathbb{R}^3 \setminus \{0\}$. Denote by $\mu_1(x) \leq \mu_2(x)$ the other two eigenvalues. These are also the eigenvalues of the spherical Hessian $\nabla^2_{\mathbb{S}^2} v$ of the function $v := u(x/|x|) \in C^2(\mathbb{S}^2)$ at the point $\eta = x/|x|$ (see e.g. [5]). Here, the spherical Hessian of v is defined by $\nabla^2_{\mathbb{S}^2} v(\eta) = (v_{ij}(\eta) + v(\eta)\delta_{ij})$, where v_{ij} are covariant derivatives with respect to a local orthonormal frame in $\mathbb{S}^2$. Then the property that the coefficients a_{ij} of (1.1) satisfy the degenerate ellipticity conditions (i), (ii) of (1.5) is equivalent to $\mu_1\mu_2 \leq 0$ everywhere, i.e. $\det(\nabla^2_{\mathbb{S}^2} v) \leq 0$ on $\mathbb{S}^2$. This follows from the argument indicated after equation (2.6).

Consider the *hedgehog* in $\mathbb{R}^3$ given by the restriction of the gradient mapping of u to the unit sphere, $\nabla u : \mathbb{S}^2 \to \mathbb{R}^3$. It can be regarded as a compact surface (with singularities) in $\mathbb{R}^3$ (see [9]). By compactness, $\nabla u(\mathbb{S}^2)$ admits a support plane in any direction, where by a *support plane* in the direction $\xi \in \mathbb{S}^2$ we mean a plane $\Pi_\xi \subset \mathbb{R}^3$ orthogonal to ξ that touches $\nabla u(\mathbb{S}^2)$ at some point q_ξ and $\langle \nabla u - q_\xi, \xi\rangle \leq 0$ on $\mathbb{S}^2$. Observe that $\nabla u(\mathbb{S}^2)$ cannot be constant, since D^2u is not identically zero. Thus, the convex hull C_0 of $\nabla u(\mathbb{S}^2)$ is not a single point of $\mathbb{R}^3$. Trivially, every support plane of C_0 must also intersect $\nabla u(\mathbb{S}^2)$. By a basic result in convex geometry, for almost every $\xi \in \mathbb{S}^2$ (in the measure sense), the support plane to C_0 in the ξ direction intersects C_0 in a unique point. Thus, for almost every direction $\xi \in \mathbb{S}^2$, the two associated support planes to ξ and $-\xi$ are different, and each of them intersects $\nabla u(\mathbb{S}^2)$ in a unique point.

Given arbitrary Euclidean coordinates (x, y, z) in $\mathbb{R}^3$, the hedgehog $\nabla u : \mathbb{S}^2 \to \mathbb{R}^3$ can be parametrized as the map in (2.2), for all $\nu \in \mathbb{S}^2$ with positive z-coordinate, that is,

$$\psi(x, y) := \nabla u(\nu) = (h_x, h_y, h - xh_x - yh_y), \tag{2.14}$$

where

$$\nu := \frac{(x, y, 1)}{\sqrt{1 + x^2 + y^2}}. \tag{2.15}$$

Recall that, by (2.6), $h_{xx}h_{yy} - h_{xy}^2 \leq 0$. Obviously, $\psi(x, y)$ is an immersion with unit normal ν at the points where $\det(D^2 h) < 0$. We call these points *regular points* of the hedgehog. We should note that, although ψ is at first only of class C^1, it can be easily checked using the inverse function theorem that any regular point q of ψ has a neighborhood $\mathcal{U} \subset \mathbb{R}^2$ such that $\psi(\mathcal{U})$ is a C^2 graph over an open set of its tangent plane at q. Thus, it makes sense to talk about the second fundamental form $II := -\langle d\psi, d\nu\rangle$ of (2.14) at regular points, and a computation from (2.14), (2.15) shows that

$$\begin{pmatrix} -\langle \psi_x, \nu_x\rangle & -\langle \psi_x, \nu_y\rangle \\ -\langle \psi_y, \nu_x\rangle & -\langle \psi_y, \nu_y\rangle \end{pmatrix} = \frac{-1}{\sqrt{1 + x^2 + y^2}} D^2 h(x, y). \tag{2.16}$$

In particular, the hedgehog has negative curvature at its regular points, and therefore such points cannot arise as contact points of $\nabla u(\mathbb{S}^2)$ with a support plane. Note that the hedgehog $\nabla u(\mathbb{S}^2)$ is regular at a point $\nu \in \mathbb{S}^2$ if and only if the two non-trivial eigenvalues $\mu_1 \leq \mu_2$ of $D^2 u(\nu)$ are non-zero (and so, necessarily, of opposite signs), i.e. if and only if $D^2 u(\nu)$ has rank 2.

Definition 2.2. We say that $p_0 \in \nabla u(\mathbb{S}^2)$ is a *Pogorelov point* if there exists a direction $\xi \in \mathbb{S}^2$ such that $\nabla u(\mathbb{S}^2) \cap \Pi_\xi = \{p_0\}$ and $p_0 \notin \{\nabla u(\xi), \nabla u(-\xi)\}$.

Assertion 2.3. *There exists a Pogorelov point of* $\nabla u(\mathbb{S}^2)$.

Proof. We first note that $\nabla u : \mathbb{S}^2 \to \mathbb{R}^3$ has a regular point. Indeed, otherwise we would have $\mu_1 \mu_2 = 0$ on $\mathbb{S}^2$. Thus, the function $f := u|_{\mathbb{S}^2}$ would satisfy $\det(\nabla^2_{\mathbb{S}^2} f) = 0$ everywhere on $\mathbb{S}^2$. By [8, Theorem 1], f would be linear on $\mathbb{S}^2$. So, u would also be linear, a contradiction.

Let then $\xi \in \mathbb{S}^2$ be a regular point of ∇u. By slightly varying ξ, we can assume additionally that each of the support planes Π_ξ and $\Pi_{-\xi}$ intersects $\nabla u(\mathbb{S}^2)$ in a unique point, say q_1 and q_2. As $\nabla u(\xi)$ cannot lie in any of these two planes (by regularity), either q_1 or q_2 is a Pogorelov point for $\nabla u(\mathbb{S}^2)$. ∎

Step 3: Setup for the rest of the proof. We fix from now on a Pogorelov point $p_0 \in \nabla u(\mathbb{S}^2)$, with associated direction $\xi \in \mathbb{S}^2$. Take $\nu_0 \in \mathbb{S}^2$ with $\nabla u(\nu_0) = p_0$. We consider Euclidean coordinates (x, y, z) with $\xi = (1, 0, 0)$ and $\nu_0 = (\nu_0^1, 0, \nu_0^3)$, with $\nu_0^3 > 0$. One should observe that ν_0 is not uniquely determined by ξ, since the subset $(\nabla u)^{-1}(p_0)$ of $\mathbb{S}^2$ might be large. As a matter of fact, we seek to show that it contains a geodesic semicircle. At this stage of the proof we will not require any additional information on ν_0, but in Step 8 we will discuss how to choose it in a convenient way.

Since $\xi = (1, 0, 0)$, the support plane Π_ξ leaves $\nabla u(\mathbb{S}^2)$ on its left side, i.e., Π_ξ is of the form $x = \mu_{\max}$, and

$$\mu_{\min} \leq u_x(p) \leq \mu_{\max} \quad \forall p \in \mathbb{S}^2, \tag{2.17}$$

for some $\mu_{\min}, \mu_{\max} \in \mathbb{R}$. The points $\nabla u(\pm\xi)$ do not lie in $x = \mu_{\max}$, since p_0 is a Pogorelov point. Thus, there exist $\mu_0 < \mu_{\max}$ and $\varepsilon > 0$ such that $u_x(p) \leq \mu_0$ for every $p \in B(\xi;\varepsilon) \cup B(-\xi;\varepsilon)$, where $B(a;\varepsilon)$ denotes a geodesic ball in $\mathbb{S}^2$ of center a and radius ε. By homogeneity, $u_x(x,y,z) \leq \mu_0$ on a subset of $\mathbb{R}^3$ of the form $x^2 \geq \delta(y^2+z^2)$ for some $\delta = \delta(\varepsilon) > 0$.

From now on, let Σ be the entire saddle graph in $\mathbb{R}^3$ given by $z = h(x,y)$, where h is defined by (2.1); note that Σ has non-positive curvature at every point, by (2.6). By (2.2) and the compactness of $\nabla u(\mathbb{S}^2)$, we see that ∇h is uniformly bounded in $\mathbb{R}^2$. Moreover, by (2.17), (2.2) and the definition of μ_0, we have

$$\mu_{\min} \leq h_x(x,y) \leq \mu_{\max} \quad \forall (x,y) \in \mathbb{R}^2, \tag{2.18}$$

and

$$h_x(x,y) \leq \mu_0 < \mu_{\max} \quad \forall (x,y) \in \mathbb{R}^2 \text{ with } x^2 \geq \delta(y^2+1). \tag{2.19}$$

We will denote by Ω^+ (for $x > 0$) and Ω^- (for $x < 0$) the two connected components of the set $x^2 \geq \delta(y^2+1)$ in $\mathbb{R}^2$. Also, note that

$$h_x(x_0,0) = \mu_{\max}, \quad \text{where} \quad \nu_0 = (\nu_0^1, 0, \nu_0^3) = \frac{(x_0,0,1)}{\sqrt{1+x_0^2}}. \tag{2.20}$$

We will frequently use the notation

$$\varphi(x,y) := (x,y,h(x,y)). \tag{2.21}$$

Step 4: A transverse line L_n^* to $\Sigma \cap \{y=0\}$ with almost maximum slope. Consider a plane Π given by $z = P(x,y) := ax + by + c$ with $a > \mu_0$. Then, for any $y_0 \in \mathbb{R}$, by (2.19) and $a > \mu_0$, the line $L_{y_0} \equiv \Pi \cap \{y = y_0\}$ is above (resp. below) the graph $z = h(x,y_0)$ as $x \to \infty$ (resp. $x \to -\infty$). Therefore, there exist points $x_1(y_0) \leq x_2(y_0)$ such that

$$h(x,y_0) > P(x,y_0) \text{ for } x < x_1(y_0), \quad h(x,y_0) < P(x,y_0) \text{ for } x > x_2(y_0). \tag{2.22}$$

In particular, there exist points $(x_1,0) \in \Omega^-$ and $(x_2,0) \in \Omega^+$ such that $h(x,0) > P(x,0)$ for all $x \leq x_1$, and $h(x,0) < P(x,0)$ for all $x \geq x_2$.

Assertion 2.4. *There exist continuous curves $x = \alpha^-(y)$, $x = \alpha^+(y)$ in $\mathbb{R}^2$, depending on the initial plane Π, such that $\alpha^-(0) = x_1$, $\alpha^+(0) = x_2$, and*

$$h(\alpha^-(y),y) > P(\alpha^-(y),y), \quad h(\alpha^+(y),y) < P(\alpha^+(y),y), \quad \forall y \in \mathbb{R}. \tag{2.23}$$

Proof. Take $\bar{a} \in (\mu_0, a)$ and denote by $\bar{\mu}_{\min}$ the minimum value of h_y in $\mathbb{R}^2$. Choose $\lambda < 0$ so that the half-line $\mathcal{L}_\lambda \subset \mathbb{R}^2$ given by $x = x_1 + \lambda y$ for $y \geq 0$ is contained in Ω^-. We can obviously choose λ so that additionally $(a-\bar{a})\lambda < \bar{\mu}_{\min} - b$ (see Figure 2.1). Then $\alpha^-(y) := x_1 + \lambda y$ satisfies the first inequality in (2.23) for all $y \geq 0$; indeed, if $(x,y) \in \mathcal{L}_\lambda$, then integrating ∇h along $\mathcal{L}_\lambda$, and using $h(x_1,0) > P(x_1,0)$ together with

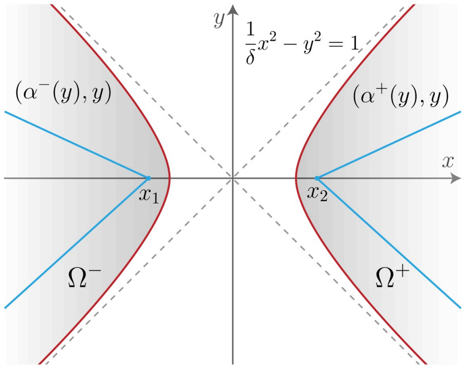

Fig. 2.1. The curves $(\alpha^{\pm}(y), y)$ in $\mathbb{R}^2$.

the previous inequalities, we have

$$h(x, y) > h(x_1, 0) + (\bar{a}\lambda + \overline{\mu}_{\min})y \ > P(x_1, 0) + (a\lambda + b)y = P(x, y).$$

The first inequality for $y < 0$, and the second inequality in (2.23) are obtained similarly. This proves Assertion 2.4. ■

Remark 2.5. Observe that if we consider the continuous curves $x = \alpha^{\pm}(y)$ defined in Assertion 2.4 with respect to the plane Π, then all points $\varphi(x, y) \in \Sigma$ where $x < \alpha^-(y)$ (resp. $x > \alpha^+(y)$) lie above (resp. below) Π. To see this, it suffices to realize that the proof of Assertion 2.4 also holds if instead of $(x_1, 0) \in \Omega^-$ we consider as initial point of $x = \alpha^-(y)$ any point $(x, 0)$ with $x < x_1$ (and a similar argument for $x > x_2$ with $(x_2, 0) \in \Omega^+$).

Take next a sequence $\{\mu_n\}_n \to \mu_{\max}$ with $\mu_n \in (\mu_0, \mu_{\max})$ for all n. Consider the line L_n in the vertical plane $y = 0$ given by $z = \mu_n(x - x_0) + h(x_0, 0)$. Note that L_n intersects $\Sigma_0 := \Sigma \cap \{y = 0\}$ transversely at $\varphi(x_0, 0)$, by (2.20). More specifically, since $\mu_n < \mu_{\max}$, we see that Σ_0 lies below L_n in the plane $y = 0$ for values of $x < x_0$ near x_0, and above L_n for $x > x_0$ near x_0. Moreover, it is clear from (2.22) that Σ_0 lies above (resp. below) L_n as $x \to -\infty$ (resp. as $x \to \infty$). This shows, in particular, that the planar set $\Sigma_0 \setminus L_n$ has at least four connected components, each homeomorphic to an open interval.

By the transversality of Σ_0 and L_n at $\varphi(x_0, 0)$, there exists some $\varepsilon > 0$ such that $h_x(x, 0) > \mu_n$ and $\varphi(x, 0) \notin L_n$ for all $x \neq x_0$ with $|x - x_0| < \varepsilon$. By Sard's theorem, if necessary, we can make a small parallel translation of L_n in the plane $y = 0$ to obtain a new straight line L_n^* which might not pass through $(x_0, 0, h(x_0, 0))$ anymore, but which intersects Σ_0 transversely at every intersection point. Specifically, we may take L_n^* so that it contains a point $\varphi(x_0^*, 0)$ with $|x_0 - x_0^*| < \varepsilon$, and so that the distance between $\varphi(x_0^*, 0)$ and $\varphi(x_0, 0)$ is smaller than $1/n$. Here, $x_0^* = x_0^*(n)$, i.e., x_0^* depends on n.

Note that, by (2.22), L_n^* lies either above or below Σ_0 as $x \to \infty$ or $x \to -\infty$. Then, by transversality, $\Sigma_0 \setminus L_n^*$ has a finite number of connected components. By the above

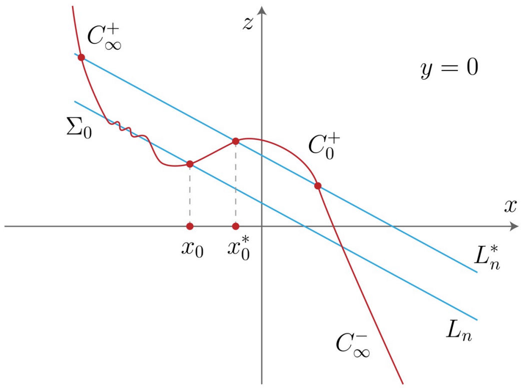

Fig. 2.2. The connected components C_∞^+, C_∞^- and C_0^+.

arguments, we also know that the number of such connected components is at least 4, and that $\varphi(x_0^*, 0)$ lies at the common boundary of two such *bounded* connected components. We will use the following notations for some special connected components of $\Sigma_0 \setminus L_n^*$ (see Figure 2.2).

- C_∞^+ is the unbounded component that lies strictly above L_n^*.
- C_∞^- is the unbounded component that lies strictly below L_n^*.
- C_0^+ is the *bounded* component that lies strictly above L_n^*, and has $\varphi(x_0^*, 0)$ as a boundary point.

Observe that C_∞^+ lies in the set $\{x < x_0^*\}$, while C_∞^- and C_0^+ lie in $\{x > x_0^*\}$.

Step 5: Study of the intersection of Σ with the sheaf of planes containing L_n^*. Let us now fix the straight line L_n^*, and consider all the planes in $\mathbb{R}^3$, excluding $y = 0$, that contain L_n^*. They are given by

$$z = P_b(x, y) = \mu_n(x - x_0^*) + by + h(x_0^*, 0) \tag{2.24}$$

for each $b \in \mathbb{R}$. Let Π_b be the plane determined by b. We next study $\Sigma \cap \Pi_b$.

Fix some point $q_0 \in C_0^+$. Let I_n (resp. J_n) denote the set of values $b \in \mathbb{R}$ for which q_0 can be joined to a point $\varphi(x, y) \in \Sigma$ with $y > n$ (resp. with $y < -n$) through an arc contained in $\Sigma \setminus (\Pi_b \cup C_\infty^+)$. The statement of the next assertion uses the fact that $\overline{\mu}_{\min} \leq h_y(x, y) \leq \overline{\mu}_{\max}$ for appropriate constants, for all $(x, y) \in \mathbb{R}^2$. It states that for any $n \in \mathbb{N}$ there exists a plane Π_{b_n} such that we can find an arc in Σ joining q_0 to points $\varphi(x, y)$ where $y > n$ and $y < -n$, while avoiding both Π_{b_n} and the connected component C_∞^+.

Assertion 2.6. *There exists $b_n \in I_n \cap J_n$ with $\overline{\mu}_{\min} \leq b_n \leq \overline{\mu}_{\max}$.*

Proof. Write $q_0 = \varphi(q_0^1, 0)$. By construction, q_0 lies above L_n^*. If we choose $b \leq \overline{\mu}_{\min}$, then $\varphi(q_0^1, y) \in \Sigma$ lies above Π_b for all $y > 0$. Since $\varphi(q_0^1, 0) \notin C_\infty^+$, this means that

$\overline{\mu}_{\min} \in I_n$. By the same argument, $\overline{\mu}_{\max} \in J_n$. Thus I_n and J_n are non-empty, and they both intersect the closed interval $[\overline{\mu}_{\min}, \overline{\mu}_{\max}]$.

We check next that I_n is open. Let $b_0 \in I_n$. Then there exists an arc in $\Sigma \setminus (\Pi_{b_0} \cup C_\infty^+)$ joining q_0 to a point $p = \varphi(x, y)$ with $y > n$. By compactness, this arc lies above Π_{b_0} at a certain distance $d > 0$. In particular, for values of b near b_0, this arc also avoids $\Pi_b \cup C_\infty^+$. Therefore, I_n is open. By the same argument, J_n is open.

Finally, we prove that $I_n \cup J_n = \mathbb{R}$, which, together with the already proved properties and the fact that $[\overline{\mu}_{\min}, \overline{\mu}_{\max}]$ is connected, yields Assertion 2.6. Arguing by contradiction, assume that there exists $b \in \mathbb{R} \setminus (I_n \cup J_n)$. We are going to prove next that the (open) connected component of $\Sigma \setminus \Pi_b$ that contains q_0, which we will denote by $\Sigma(C_0^+)$, is *bounded*. This will contradict the fact that Σ is a saddle graph.

To do this, we start by fixing some notation and making some elementary comments. First, note that $\Sigma(C_0^+)$ lies above Π_b, since $q_0 \in C_0^+$. Also, denote by $\Sigma(C_\infty^+)$ the connected component of $\Sigma \setminus \Pi_b$ that contains C_∞^+. By Remark 2.5, if we consider the continuous curves $x = \alpha^\pm(y)$ defined in Assertion 2.4 with respect to the plane Π_b, then all points $\varphi(x, y) \in \Sigma$ where $x < \alpha^-(y)$ (resp. $x > \alpha^+(y)$) lie above (resp. below) Π_b. In this way, the curve $\Gamma^- := \{\varphi(\alpha^-(y), y) \in \Sigma : y \in \mathbb{R}\}$ is contained in $\Sigma(C_\infty^+)$.

First of all, we prove that every point $\varphi(x, y)$ of $\Sigma(C_0^+)$ satisfies $y \in [-n, n]$. Indeed, otherwise there would exist an arc γ in Σ, starting at q_0, that reaches either $\{y < -n\}$ or $\{y > n\}$, and that intersects C_∞^+, since $b \notin I_n \cap J_n$. Let $\overline{z}_0$ denote the first point where γ touches C_∞^+. Then a neighborhood of $\overline{z}_0$ trivially lies in $\Sigma(C_\infty^+)$ (see Figure 2.3). In particular, $\Sigma(C_\infty^+) = \Sigma(C_0^+)$.

Let $z_0 = \varphi(z_0^1, z_0^2)$ be a point of that neighborhood that also lies in the interior of the arc of γ between q_0 and $\overline{z}_0$. Assume that $z_0^2 < 0$ (the argument for $z_0^2 > 0$ is similar).

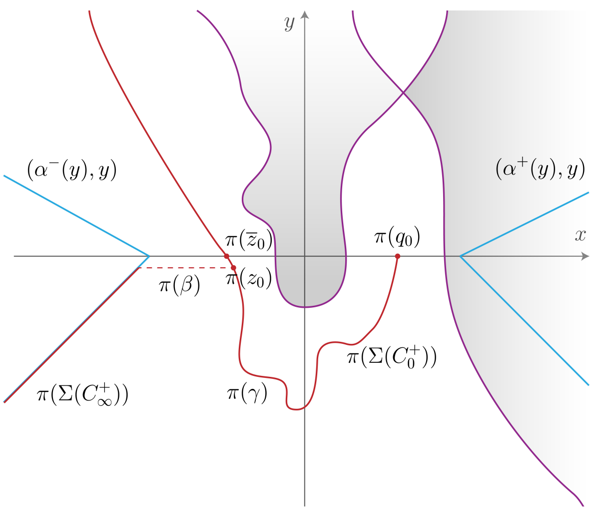

Fig. 2.3. Proof that $\Sigma(C_0^+)$ lies in the slab of $\mathbb{R}^3$ given by $|y| \leq n$. In the figure, π denotes the projection $\pi(x, y, z) = (x, y)$ onto the x, y-plane.

Then we can join the curve $\Gamma^- \subset \Sigma$ defined above to the point z_0 along an arc β contained in $\Sigma(C_\infty^+)$ and so that every point of the arc β has negative y-coordinate (see Figure 2.3). This implies that β does not touch C_∞^+, which is contained in the $y = 0$ plane. Now, the union of the arc of γ joining q_0 to z_0, the arc β, and a suitable arc of the curve Γ^- produces an arc in $\Sigma(C_\infty^+)$ that avoids C_∞^+ and joins q_0 to a point in $\Sigma \cap \{y < -n\}$ (see Figure 2.3). This would mean that $b \in J_n$, a contradiction. Thus, $\Sigma(C_0^+)$ lies in the slab of $\mathbb{R}^3$ given by $|y| \le n$, as desired.

Recall that all points of the form $\varphi(\alpha^+(y), y)$ lie below Π_b, by Assertion 2.4. Since all points $\varphi(x, y) \in \Sigma(C_0^+)$ satisfy $|y| \le n$ and lie above Π_b, we conclude that their x-coordinates are bounded from above by $\alpha^+(y)$.

On the other hand, assume that there exists an arc in $\Sigma(C_0^+)$ that joins q_0 to a point of the form $\varphi(\alpha^-(y), y)$. By Assertion 2.4, we have $\Sigma(C_0^+) = \Sigma(C_\infty^+)$. But now, as $\Sigma(C_\infty^+)$ has points of the form $\varphi(x, y)$ with $|y|$ arbitrarily large, we contradict the fact that $\Sigma(C_0^+)$ lies in the slab $|y| \le n$.

We have thus proved that $\Sigma(C_0^+)$ is contained in the compact set

$$\{\varphi(x, y) : \alpha^-(y) \le x \le \alpha^+(y),\ |y| \le n\} \subset \Sigma.$$

Thus $\Sigma(C_0^+)$ is a bounded connected component of $\Sigma \setminus \Pi_b$, in contradiction with the saddleness of Σ. This proves Assertion 2.6. ∎

Step 6: Study of the intersection of Σ with the limit plane Π_∞. For each n, let $b_n \in \mathbb{R}$ be given by Assertion 2.6, and consider the associated plane Π_{b_n} given by (2.24) for $b = b_n$. Since $\overline{\mu}_{\min} \le b_n \le \overline{\mu}_{\max}$, we have $\{b_n\}_n \to b_\infty \in [\overline{\mu}_{\min}, \overline{\mu}_{\max}]$ up to a subsequence. Since $|x_0^* - x_0| < 1/n$ and $\{\mu_n\}_n \to \mu_{\max}$, the planes Π_{b_n} converge to the limit plane

$$\Pi_\infty : z = P_\infty(x, y) := \mu_{\max}(x - x_0) + b_\infty y + h(x_0, 0), \tag{2.25}$$

which passes through $\varphi(x_0, 0) \in \Sigma$ with maximum slope $\mu_{\max}$ in the x-direction.

We next study $\Sigma \cap \Pi_\infty$. Fix any $y_0 \in \mathbb{R}$. Then, taking $\Pi = \Pi_\infty$ in Assertion 2.4, it is a consequence of (2.23) that the curve $\Sigma \cap \{y = y_0\}$ intersects Π_∞.

Assertion 2.7. *Either for all $y_0 \ge 0$, or for all $y_0 \le 0$, there exist $x^-(y_0) \le x^+(y_0)$ such that*

$$\Pi_\infty \cap \Sigma \cap \{y = y_0\} = \{\varphi(x, y_0) : x \in [x^-(y_0), x^+(y_0)]\}.$$

Moreover, $h_x(x, y_0) = \mu_{\max}$ for every $x \in [x^-(y_0), x^+(y_0)]$, and Σ lies above Π_∞ (resp. below Π_∞) when $x < x^-(y_0)$ (resp. $x > x^+(y_0)$).

Proof. Fix $y_0 \in \mathbb{R}$. We distinguish two possible situations.

Case 1: $\Pi_\infty \cap \Sigma \cap \{y = y_0\}$ is not a unique point. In that case, given two points $\varphi(x_1, y_0)$, $\varphi(x_2, y_0)$ of that intersection, we find that all points of the form $\varphi(x, y_0)$ with $x \in [x_1, x_2]$ also lie in $\Pi_\infty \cap \Sigma \cap \{y = y_0\}$. This follows since $h_x(x, y_0) \le \mu_{\max} = (P_\infty)_x(x, y_0)$ and $h(x_i, y_0) = P_\infty(x_i, y_0)$ for $i = 1, 2$. Thus, if $\Pi_\infty \cap \Sigma \cap \{y = y_0\}$ has at least two points, there exist $x^-(y_0) < x^+(y_0)$ such that

- $\varphi(x, y_0)$ lies above Π_∞ for all $x < x^-(y_0)$.
- $\varphi(x, y_0)$ lies below Π_∞ for all $x > x^+(y_0)$.
- $\varphi(x, y_0) \in \Pi_\infty$ for all $x \in [x^-(y_0), x^+(y_0)]$.

Note that $h_x(x, y_0) = \mu_{\max}$ for all (x, y_0) in the last situation above. So, the statement of Assertion 2.7 holds for every $y_0 \in \mathbb{R}$ such that $\Pi_\infty \cap \Sigma \cap \{y = y_0\}$ is not a unique point. No sign assumption is needed here for y_0.

Case 2: $\Pi_\infty \cap \Sigma \cap \{y = y_0\}$ is a unique point. This situation is subtler, and needs additional control on the intersections $\Sigma \cap \Pi_{b_n}$ before passing to the limit.

Let b_n be given by Assertion 2.6, with $\{b_n\}_n \to b_\infty$. As $b_n \in I_n \cap J_n$, there exists an arc $\gamma^+ = \gamma^+(n)$ in Σ that lies above Π_{b_n}, does not intersect C_∞^+, and whose endpoints have y-coordinate equal to n and $-n$, respectively. Since L_n^* intersects $\Sigma_0 = \Sigma \cap \{y = 0\}$ transversely in a finite number of points, there obviously exists a unique connected component C_1^- of $\Sigma_0 \setminus L_n^*$ that has as a boundary point the unique boundary point of C_∞^+, and lies *below* Π_{b_n} (since C_1^- lies below L_n^*). As γ^+ lies above Π_{b_n} and does not intersect C_∞^+, we easily deduce that every point in $\gamma^+ \cap \Sigma_0$ is of the form $\varphi(x, 0)$ with $x > \sup\{x : \varphi(x, 0) \in C_1^-\}$. Obviously, $\gamma^+ \cap \Sigma_0$ is non-empty since γ^+ goes from $y = n$ to $y = -n$.

Let $\Sigma(C_1^-)$ denote the connected component of $\Sigma \setminus \Pi_{b_n}$ that contains C_1^- (thus, it lies below Π_{b_n}). For each n, let $\alpha_n^+(y), \alpha_n^-(y)$ be the functions $\alpha^+(y), \alpha^-(y)$ defined by Assertion 2.4 with respect to $\Pi = \Pi_{b_n}$. Then $\Sigma(C_1^-)$ must intersect either $\Sigma \cap \{y = n\}$ or $\Sigma \cap \{y = -n\}$; indeed, otherwise, $\Sigma(C_1^-)$ would be a connected component contained in a compact region of Σ bounded by γ^+, $\Sigma \cap \{y = \pm n\}$ and $\{\varphi(\alpha_n^-(y), y) : y \in \mathbb{R}\}$, and this contradicts the saddleness of Σ.

In this way, we can take an arc $\gamma^- = \gamma^-(n)$ contained in $\Sigma(C_1^-)$ that joins a point of C_1^- to a point q_n of Σ with y-coordinate n or $-n$. Up to a subsequence of $\{b_n\}_n$, we can assume that one of these two situations holds for all n. For definiteness, *we will assume that the y-coordinate of q_n is equal to n for all n.*

Then, obviously, any plane $\{y = y_0\}$ with $y_0 \in [0, n]$ is intersected by the curves γ^-, γ^+, and $\{\varphi(\alpha_n^+(y), y) : y \in \mathbb{R}\}$. Using again $\gamma^+ \cap C_\infty^+ = \emptyset$, we deduce the existence of points $x_1 < x_2 < \alpha_n^+(y_0)$, with each x_1, x_2 depending on y_0 and n, such that

$$\varphi(x_1, y_0) \in \gamma^-, \quad \varphi(x_2, y_0) \in \gamma^+.$$

Therefore, there exist $x_3 \in (x_1, x_2)$ and $x_4 \in (x_2, \alpha^+(y))$ such that both $\varphi(x_3, y_0)$ and $\varphi(x_4, y_0)$ lie in $\Sigma \cap \Pi_{b_n} \cap \{y = y_0\}$. Moreover, since the line $\Pi_{b_n} \cap \{y = y_0\}$ has slope μ_n and $\varphi(x_2, y_0)$ lies above Π_{b_n} with $x_2 \in (x_3, x_4)$, by the mean value theorem there must exist $x_5 \in (x_3, x_4)$ such that $\varphi(x_5, y_0)$ lies above Π_{b_n} and $h_x(x_5, y_0) = \mu_n$.

From now on, we denote $s_n(y_0) := x_3 < t_n(y_0) := x_5$. Thus, for every $n \in \mathbb{N}$ and every $y \in [0, n]$, we have

- $\varphi(s_n(y), y) \in \Sigma \cap \Pi_{b_n}$,
- $\varphi(t_n(y), y)$ lies above Π_{b_n}, and $h_x(t_n(y), y) = \mu_n$.

We now pass to the limit, and show that the statement of Assertion 2.7 holds for every $y_0 \geq 0$; if we had assumed that the y-coordinate of q_n is $-n$, the next argument would show that Assertion 2.7 holds for every $y_0 \leq 0$.

Fix then $y_0 \geq 0$. By our hypothesis in the present Case 2 and (2.22), there exists a certain value $x(y_0)$ such that $\varphi(x, y_0)$ lies above Π_∞ for all $x < x(y_0)$, and below Π_∞ for all $x > x(y_0)$.

Take $(c(y_0), y_0) \in \Omega^-$ with $c(y_0) < x(y_0)$. Since $\{\Pi_{b_n}\}_n \to \Pi_\infty$, there exists $n_0 \in \mathbb{N}$ such that $\varphi(c(y_0), y_0)$ lies above Π_{b_n} for every $n \geq n_0$. Now, as $(c(y_0), y_0) \in \Omega^-$, we see by (2.19) and $\mu_0 < \mu_n$ that $\varphi(x, y_0)$ lies above Π_{b_n} for all $x < c(y_0)$ and all $n \geq n_0$. In particular, $c(y_0) < s_n(y_0) < t_n(y_0)$ for all n large enough, since $\varphi(s_n(y_0), y_0) \in \Pi_{b_n}$.

Arguing in a similar way for large positive values of x, we deduce that the sequences $\{s_n(y_0)\}_n$ and $\{t_n(y_0)\}_n$ are bounded. Thus, up to a subsequence, $\{\varphi(s_n(y_0), y_0)\}_n \to \varphi(x(y_0), y_0)$, by uniqueness of the point $\varphi(x(y_0), y_0)$.

On the other hand, the points $\varphi(t_n(y_0), y_0)$ converge to some point that is not below Π_∞, since $\varphi(t_n(y_0), y_0)$ lies above Π_{b_n} and $\{\Pi_{b_n}\}_n \to \Pi_\infty$. But as $t_n(y_0) > s_n(y_0) \to x(y_0)$ and $\varphi(x, y_0)$ lies below Π_∞ for all $x > x(y_0)$, we deduce that $\{t_n(y_0)\}_n \to x(y_0)$. In particular, $h_x(x(y_0), y_0) = \mu_{\max}$, since $h_x(t_n(y_0), y_0) = \mu_n$. This proves Assertion 2.7 in Case 2, and thus completes the proof. ■

Step 7: Existence of a half-line of maximal slope in $\Sigma \cap \Pi_\infty$. In this step, we show that the set $\Sigma \cap \Pi_\infty$ contains some half-line $\mathcal{L}^*$, and moreover $h_x(x, y) = \mu_{\max}$ for all $(x, y) \in \mathbb{R}^2$ with $\varphi(x, y) \in \mathcal{L}^*$.

To start, assume for definiteness that Assertion 2.7 holds for $y_0 \geq 0$ (the case $y_0 \leq 0$ is treated analogously). Let $\mathcal{J}$ be the set of values $y_0 \in \mathbb{R}$ such that $\Pi_\infty \cap \Sigma \cap \{y = y_0\}$ is a unique point $\varphi(x(y_0), y_0)$ at which $h_x(x(y_0), y_0) < \mu_{\max}$. Then, by Assertion 2.7, we have $\mathcal{J} \subset (-\infty, 0)$. Let $\delta_0 \leq 0$ denote the supremum of $\mathcal{J}$, where we use the convention that $\delta_0 = -\infty$ if $\mathcal{J}$ is empty.

It follows from Assertion 2.7 that there exist two (at first, maybe non-continuous) functions $x^-(y) < x^+(y)$, defined for all $y > \delta_0$, and such that the following properties hold:

$$\begin{cases} \text{(i)} \;\; h(x,y) > P_\infty(x,y) \text{ if } x < x^-(y), \\ \text{(ii)} \;\; h(x,y) < P_\infty(x,y) \text{ if } x > x^+(y), \\ \text{(iii)} \;\; h(x,y) = P_\infty(x,y) \text{ and } h_x(x,y) = \mu_{\max} \text{ if } x \in [x^-(y), x^+(y)]. \end{cases} \tag{2.26}$$

To see this, one should recall that our conclusion in Case 1 in the proof of Assertion 2.7 holds for all $y_0 \in \mathbb{R}$, not only for $y_0 \geq 0$ or $y_0 \leq 0$.

Assertion 2.8. *The sets*

$$D^- = \{(x,y) \in \mathbb{R} \times (\delta_0, \infty) : h(x,y) > P_\infty(x,y)\},$$
$$D^+ = \{(x,y) \in \mathbb{R} \times (\delta_0, \infty) : h(x,y) < P_\infty(x,y)\}$$

are open convex subsets of $\mathbb{R}^2$. *In particular,* $x^+(y)$, $x^-(y)$ *are continuous.*

Proof. We will prove the result for D^+; the argument for D^- is analogous. Let $p_i := (x_i, y_i) \in D^+$, $i = 1, 2$. If $y_1 = y_2$, the segment that joins the two points lies in D^+, by (ii) of (2.26).

Assume that $y_1 \neq y_2$, and that the segment that joins p_1 to p_2 is not contained in D^+. As $x^+(y) < \alpha^+(y)$ and $\alpha^+(y)$ is continuous, we can take a translation of $\overline{p_1 p_2}$ in the positive x-direction so that the resulting segment is contained in D^+. Next, translate that segment back in the negative x-direction, until reaching a first contact point with the set $D_0 := \{(x, y) : h(x, y) = P_\infty(x, y)\}$. We will denote the resulting segment by $\mathcal{S}_0$.

Note that the endpoints of $\mathcal{S}_0$ lie in D^+, and D^+ is connected by (i)–(iii) of (2.26). Let γ denote a compact arc in D^+ joining the endpoints of $\mathcal{S}_0$. Then there exists $\varepsilon > 0$ such that $h \leq P_\infty - \varepsilon$ for any point of γ. In this way, if we let r_∞ denote the line in the intersection of Π_∞ with the vertical plane that projects over the segment $\mathcal{S}_0$, since $h \leq P_\infty$ along $\mathcal{S}_0$, we obtain the existence of a plane Π_1 that contains r_∞, has slope smaller than $\mu_{\max}$ in the x-direction, and does not touch $\varphi(\gamma)$; see Figure 2.4.

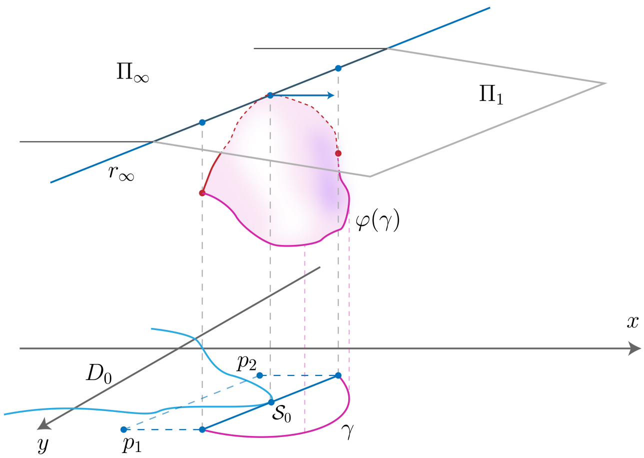

Fig. 2.4. The argument in the proof of Assertion 2.8.

Consider next the graph $\mathcal{G}$ in $\mathbb{R}^3$ given by the restriction of $z = h(x, y)$ to the compact domain of $\mathbb{R}^2$ bounded by the segment $\mathcal{S}_0$ and the curve γ. Since $\mathcal{G}$ is saddle and its boundary does not touch the half-space of $\mathbb{R}^3$ above Π_1, $\mathcal{G}$ also has this property. But now observe that at the points of the non-empty set $\mathcal{S}_0 \cap D_0$ we have $h_x = \mu_{\max}$. Since the slope of Π_1 in the x-direction is smaller than $\mu_{\max}$, this implies that there should exist points of $\mathcal{G}$ above Π_1, a contradiction. This proves Assertion 2.8. ∎

Since D^-, D^+ are disjoint, open convex sets of $\mathbb{R}^2$, there exists a line $\mathcal{L} \subset \mathbb{R}^2$ that separates them strictly, i.e., D^- and D^+ lie in different connected components of $\mathbb{R}^2 \setminus \mathcal{L}$. In particular, any point of the straight half-line $\mathcal{L}^* := \mathcal{L} \cap \{y \geq \delta_0\}$ lies in the set

$$\mathcal{D} = \{(x, y) : y \geq \delta_0,\ x \in [x^-(y), x^+(y)]\}. \tag{2.27}$$

Observe that, by (iii) of (2.26), we have $h_x = \mu_{\max}$ and $h = P_\infty$ on $\mathcal{D}$, i.e., $\varphi(\mathcal{D}) \subset \Pi_\infty \cap \Sigma$. Since the intersection of $\nabla u(\mathbb{S}^2)$ with the support plane $x = \mu_{\max}$ of $\mathbb{R}^3$ is just the point p_0, we deduce that $\psi(\mathcal{D}) = \{p_0\}$, where ψ is given by (2.14). Thus, h_y is constant on $\mathcal{D}$. In particular, h_x and h_y are constant along $\mathcal{L}^*$, with $h_x = \mu_{\max}$. Then $\varphi(\mathcal{L}^*)$ is a straight half-line that lies in $\Sigma \cap \Pi_\infty$, and we deduce that $h_y = b_\infty$ on $\mathcal{D}$, where b_∞ is defined in (2.25). In particular, the limit plane Π_∞ is tangent to Σ at every point of $\varphi(\mathcal{D})$. Also,

$$p_0 = (\mu_{\max}, b_\infty, *) \in \mathbb{R}^3. \tag{2.28}$$

Note that if $\delta_0 = -\infty$, both $\mathcal{L}^*$ and $\varphi(\mathcal{L}^*)$ are (complete) lines.

Step 8: Existence of a geodesic semicircle in $(\nabla u)^{-1}(p_0)$. In this step we show that, by choosing in a more careful way the initial direction $\nu_0 \in (\nabla u)^{-1}(p_0)$ that we fixed at the beginning of Step 3, we can ensure that $\Omega_\xi := (\nabla u)^{-1}(p_0)$ contains a geodesic semicircle of $\mathbb{S}^2$.

Assume that this last property is not true. Let β be any geodesic arc of $\mathbb{S}^2$ contained in Ω_ξ, and denote its endpoints by $\{\beta_0^1, \beta_0^2\}$. Note that, by our choice of the direction ξ in Step 3, the distance in $\mathbb{S}^2$ between the compact subsets Ω_ξ and $\{\xi, -\xi\}$ is positive (since p_0 is a Pogorelov point). Thus, we can consider the *angle* $\theta(\beta) \in [0, \pi]$ at ξ defined by the two geodesic semicircles γ_1, γ_2 of $\mathbb{S}^2$ with endpoints $\{\xi, -\xi\}$ that satisfy $\beta_0^i \in \gamma_i$ (see Figure 2.5). Since β has length $< \pi$ by hypothesis, this angle is $< \pi$.

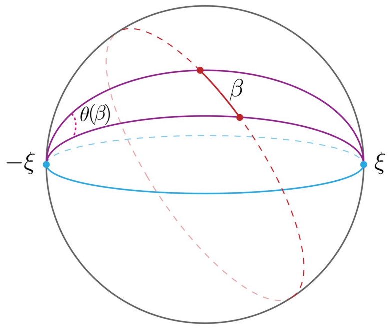

Fig. 2.5. The definition of the angle $\theta(\beta)$.

Observe first of all that there exists at least one geodesic arc (of positive length) β^* contained in Ω_ξ. To see this, let $\mathcal{L}^*$ denote the straight half-line of the (x, y)-plane whose existence was shown in Step 7. Let β^* be the geodesic arc in $\mathbb{S}^2$ that corresponds to $\mathcal{L}^*$ via the totally geodesic bijection $\mathbb{R}^2 \to \mathbb{S}^2_+$ given by (2.15). Since $h_x = \mu_{\max}$ along $\mathcal{L}^*$, we find from (2.14) and (2.28) that

$$\beta^* \subset (\nabla u)^{-1}(p_0) = \Omega_\xi. \tag{2.29}$$

Since $\mathcal{L}^*$ is not parallel to the y-axis, clearly $\theta(\beta^*) > 0$.

We next prove that there exists a geodesic arc β_∞ of maximum angle in Ω_ξ. Let $\theta_0 \in (0,\pi]$ denote the supremum of the angles $\theta(\beta)$ among all possible choices of geodesic arcs β contained in Ω_ξ. Take any sequence $\{\beta_n\}_n$ of geodesic arcs in Ω_ξ with $\theta(\beta_n) \to \theta_0$. Then, up to a subsequence, the endpoints a_n, b_n and the midpoint c_n of the β_n converge to three geodesically aligned points $\{a_1, a_2, a_3\}$ in Ω_ξ. Since any point of β_n is a convex combination of its endpoints, we deduce that $\{\beta_n\}_n$ converges to the geodesic arc β_∞ contained in Ω_ξ with endpoints $\{a_1, a_2\}$ and midpoint a_3. In particular, β_∞ has positive length $< \pi$, and $\theta(\beta_\infty) = \theta_0$. We then conclude that $\theta_0 < \pi$.

Once we know this, it is clear that we can choose the original $\nu_0 \in (\nabla u)^{-1}(p_0)$, which was initially chosen in Step 3 without any a priori limitation, as follows: ν_0 *is the unique point of the geodesic arc* $\beta_\infty \subset \Omega_\xi$ *with the property that the angles* θ_1, θ_2 *of the two geodesic arcs of* β_∞ *joining* ν_0 *to each of the endpoints* $\{a_1, a_2\}$ *of* β_∞ *satisfy* $\theta_i = \theta_0/2 < \pi/2$ *for* $i = 1, 2$ (see Figure 2.6). This choice for ν_0 lets us choose in a more specific way the coordinates (x, y, z) at the beginning of Step 3. Recall that, in these (x, y, z) coordinates, we had $\xi = (1, 0, 0)$, $\nu_0 = (\nu_0^1, 0, \nu_0^3)$ with $\nu_0^3 > 0$. By our new specific choice of ν_0, after a suitable rotation of the (x, y, z) coordinates around the x-axis, we can additionally suppose that the arc β_∞ lies in the hemisphere $\mathbb{S}^2 \cap \{z > 0\}$. Note that $\nu_0 \in \beta_\infty$, and that every point of β_∞ lies in $(\nabla u)^{-1}(p_0)$.

Consider the totally geodesic bijection $\mathbb{R}^2 \to \mathbb{S}^2_+$ given by (2.15). This bijection takes ν_0 to $(x_0, 0)$ for some $x_0 \in \mathbb{R}$, and β_∞ to a compact line segment L_∞ passing through $(x_0, 0)$ (see Figure 2.6). In the same way, the geodesic semicircles γ_1, γ_2 in $\mathbb{S}^2 \cap \{z \geq 0\}$ that pass through the points $\{\xi, -\xi, a_i\}$ are projected into two parallel lines in $\mathbb{R}^2$ of the form $y = r_i$ for some $r_1 < 0 < r_2$. Obviously, each of the endpoints of L_∞ lies in one of these lines.

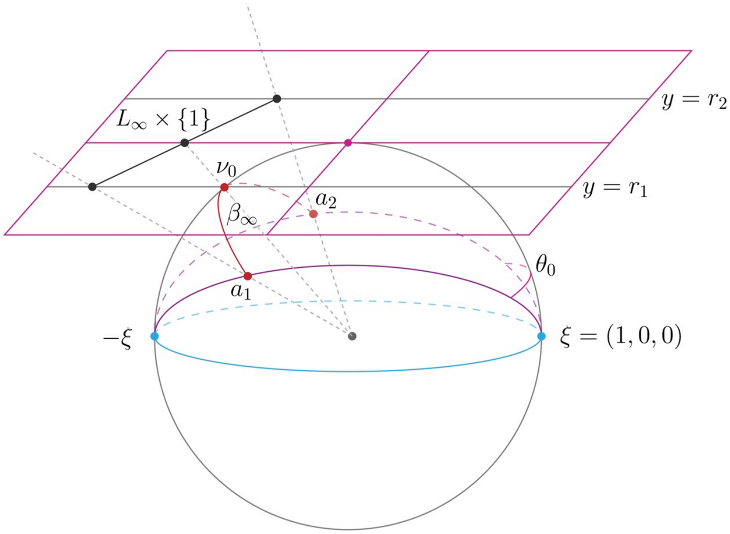

Fig. 2.6. Choice of $\nu_0 \in (\nabla u)^{-1}(p_0)$.

We can now carry out the argument in Steps 3 through 7 for this new choice of ν_0. Let $\mathcal{D} \subset \mathbb{R}^2 \cap \{y \geq \delta_0\}$ denote the subset given by (2.27) in Step 7 of the proof. Since $\psi(L_\infty) = \{p_0\}$, where ψ is given by (2.14), we deduce from (2.28) that $(h_x, h_y) = (\mu_{\max}, b_\infty)$, constant along L_∞. Also, observe that $(x_0, 0) \in \mathcal{D} \cap L_\infty$ and recall that $\varphi(\mathcal{D}) \subset \Pi_\infty \cap \Sigma$. In this way, $\varphi(L_\infty) \subset \Pi_\infty \cap \Sigma$. Since $h_x = \mu_{\max}$ along L_∞, we conclude by the definition of δ_0 that $\delta_0 \leq r_1$.

Consider next the geodesic arc β^* in (2.29). It corresponds via (2.15) to the half-line $\mathcal{L}^* = \mathcal{L} \cap \{y \geq \delta_0\}$. Since we have proved that $[r_1, r_2] \subset [\delta_0, \infty)$, this geodesic arc β^* has angle $\theta(\beta^*)$ greater than $\theta(\beta_\infty) = \theta_0$. This contradicts the definition of θ_0. Therefore, $(\nabla u)^{-1}(p_0)$ contains a geodesic semicircle of $\mathbb{S}^2$.

Step 9: Existence of a geodesic semicircle in $(\nabla u)^{-1}(p)$ *for at least four different points.* We have seen that, for any Pogorelov point $p_0 \in \nabla u(\mathbb{S}^2)$ of the hedgehog $\nabla u(\mathbb{S}^2)$, the set $(\nabla u)^{-1}(p_0)$ contains a geodesic semicircle. We will next show that there exist at least four different Pogorelov points for $\nabla u(\mathbb{S}^2)$, which proves the statement above.

Let p be a contact point of $\nabla u(\mathbb{S}^2)$ with one of its support planes, and consider the set $\mathcal{N}_p := \{\xi \in \mathbb{S}^2 : p \in \Pi_\xi\}$. Note that the convex hull $\mathcal{C}$ of $\nabla u(\mathbb{S}^2)$ is not contained in a plane, since ∇u has some regular point of negative curvature (see the proof of Assertion 2.3). In these conditions, it is well known that $\mathcal{N}_p$ is a compact, convex subset of an open hemisphere of $\mathbb{S}^2$.

Arguing by contradiction, assume that $\nabla u(\mathbb{S}^2)$ has at most three (distinct) Pogorelov points p_1, p_2, p_3. Then $\mathcal{V} := \mathbb{S}^2 \setminus \bigcup_{i=1}^3 \mathcal{N}_{p_i}$ is a non-empty open set, since each $\mathcal{N}_{p_i}$ lies in an open hemisphere. For almost every $\xi \in \mathcal{V}$, the intersection $\Pi_\xi \cap \nabla u(\mathbb{S}^2)$ is a unique point q_ξ, which is not a Pogorelov point. Thus, from the definition of Pogorelov point, either $\nabla u(\xi) = q_\xi$, or $\nabla u(-\xi) = q_\xi$, for almost all $\xi \in \mathcal{V}$. If $\nabla u(-\xi_0) \neq q_{\xi_0}$ for any such ξ_0, then, by definition of support plane,

$$\langle \nabla u(-\xi_0) - q_{\xi_0}, \xi_0 \rangle < \langle \nabla u(\xi_0) - q_{\xi_0}, \xi_0 \rangle = 0,$$

and so

$$\langle \nabla u(-\xi_0), \xi_0 \rangle < \langle \nabla u(\xi_0), \xi_0 \rangle.$$

Hence, this property holds in a neighborhood $\mathcal{W} \subset \mathcal{V}$ of ξ_0, and it implies that for almost every $\xi \in \mathcal{W}$, we have $\nabla u(\xi) = q_\xi$. In particular, ∇u is singular in a neighborhood of ξ_0, since regular points of $\nabla u(\mathbb{S}^2)$ never touch support planes. If $\nabla u(\xi_0) \neq q_{\xi_0}$, the same argument shows that ∇u is singular in a neighborhood $\mathcal{W}$ of $-\xi_0$, and $\nabla u(\xi) = q_{-\xi}$ for almost every $\xi \in \mathcal{W}$.

Finally, if $\nabla u(\xi) = \nabla u(-\xi) = q_\xi$ for almost all $\xi \in \mathcal{V}$, then ∇u is singular in $\mathcal{V}$.

In other words, we have shown that there exists an open set $\mathcal{W} \subset \mathbb{S}^2$ such that ∇u is singular everywhere on $\mathcal{W}$, and for almost every $\xi \in \mathcal{W}$, $\nabla u(\xi)$ is the unique contact point of $\nabla u(\mathbb{S}^2)$ with one of the support planes Π_ξ or $\Pi_{-\xi}$.

Recall that, by homogeneity, $D^2 u$ always has a zero eigenvalue at every point, corresponding to the radial direction, and that the regular points of the hegdehog $\nabla u(\mathbb{S}^2)$ are those where the rank of $D^2 u$ is 2; see the paragraph before Definition 2.2. Since ∇u

is singular on $\mathcal{W}$, by reducing $\mathcal{W}$ if necessary, we can assume additionally that the rank of D^2u is constantly equal to 0 or 1 in $\mathcal{W}$. We rule out these two cases separately.

Assertion 2.9. *The rank of D^2u cannot be zero in $\mathcal{W}$.*

Proof. Assume that $D^2u = 0$ in $\mathcal{W}$, and choose $\xi \in \mathcal{W}$. Suppose, for definiteness, that $\nabla u(\xi) = q_\xi \in \Pi_\xi$; the discussion is similar if $\nabla u(\xi) \in \Pi_{-\xi}$.

We will start by arguing as in Step 3. Consider Euclidean coordinates (x, y, z) in $\mathbb{R}^3$ such that $\xi = (1, 0, 0)$, and let Σ be the entire saddle graph in $\mathbb{R}^3$ given by $z = h(x, y)$, where h is defined by (2.1). Then equations (2.17) and (2.18) at the beginning of Step 3 hold, but (2.19) does not. Since u is linear in a neighborhood of ξ, with $u_x = \mu_{\max}$, we deduce that instead of (2.19), in our context we have

$$h_x(x, y) = \mu_{\max} \quad \forall (x, y) \in (0, \infty) \times \mathbb{R} \text{ with } x^2 \geq \delta(y^2 + 1), \tag{2.30}$$

for some $\delta > 0$. In this way, if we choose $(x_0, 0)$ with $x_0 > \delta$ and define the linear function

$$P(x, y) := \mu_{\max}(x - x_0) + h_y(x_0, 0)y + h(x_0, 0),$$

we have $h(x, y) = P(x, y)$ in a connected planar subset $\Omega \subset \mathbb{R}^2$ that contains the set defined in (2.30), and $h(x, y) > P(x, y)$ in $\mathbb{R}^2 \setminus \Omega$.

By the argument in Assertion 2.8, we deduce that $\mathbb{R}^2 \setminus \Omega$ is an open convex set. Consider the set $\Theta_0 \subset \mathbb{S}^2$ given by the points ν of the form (2.15) with $(x, y) \in \Omega$. Since (2.15) is a totally geodesic mapping, this means that if $\mathbb{S}^2_+ := \mathbb{S}^2 \cap \{z > 0\}$, then $\mathbb{S}^2_+ \setminus \Theta_0$ is a convex subset of $\mathbb{S}^2_+$. But note that the Euclidean coordinates (x, y, z) were chosen arbitrarily except for the condition $\xi = (1, 0, 0)$. Thus, if we define $\Theta \subset \mathbb{S}^2$ as the set of points $\nu \in \mathbb{S}^2$ that are given by (2.15) for some $(x, y) \in \Omega$ with respect to *some* Euclidean coordinates (x, y, z) with $\xi = (1, 0, 0)$, we deduce that $\mathbb{S}^2 \setminus \Theta$ is a convex subset of $\mathbb{S}^2$, and u is linear on Θ. Then $\mathbb{S}^2 \setminus \Theta$ lies in an open hemisphere. Consequently, u is linear on a closed hemisphere H of $\mathbb{S}^2$, with $\nabla u = q_\xi$. Consider next the homogeneous function $v(p) := u(p) - \langle p, q_\xi \rangle$, defined for all $p \in \mathbb{R}^3$. Note that $D^2v = D^2u$ everywhere, and that v vanishes along the geodesic ∂H of $\mathbb{S}^2$. By [13, Theorem 1.6.4] or [8, Theorem 2], v must be linear. Hence, u is linear, a contradiction. ■

Assertion 2.10. *The rank of D^2u cannot be* 1 *in $\mathcal{W}$.*

Proof. In order to prove the assertion, we use some results of hegdehog theory developed by Martinez-Maure [10], which we explain next. Given $h \in C^2(\mathbb{S}^2)$, let $\mathcal{H}$ be the *hedgehog* in $\mathbb{R}^3$ with support function h, i.e., $\mathcal{H}$ is given by

$$\chi(\nu) := \nabla_{\mathbb{S}} h(\nu) + h(\nu)\nu : \mathbb{S}^2 \to \mathcal{H} \subset \mathbb{R}^3,$$

where $\nabla_{\mathbb{S}}$ denote the gradient in $\mathbb{S}^2$. We assume that the curvature of χ is negative at its regular points, and that χ is not constant. Note that the hedgehog $\mathcal{H} := \nabla u(\mathbb{S}^2)$ of our problem satisfies these conditions.

For any $\omega \in \mathbb{S}^2$, consider the plane $P = \{\omega\}^\perp$, and let $\pi : \mathbb{R}^3 \to P$ denote the orthogonal projection. Define $\chi_\omega : \mathbb{S}^1 \equiv \mathbb{S}^2 \cap P \to P$ by

$$\chi_\omega(\theta) := \pi(\chi(\theta)). \tag{2.31}$$

Then χ_ω defines a *planar hedgehog* in P, which we denote by $\mathcal{H}_\omega = \chi_\omega(\mathbb{S}^1)$. Since $\mathcal{H}$ has negative curvature at its regular points, this projected hedgehog $\mathcal{H}_\omega$ has empty *convex interior*; see [10, Theorem 2 and Corollary 1], where the definition of convex interior of a planar hedgehog (which we will not use explicitly) is also presented; see also [11, Corollary 1].

We now prove Assertion 2.10 using this information. Since D^2u has rank one in the open set $\mathcal{W} \subset \mathbb{S}^2$, then $\nabla u(\mathcal{W})$ is a regular curve γ. Also, note that for almost every $q \in \gamma$ we have either $\{q\} = \Pi_\xi \cap \nabla u(\mathbb{S}^2)$ or $\{q\} = \Pi_{-\xi} \cap \nabla u(\mathbb{S}^2)$.

Let T be the unit tangent vector to γ at q, and define $\omega := T \times \xi$. Let $\pi : \mathbb{R}^3 \to \{\omega\}^\perp$ denote the orthogonal projection onto $P = \{\omega\}^\perp$. Then $\beta := \pi(\gamma)$ is a regular curve in $P = \{\omega\}^\perp$ around $\pi(q)$, and $\pi(q) \in \beta \cap \mathcal{H}_\omega$ (since $\langle T, \omega\rangle = 0$), where $\mathcal{H}_\omega$ is the planar hedgehog given by (2.31). Note that $\pi(q)$ is a regular point of $\mathcal{H}_\omega$, since $\chi_\omega(T) = \pi(q)$ and $\langle \nabla u(q), T\rangle \neq 0$, by regularity of γ. Also, either $\mathcal{H}_\omega$ lies on one side of the line $L_\xi = \Pi_\xi \cap P$, and in that case $\pi(q) \in L_\xi \cap \mathcal{H}_\omega$, or else $\mathcal{H}_\omega$ lies on one side of $L_{-\xi} = \Pi_{-\xi} \cap P$, and $\pi(q) \in L_{-\xi} \cap \mathcal{H}_\omega$. In this way, in any of these two cases, the planar hedgehog $\mathcal{H}_\omega \subset P$ touches one of its support lines at the regular point $\pi(q)$. Since $\mathcal{H}_\omega$ has empty convex interior, we obtain a contradiction with [10, Proposition 1]. ∎

Thus, we have proved that $\nabla u(\mathbb{S}^2)$ has at least four Pogorelov points, as claimed.

Step 10: The final contradiction. We now conclude the proof of Theorem 1.3. Recall that we had initially assumed that u is not a linear function, and we were arguing by contradiction.

We have shown in Step 9 that there exist at least four different points $p_1, \dots, p_4 \in \nabla u(\mathbb{S}^2)$ for which $(\nabla u)^{-1}(p_j)$ contains a geodesic semicircle Γ_j of $\mathbb{S}^2$. The geodesic semicircles $\Gamma_1, \dots, \Gamma_4$ are disjoint, since the p_j are different.

Consider the region $\mathcal{O} \subset \mathbb{S}^2$ defined below (1.4). By hypothesis on $\mathcal{O}$, we have $\mathcal{O} \cap \Gamma_j \neq \emptyset$ for some $j \in \{1, \dots, 4\}$. Let Ω_j denote the compact set $(\nabla u)^{-1}(p_j)$. Thus, $\Omega_j \cap \mathcal{O} \neq \emptyset$, and since $\mathcal{O}$ is connected, either $\partial\Omega_j \cap \mathcal{O} \neq \emptyset$ or $\mathcal{O} \subset \Omega_j$.

Suppose first that $\mathcal{O} \subset \Omega_j$. Then it is clear that the distance from $\mathcal{O}$ to any of the semicircles Γ_k, $k \neq j$, is positive. In particular, there exists $\varepsilon > 0$ such that $\mathcal{O}$ does not intersect the open set $\mathcal{U}_\varepsilon := \{\nu \in \mathbb{S}^2 : \operatorname{dist}(\nu, \Gamma_k) < \varepsilon\}$. But on the other hand, it is clear that there exist infinitely many closed disjoint geodesic semicircles contained in $\mathcal{U}_\varepsilon$. This contradicts the hypothesis that $\mathcal{O}$ intersects any configuration of four disjoint geodesic semicircles. Thus, $\mathcal{O}$ is not contained in Ω_j.

Hence, there must exist some $w_j \in \partial\Omega_j \cap \mathcal{O}$. Since $w_j \in (\nabla u)^{-1}(p_j)$, we can choose w_j as the vector $\nu_0 \in \mathbb{S}^2$ in the argument that we carried out in Steps 3 through 7. Specifically, choose Euclidean coordinates (x, y, z) so that $\xi_j = (1, 0, 0)$ and $w_j =: \nu_0 = (\nu_0^1, 0, \nu_0^3)$, with $\nu_0^3 > 0$. Denote $\mathbb{S}^2_+ = \mathbb{S}^2 \cap \{z > 0\}$. Then, by the argument in Steps 3 through 7, the connected component of the set $(\nabla u)^{-1}(p_j) \cap \mathbb{S}^2_+$ that contains ν_0 is made up of the points $\nu \in \mathbb{S}^2$ given by (2.15) with (x, y) a point of the planar set $\mathcal{D}$ defined in (2.27). Also, (2.28) holds for $p_0 := p_j$.

Take $x_0 \in \mathbb{R}$ given by $\nu_0 = \frac{1}{\sqrt{1+x_0^2}}(x_0, 0, 1)$. Since $\nu_0 \in \partial\Omega_j$, obviously $(x_0, 0) \in \partial\mathcal{D}$, and $h_x(x_0, 0) = \mu_{\max}$ by (2.28) and (2.14). Thus, h_x has an absolute maximum at $(x_0, 0)$. Hence, as $\nu_0 := w_j$ lies in $\mathcal{O}$, it follows by Assertion 2.1 that h_x is constant around $(x_0, 0)$, since ∇h cannot be an open mapping. Then, by (2.14), ν_0 lies in the interior of Ω_j, in contradiction with $\nu_0 \in \partial\Omega_j$.

By this final contradiction, the function u must be linear, and this proves Theorem 1.3.

3. Proof of Theorem 1.4

In Steps 2 through 9 of our proof of Theorem 1.3 we actually showed the following result. Let $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ be a 1-homogeneous solution to a linear equation (1.1). Assume that the coefficients a_{ij} of (1.1) satisfy the degenerate ellipticity conditions (i), (ii) of (1.5). Let $\nabla u : \mathbb{S}^2 \to \mathbb{R}^3$ be the restriction of the gradient of u to $\mathbb{S}^2$. Then there exist at least four different points $p_1, \ldots, p_4$ in $\mathbb{R}^3$ such that each $(\nabla u)^{-1}(p_j)$ contains a geodesic semicircle Γ_j for $j = 1, \ldots, 4$. These semicircles are disjoint, and $D^2 u$ vanishes along the configuration $\Gamma = \bigcup_{i=1}^4 \Gamma_i$.

As explained at the beginning of Step 2, there is an equivalence between 1-homogeneous solutions $u \in C^2(\mathbb{R}^3 \setminus \{0\})$ of (1.1) whose coefficients satisfy conditions (i), (ii) of (1.5) and C^2 saddle functions $v(x) = u(x/|x|)$ on $\mathbb{S}^2$. Taking into account this equivalence, it is then clear that the result obtained in Steps 2 through 9 that we just recalled directly proves Theorem 1.4.

Theorem 1.4 is equivalent to the geometric statement below. Indeed, if $\rho \in C^2(\mathbb{S}^2)$ denotes the support function of an ovaloid satisfying (3.1), then $v := \rho - c$ is a saddle function in $\mathbb{S}^2$, thus satisfying the conditions of Theorem 1.4 (and conversely).

Theorem 3.1. *Let $S \subset \mathbb{R}^3$ be a C^2 ovaloid in $\mathbb{R}^3$ whose principal curvatures κ_1, κ_2 satisfy*

$$(\kappa_1 - c)(\kappa_2 - c) \leq 0 \tag{3.1}$$

for some $c > 0$. Then S is round along four geodesic semicircles. Specifically, S is tangent up to the second order to four spheres $\Sigma_1^c, \ldots, \Sigma_4^c$ of radius $1/c$ along four disjoint geodesic semicircles $\alpha_j \subset \Sigma_j^c \cap S$ for $j = 1, \ldots, 4$.

In other words, there exist four disjoint geodesic semicircles $\Gamma_1, \ldots, \Gamma_4$ in $\mathbb{S}^2$ such that if $\eta : S \to \mathbb{S}^2$ is the Gauss map of S, then each $\eta^{-1}(\Gamma_j) = \alpha_j$ is made up of umbilic points of S, and coincides with a geodesic semicircle of a sphere of radius $1/c$ in $\mathbb{R}^3$.

Acknowledgments. The authors are grateful to Yves Martinez-Maure for enlightening comments and discussions.

Funding. Research financially supported by: Projects PID2020-118137GB-I00 and CEX2020-001105-M, funded by MCIN/AEI /10.13039/501100011033; CARM, Programa Regional de Fomento de la Investigación, Fundación Seneca - Agencia de Ciencia y Tecnología de la Región de Murcia, reference 21937/PI/22.

References

[1] Alexandrov, A. D.: Sur les théorèmes d'unicité pour les surfaces fermées. C. R. Dokl. Acad. Sci. URSS **22**, 99–102 (1939) JFM 65.0828.03

[2] Aleksandrov, A. D.: Uniqueness theorems for surfaces in the large, I. Vestnik Leningrad. Univ. **11**, 5–17 (1956) (in Russian); English transl.: Amer. Math. Soc. Transl. **21**, 341–354 (1962) Zbl 0101.13801 MR 0086338

[3] Astala, K., Iwaniec, T., Martin, G.: Elliptic partial differential equations and quasiconformal mappings in the plane. Princeton Math. Ser. 48, Princeton University Press, Princeton, NJ (2009) Zbl 1182.30001 MR 2472875

[4] Gálvez, J. A., Mira, P., Tassi, M. P.: A quasiconformal Hopf soap bubble theorem. Calc. Var. Partial Differential Equations **61**, article no. 129, 20 pp. (2022) Zbl 1502.53013 MR 4417394

[5] Guan, P., Wang, Z., Zhang, X.: A proof of Alexandrov's uniqueness theorem for convex surfaces in $\mathbb{R}^3$. Ann. Inst. H. Poincaré C Anal. Non Linéaire **33**, 329–336 (2016) Zbl 1335.53092 MR 3465378

[6] Han, Q., Nadirashvili, N., Yuan, Y.: Linearity of homogeneous order-one solutions to elliptic equations in dimension three. Comm. Pure Appl. Math. **56**, 425–432 (2003) Zbl 1236.35025 MR 1949137

[7] Iwaniec, T., Šverák, V.: On mappings with integrable dilatation. Proc. Amer. Math. Soc. **118**, 181–188 (1993) Zbl 0784.30015 MR 1160301

[8] Koutroufiotis, D.: On a conjectured characterization of the sphere. Math. Ann. **205**, 211–217 (1973) Zbl 0257.53056 MR 0348684

[9] Langevin, R., Levitt, G., Rosenberg, H.: Hérissons et multihérissons (enveloppes parametrées par leur application de Gauss). In: Singularities (Warszawa, 1985), Banach Center Publ. 20, PWN, Warszawa, 245–253 (1988) Zbl 0658.53004 MR 1101843

[10] Martinez-Maure, Y.: Indice d'un hérisson: étude et applications. Publ. Mat. **44**, 237–255 (2000) Zbl 0974.53003 MR 1775763

[11] Martinez-Maure, Y.: Contre-exemple à une caractérisation conjecturée de la sphère. C. R. Acad. Sci. Paris Sér. I Math. **332**, 41–44 (2001) Zbl 1008.53002 MR 1805625

[12] Mooney, C.: Minimizers of convex functionals with small degeneracy set. Calc. Var. Partial Differential Equations **59**, article no. 74, 19 pp. (2020) Zbl 1456.49029 MR 4079759

[13] Nadirashvili, N., Tkachev, V., Vlăduţ, S.: Nonlinear elliptic equations and nonassociative algebras. Math. Surveys Monogr. 200, American Mathematical Society, Providence, RI (2014) Zbl 1357.35002 MR 3243534

[14] Nadirashvili, N., Vlăduţ, S.: Homogeneous solutions of fully nonlinear elliptic equations in four dimensions. Comm. Pure Appl. Math. **66**, 1653–1662 (2013) Zbl 1284.35178 MR 3084701

[15] Panina, G.: New counterexamples to A. D. Alexandrov's hypothesis. Adv. Geom. **5**, 301–317 (2005) Zbl 1077.52003 MR 2131822

[16] Panina, G.: Isotopy problems for saddle surfaces. Eur. J. Combin. **31**, 1160–1170 (2010) Zbl 1213.52002 MR 2607258

[17] Pogorelov, A. V.: Solution of a problem of A. D. Aleksandrov. Dokl. Akad. Nauk **360**, 317–319 (1998) (in Russian); English transl.: Doklady Math. **57**, 398–399 (1998) Zbl 0980.53088 MR 1673360

[18] Pucci, C.: Un problema variazionale per i coefficienti di equazioni differenziali di tipo ellittico. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) **16**, 159–172 (1962) Zbl 0197.08802 MR 0141871

[19] Safonov, M. V.: Unimprovability of estimates of Hölder constants for solutions of linear elliptic equations with measurable coefficients. Mat. Sb. (N.S.) **132**, 275–288 (1987) (in Russian); English transl.: Math. USSR-Sb. **60**, 269–281 (1988) Zbl 0656.35027 MR 0882838

[20] Safonov, M. V.: Nonlinear elliptic equations of second order. Lecture Notes, University of Florence (1991)