MATHEMATICS OF COMPUTATION
Volume 85, Number 302, November 2016, Pages 2743–2775
http://dx.doi.org/10.1090/mcom/3080
Article electronically published on March 22, 2016

MONOTONE AND CONSISTENT DISCRETIZATION OF THE MONGE-AMPÈRE OPERATOR

JEAN-DAVID BENAMOU, FRANCIS COLLINO, AND JEAN-MARIE MIREBEAU

Abstract. We introduce a novel discretization of the Monge-Ampère operator, simultaneously consistent and degenerate elliptic, hence accurate and robust in applications. These properties are achieved by exploiting the arithmetic structure of the discrete domain, assumed to be a two dimensional cartesian grid. The construction of our scheme is simple, but its analysis relies on original tools seldom encountered in numerical analysis, such as the geometry of two dimensional lattices and an arithmetic structure called the Stern-Brocot tree. Numerical experiments illustrate the method's efficiency.

1. Introduction

We introduce a new discretization of the Monge-Ampère operator, on two dimensional cartesian grids, which is consistent and preserves at the discrete level a fundamental property of the continuous operator: degenerate ellipticity. Discrete degenerate ellipticity [20] implies strong guarantees for the numerical scheme: a comparison principle, convergence of discrete solutions towards the continuous one in the setting of viscosity solutions, and convergence of Euler iterative solvers for the discrete system [20]. Some Degenerate Elliptic (DE) schemes for the Monge-Ampère (MA) Partial Differential Equation (PDE) already exist [9, 20], but they suffer from several flaws: they are strongly non-local and only approximately consistent. Consistent *non-DE* schemes such as [3, 15] offer better accuracy but require the PDE solution to be sufficiently smooth and the discrete numerical solver to be well initialized. Filtered schemes [10] non-linearly combine several existing schemes in order to cumulate their advantages (here degenerate ellipticity and consistency) or mitigate their defects. Their definition and their analysis are however complex, and their application requires adjusting several parameters. For a recent overview of the numerical approaches to solving the Monge-Ampère equation, see Glowinski, Feng and Neilan [8].

We introduce a new numerical scheme, Monge-Ampère using Lattice Basis Reduction (MA-LBR), which is both consistent[1] *and* degenerate elliptic. Lattice Basis reduction is a tool from discrete geometry, which arises here due to the interaction of the cartesian discretization grid, with the anisotropic nature of the Monge-Ampère operator. This operator is indeed invariant under all linear changes of variables with unit determinant, unlike e.g. the Laplacian, which is merely invariant under

Received by the editor September 23, 2014 and, in revised form, May 11, 2015.
2010 *Mathematics Subject Classification.* Primary 35J96, 65N06.
Key words and phrases. Monge-Ampère PDE, monotone finite differences scheme, lattice basis reduction, Stern-Brocot tree.
This work was partially supported by the ANR grant NS-LBR ANR-13-JS01-0003-01.

[1] Assuming the solution hessian condition number is uniformly bounded.

orthogonal transformations. The MA-LBR is inspired by the Wide Stencil [20] family of schemes. Using another arithmetic tool, the Stern-Brocot tree, we solve a second issue plaguing these methods (in addition to consistency errors): our discretization stencil need not be chosen a priori (which usually involves a difficult arbitrage between scheme locality, consistency error and available CPU time), but can be generated automatically in a guaranteed, parameter free and solution adapted manner. Numerical experiments in §4 illustrate the MA-LBR accuracy and robustness.

We fix throughout this paper a convex open bounded domain $\Omega \subseteq \mathbb{R}^2$. Given a positive density $\rho \in C^0(\overline{\Omega}, \mathbb{R}_+^*)$ and some Dirichlet data $\sigma \in C^0(\partial\Omega, \mathbb{R})$, we set the goal of approximating numerically the unique viscosity solution [6, 12] of

$$\begin{cases} \det(\nabla^2 u) = \rho & \text{on } \Omega, \\ u = \sigma & \text{on } \partial\Omega, \\ u \text{ convex}. \end{cases} \tag{1.1}$$

Our framework admittedly does not encompass solutions of the weaker Alexandrov type, where ρ is merely a non-negative measure. If Ω is convex but not strictly convex, then the Dirichlet data σ is assumed to be convex on any segment of $\partial\Omega$. Let us point out that optimal transport, from Ω to another domain Ω', equipped with densities ρ, ρ', admits a PDE formulation similar in spirit to (1.1): $\rho'(\nabla u)\det(\nabla^2 u) = \rho$, $\nabla u(\Omega) \subseteq \Omega'$, u convex. The gradient non-linearity and the second boundary condition raise difficulties [1, 24] that we choose not to address in the present paper, focusing instead on the Monge-Ampère operator $\det(\nabla^2 u)$.

We assume that the PDE domain Ω is discretized on a cartesian grid: $\Omega \cap hR(\xi + \mathbb{Z}^2)$, where $h > 0$ is the grid scale, R is an arbitrary rotation, and ξ is an offset. For notational simplicity, and up to a linear change of coordinates, we limit our attention to the canonical values of these parameters, so that the discrete domain is

$$X := \Omega \cap \mathbb{Z}^2.$$

Definition 1.1. We denote by $\mathbb{U}$ the collection of discrete maps $u : X \cup \partial\Omega \to \mathbb{R}$. A (discrete) operator is a map $\mathcal{D} : \mathbb{U} \to \mathbb{R}^X$. It associates to each $u \in \mathbb{U}$ a collection of values $\mathcal{D}u(x)$, $x \in X$.

The notation $\mathcal{D}u$ and $\mathcal{D}(u)$ refers to the same object, which is a map $X \to \mathbb{R}$, and is used interchangeably with the aim of improving readability. In numerical experiments, the values of $u \in \mathbb{U}$ on X are the unknowns, while the values on $\partial\Omega$ are the supplied boundary data: $u_{|\partial\Omega} = \sigma$. For each $e \in \mathbb{Z}^2$ we introduce a second order differences operator Δ_e, built so that $\Delta_e u(x) \approx \langle e, (\nabla^2 u(x))e\rangle$, where $u \in \mathbb{U}$ and $x \in X$, and where $\langle \cdot, \cdot \rangle$ denotes the euclidean scalar product on $\mathbb{R}^2$. In the simplest case where $x \pm e \in \Omega$, we set

$$\Delta_e u(x) := u(x+e) - 2u(x) + u(x-e). \tag{1.2}$$

When $x \in X$ is close to $\partial\Omega$, the points $x+e$ or $x-e$ may not belong to Ω. Denoting by $h^\pm$ the only element of $]0,1]$ such that $x \pm h^\pm e \in X \cup \partial\Omega$, we define

$$\Delta_e u(x) := \frac{2}{h^+ + h^-}\left(\frac{u(x+h^+e) - u(x)}{h^+} + \frac{u(x-h^-e) - u(x)}{h^-}\right). \tag{1.3}$$

Let us again point out that if $h^+ < 1$, then the value $u(x + h^+e)$ is the supplied boundary data $\sigma(x+h^+e)$. On the other hand if $h^+ = h^- = 1$, then (1.2) and (1.3)

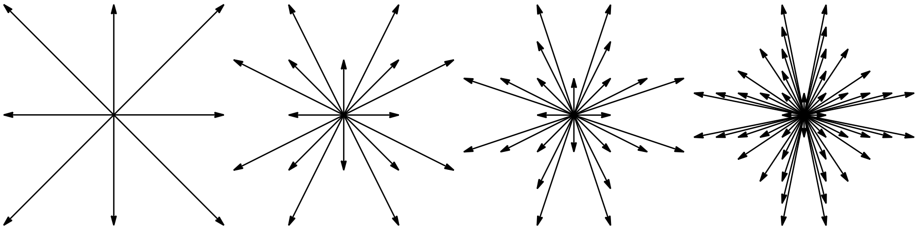

FIGURE 1. Examples of stencils $V \subseteq \mathbb{Z}^2$, containing 8, 16, 24 and 48 elements respectively. Wider stencils yield smaller consistency errors; see Figure 3.

coincide. No other consistent approximation of $\langle e, (\nabla^2 u(x))e\rangle$ can be built using the values $u(x+h^+e)$, $u(x)$ and $u(x-h^-e)$.

Discretizations of the Monge-Ampère operator $\det(\nabla^2 u)$ are typically built upon the operators Δ_e. Consider for instance the Finite Differences (FD) discretization [15]

$$\mathcal{D}^{\mathrm{FD}} := \Delta_{(1,0)}\Delta_{(0,1)} - (\Delta_{(1,1)} - \Delta_{(1,-1)})^2/16. \tag{1.4}$$

Given such a discrete operator $\mathcal{D}$, the discrete analog of (1.1) takes the form

$$\text{Find } u \in \mathbb{U} \text{ such that } \mathcal{D}u = \rho \text{ on } X \text{ and } u_{|\partial\Omega} = \sigma. \tag{1.5}$$

This discrete system lacks a counterpart of the constraint of convexity in (1.1) because (i) there is no unique notion of discrete convexity but several competing approaches (see for instance [18, 21]), and (ii) some form of discrete convexity constraint can often be embedded in the equation $\mathcal{D}u = \rho$ (see §1.2). From a theoretical and a practical standpoint, choosing $\mathcal{D}^{\mathrm{FD}}$ in (1.5) is a risky bet: second order convergence can often be observed in numerical experiments (see §4) but only on rather easy cases and with a good initialization for the numerical solver. Robustness results (existence, uniqueness, and algorithmic guarantees) are limited to discretizations obeying an additional property: a counterpart of the ellipticity of the (opposite of the) Monge-Ampère operator $-\det(\nabla^2 u)$.

We use the notion of discrete degenerate ellipticity [20], slightly specialized due to our focus on MA. Degenerate Elliptic Monge-Ampère numerical schemes cannot be strictly local, unlike (1.4), but instead need to take into account some long range second order differences, indexed by a possibly wide stencil.

Definition 1.2. A stencil is a finite set $V \subseteq \mathbb{Z}^2 \setminus \{0\}$ which is symmetric with respect to the origin (i.e. $-e \in V$ for each $e \in V$).

Definition 1.3 (DE2 scheme)**.** A numerical scheme $-\mathcal{D}$ is Degenerate Elliptic, with stencil V, iff for each $x \in X$ the quantity $\mathcal{D}u(x)$ is a non-decreasing, locally Lipschitz function of the second order differences $\Delta_e u(x)$, $e \in V$.

Observing that the second order difference $\Delta_e u(x)$ can be expressed as a non-negatively weighted sum of first order differences (1.3), we immediately find that a DE2 scheme is degenerate elliptic in the sense of [20]. DE2 schemes are also positive difference operators in the sense of [14], since they are indeed built using directional second order finite differences. In particular, for any $\varepsilon > 0$, the slightly perturbed operator $-\mathcal{D}_\varepsilon$, defined by $\mathcal{D}_\varepsilon u(x) := \mathcal{D}u(x) - \varepsilon u(x)$, is *proper* elliptic [20]. This in turn implies that the discrete system (1.5) associated with $\mathcal{D}_\varepsilon$ has a

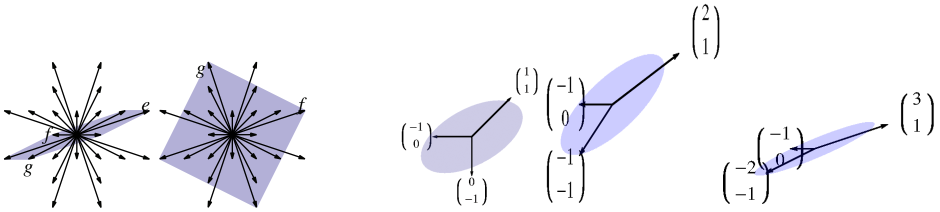

FIGURE 2. Left: A stencil V, a superbase $(e,f,g) \in V^3$, an orthogonal pair $(f,g) \in V^2$. Right: M-obtuse superbase (see Definition 1.8) and ellipse $\{v \in \mathbb{R}^2; \langle v, Mv\rangle \leq 1\}$ for some $M \in S_2^+$.

unique solution which can be computed with a geometric convergence rate using an iterative Euler scheme. The analysis for degenerate (non-proper) elliptic discrete operators is in contrast more complex and requires studying in detail the discrete operator structure, in particular the graph underlying its discretization [16]. We make no claim regarding (1.5) for the operators $\mathcal{D}$ studied below, and admittedly content ourselves with the general results holding for their perturbation $\mathcal{D}_\varepsilon$ as above. Note nevertheless that no issue was encountered numerically in the unperturbed case; see §4. In the continuous setting [6], the study of proper elliptic operators is likewise simpler and more unified than for degenerate elliptic ones. We refer to [6,16,20] for these results and will say no more on this analytic machinery in the rest of the paper, focusing instead on the algebraic structure of discrete Monge-Ampère operators.

Froese and Oberman [9] numerically address the MA PDE using a DE2 operator, referred to as the Wide Stencil (WS) scheme. Given a stencil V, and denoting $\Delta_e^+ := \max\{0, \Delta_e\}$:

$$\mathcal{D}_V^{\mathrm{WS}} u(x) := \min_{\substack{(f,g)\in V^2\\ \text{orthogonal}}} \frac{\Delta_f^+ u(x)}{\|f\|^2} \frac{\Delta_g^+ u(x)}{\|g\|^2}. \tag{1.6}$$

The minimum is taken over all pairs of vectors $(f,g) \in V^2$ which are orthogonal, in the sense that $\langle f, g\rangle = 0$, for instance $(1,0), (0,1)$ or $(2,1), (-1,2)$. We introduce a variant of this operator which does not rely on pairs of orthogonal stencil vectors, but on superbases of the lattice $\mathbb{Z}^2$.

Definition 1.4. A basis of $\mathbb{Z}^2$ is a pair $(f,g) \in (\mathbb{Z}^2)^2$ such that $|\det(f,g)| = 1$. A superbase of $\mathbb{Z}^2$ is a triplet $(e,f,g) \in (\mathbb{Z}^2)^3$ such that $e+f+g=0$, and (f,g) is a basis of $\mathbb{Z}^2$.

The MA-LBR operator, associated to a stencil V, is defined by

$$\mathcal{D}_V^{\mathrm{LBR}} u(x) := \min_{\substack{(e,f,g)\in V^3\\ \text{superbase}}} h(\Delta_e^+ u(x), \Delta_f^+ u(x), \Delta_g^+ u(x)) \tag{1.7}$$

where for $a,b,c \in \mathbb{R}_+$ we define

$$h(a,b,c) := \begin{cases} bc \text{ if } a \geq b+c, \text{ and likewise permuting } a,b,c, \\ \frac{1}{2}(ab+bc+ca) - \frac{1}{4}(a^2+b^2+c^2) \text{ otherwise.} \end{cases} \tag{1.8}$$

Remark 1.10 provides a geometric interpretation for these at first abstruse formulas.

Theorem 1.5. *The operators $\mathcal{D}_V^{WS}$ and $\mathcal{D}_V^{LBR}$, defined in* (1.6) *and* (1.7)*, are DE2.*

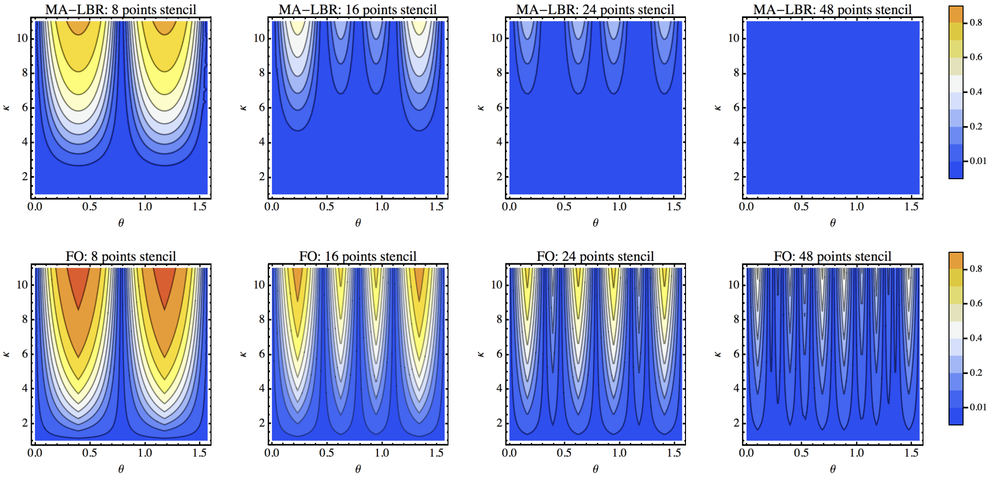

FIGURE 3. Relative consistency error $(\mathcal{D}(u_M) - \det(M))/\mathcal{D}(u_M)$ for quadratic functions (1.9), with the schemes MA-LBR (top) and WS (bottom), using the stencils of Figure 1. See (1.12) for the parametrization $M(\kappa, \theta)$ of symmetric matrices, by their condition number κ^2 and their orientation θ. Note that the MA-LBR consistency error vanishes for a large set of matrices.

Proof. The following functions are non-decreasing in each variable and locally Lipschitz: the positive part $\mathbb{R} \to \mathbb{R}_+ : \delta \mapsto \delta^+ := \max\{0, \delta\}$, the product of non-negative numbers $\times : \mathbb{R}_+^2 \to \mathbb{R}_+$, the function $h : \mathbb{R}_+^3 \to \mathbb{R}_+$ as shown in Lemma 3.6, and the minimum $\mathbb{R}^n \to \mathbb{R} : (h_1, \cdots, h_n) \mapsto \min(h_1, \cdots, h_n)$ with $n := \#(V)$. By composition, $\mathcal{D}_V^{\mathrm{WS}} u(x)$ and $\mathcal{D}_V^{\mathrm{LBR}} u(x)$ are non-decreasing, locally Lipschitz functions of the second order differences $\Delta_e u(x)$, which concludes the proof. □

Outline. We discuss in §1.1 the consistency of the MA-LBR, and show in particular that a finite stencil is sufficient to achieve consistency for all quadratic functions of condition number below a given bound. For more simplicity and efficiency we introduce in §1.2 an automatic stencil construction for the MA-LBR, which is adaptive, local, anisotropic, parameter free, and has good consistency guarantees. The proofs of the results appearing in §1.1 and §1.2 are postponed to §2 and §3 respectively.

Notation. For each $e = (a, b) \in \mathbb{R}^2$ we denote $e^\perp := (-b, a)$. If $e = (a, b) \in \mathbb{Z}^2$ then $\gcd(e) := \gcd(a, b)$. Given pairwise distinct $x_1, x_2, x_3 \in \mathbb{R}^2$, we denote by $[x_1, x_2]$ the segment of endpoints x_1, x_2, and by $[x_1, x_2, x_3]$ the triangle of vertices x_1, x_2, x_3.

1.1. **Consistency.** The consistency analysis of the numerical schemes FD, WS and the MA-LBR reveals significant differences. We denote by S_2 the collection of symmetric matrices of size 2×2, and by S_2^+ those which are positive definite. For each $M \in S_2^+$ we introduce a quadratic map $u_M \in \mathbb{U}$, defined by $u_M(x) := \langle x, Mx \rangle / 2$, $x \in X \cup \partial\Omega$. Since the second order difference operator Δ_e is consistent, for any $e \in \mathbb{Z}^2$, it is exact for u_M. Summarizing one has

$$u_M(x) := \frac{1}{2}\langle x, Mx \rangle, \qquad \Delta_e u_M(x) = \langle e, Me \rangle. \tag{1.9}$$

Definition 1.6. The consistency set of an operator $\mathcal{D}$ is the collection of matrices $M \in S_2^+$ for which $\mathcal{D}(u_M) = \det(M)$, identically on X.

Remark 1.7 (Consistency error). Let $\mathcal{D}$ be DE2 scheme, so that $\mathcal{D}u(x)$ is a non-decreasing, locally Lipschitz function of the second order differences $\Delta_e u(x)$, $e \in V$. For each $\varepsilon > 0$ let $\mathcal{D}^\varepsilon u(x)$ be the same function of the rescaled second order differences $\Delta_e^\varepsilon u(x) := \varepsilon^{-2}(u(x+\varepsilon e) - 2u(x) + u(x - \varepsilon e))$, $e \in V$. Given fixed $u \in C^4(\Omega)$, $x \in \Omega$, and $M := \nabla^2 u(x)$, a Taylor development yields as $\varepsilon \to 0$

$$|\Delta_e^\varepsilon u(x) - \langle e, Me\rangle| \leq C\|e\|^2\varepsilon^2, \tag{1.10}$$

where C depends on $\nabla^4 u(x)$. Hence by Lipschitz regularity, and since $\Delta_e u_M(x) = \langle e, Me\rangle$,

$$|\mathcal{D}^\varepsilon u(x) - \mathcal{D}u_M(x)| \leq C'\varepsilon^2. \tag{1.11}$$

If M lies in the consistency set of $\mathcal{D}$, then as $\varepsilon \to 0$ we obtain $\mathcal{D}^\varepsilon u(x) \to \mathcal{D}u_M(x) = \det(M) = \det(\nabla^2 u(x))$ as desired. On the other hand, if M does not lie in the consistency set of $\mathcal{D}$, then the limit $\mathcal{D}u_M(x)$ is not the desired Monge-Ampère operator $\det(\nabla^2 u(x))$: the scheme is inconsistent, and its error is independent of the scale ε.

One easily checks that the consistency set of the finite differences discretization $\mathcal{D}^{\mathrm{FD}}$ (see (1.4)) is the whole S_2^+. In fact the identity $\mathcal{D}^{\mathrm{FD}}(u_M) = \det(M)$ also holds for non-definite matrices $M \in S_2$, although they are irrelevant for our application. Since that scheme is not DE, this consistency does not imply convergence results. As illustrated in Figures 3 and 4, schemes WS and MA-LBR have in contrast non-trivial consistency sets, depending on the chosen stencil. Matrices $M \in S_2^+$ are parameterized in these figures by their condition number $\kappa^2 \in [1, \infty[$ and the orientation $\theta \in [0, \pi]$ of their first eigenvector e_θ:

$$M(\kappa, \theta) = \kappa^{-1}\, e_\theta \otimes e_\theta + \kappa\, e_\theta^\perp \otimes e_\theta^\perp, \qquad \text{with } e_\theta = (\cos\theta, \sin\theta). \tag{1.12}$$

The consistency analysis of $\mathcal{D}_V^{\mathrm{WS}}$ is based on Hadamard's theorem [9]: for all $M \in S_2^+$ and any pair $(f, g) \in (\mathbb{R}^2)^2$ of non-zero orthogonal vectors, one has $\langle f, Mf\rangle\langle g, Mg\rangle \geq \|f\|^2\|g\|^2 \det(M)$, with equality iff f and g are eigenvectors of M. As a result, scheme $\mathcal{D}_V^{\mathrm{WS}}$ is only consistent on a negligible subset of S_2^+: those matrices whose eigenvectors lie in V; see Figures 3 and 4. From a theoretical standpoint, convergence results are obtained in [9] by increasing the stencil size, up to infinity, as the discretization grid scale tends to zero. In practical cases, finding the optimal stencil size is non-trivial; see §4.

The key concept in the MA-LBR consistency analysis is the notion of M-obtuse superbase, which originates from lattice geometry [5] (a lattice is a discrete subgroup of $\mathbb{R}^n$ containing a basis, such as $\mathbb{Z}^n$). It was already applied to PDE discretizations in [7,19].

Definition 1.8. Let $M \in S_2^+$. A superbase (e_0, e_1, e_2) of $\mathbb{Z}^2$ is said to be M-obtuse iff $\langle e_i, Me_j\rangle \leq 0$ for all $0 \leq i < j \leq 2$.

Theorem 1.9 (Consistency). *A matrix $M \in S_2^+$ is in the consistency set of $\mathcal{D}_V^{LBR}$ iff there exists $(e, f, g) \in V^3$ which form an M-obtuse superbase.*

The following remark attempts to give a geometrical interpretation of the function (1.8) and of Theorem 1.9. The results of this section, Theorem 1.9, Remark 1.10 and Theorem 1.11, are established in §2.

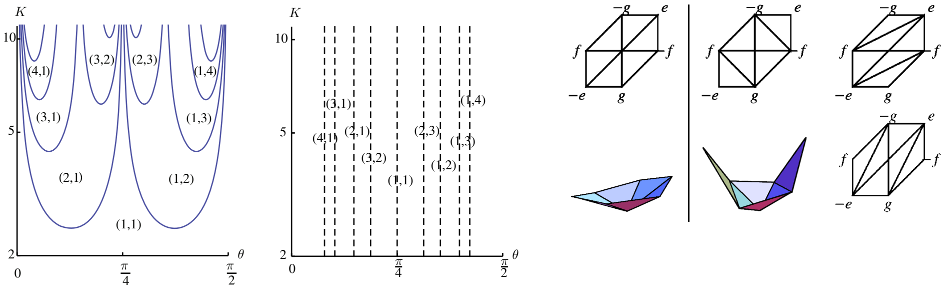

FIGURE 4. Left: element e of largest euclidean norm of an $M(\kappa, \theta)$-obtuse superbase. Center left: an eigenvector e of $M(\kappa, \theta)$. The consistency set of $\mathcal{D}_V^{\mathrm{LBR}}$ (resp. $\mathcal{D}_V^{\mathrm{WS}}$) is the union of the regions (resp. the dashed lines) corresponding to all stencil elements $e \in V$. Right: triangulations mentioned in Remark 1.10, and some 3D visualisations of the maximal convex extension U.

Remark 1.10 (Geometric interpretation). Let (e, f, g) be a superbase of $\mathbb{Z}^2$. Let $M \in S_2^+$, let $u := u_M$, and let U be the maximal convex map bounded above by u at the points $x, x \pm e, x \pm f, x \pm g$. Then $h(\Delta_e^+ u(x), \Delta_f^+ u(x), \Delta_g^+ u(x)) = \mathrm{Area}(\partial U(x))$ (the Lebesgue measure of the subgradient of U at x, which is a natural relaxation of the Monge-Ampère operator [12]). The map U is polygonal, on one of the four triangulations illustrated in Figure 4. The identity $|\partial U(x)| = \det(\nabla^2 u(x)) = \det(M)$ holds for the first triangulation only, which corresponds to an M-obtuse superbase (e, f, g).

Strikingly, one cannot hope for a DE2 scheme more localized than the MA-LBR. Finding well localized numerical schemes involving small stencils is a natural objective [13].

Theorem 1.11 (Minimality). *Let $\mathcal{D}$ be a DE2 scheme with stencil V. If the consistency set of $\mathcal{D}$ contains the neighborhood of a matrix $M \in S_2^+$, then there exists $(e, f, g) \in \mathrm{Hull}(V)^3$ which form an M-obtuse superbase.*

The following algorithm and proposition, dating back to Selling [5, 23], constructively show the existence of an M-obtuse superbase for each $M \in S_2^+$, without which Theorems 1.9 and 1.11 would be mostly vacuous. It is worth noting that this algorithm extends to dimension three [5]. Proposition 1.12 also immediately implies that all matrices $M \in S_2^+$ with condition number $\|M\|\|M^{-1}\| \leq \kappa^2$ are simultaneously in the consistency set of the MA-LBR operator $\mathcal{D}_V^{\mathrm{LBR}}$ with stencil

$$V := \{e \in \mathbb{Z}^2;\ \gcd(e) = 1, \|e\| \leq 2\kappa\}. \tag{1.13}$$

Algorithm 1 Construction of an M-obtuse superbase (Selling [5]).

```
Initialize $e_0 \leftarrow (-1,-1)$,  $e_1 \leftarrow (1,0)$,  $e_2 \leftarrow (0,1)$. (Or any other initial superbase.)
While the superbase $(e_0, e_1, e_2)$ is not $M$-obtuse do
    Find $0 \leq i < j \leq 2$ such that $\langle e_i, M e_j \rangle > 0$, and
    set $(e_0, e_1, e_2) \leftarrow (e_i - e_j,\ e_j,\ -e_i)$.
```

In order to analyze this algorithm, we associate to each $M \in S_2^+$ the norm

$$\|e\|_M := \sqrt{\langle e, Me\rangle}, \quad e \in \mathbb{R}^2.$$

Proposition 1.12 (Existence of an M-obtuse superbase [23]). *Let $M \in S_2^+$, and let $\kappa := \sqrt{\|M\| \|M^{-1}\|}$. Algorithm 2 terminates in at most κ^2 steps, and the final state of (e_0, e_1, e_2) is an M-obtuse superbase. Furthermore $\|e_i\| \leq 2\kappa$ for each $0 \leq i \leq 2$.*

Proof. Consider a step n of Selling's algorithm, which is *not* the last one. The next superbase $(e_0^{n+1}, e_1^{n+1}, e_2^{n+1})$, which equals $(e_i^n - e_j^n, e_j^n, -e_i^n)$ for some $0 \leq i < j \leq 2$, is therefore *not* obtuse. At least one of the following three scalar products is thus positive: with $s_n := \langle e_i^n, Me_j^n\rangle$,

$$\begin{aligned}\langle e_0^{n+1}, Me_1^{n+1}\rangle &= s_n - \|e_j^n\|_M^2,\\ \langle e_0^{n+1}, Me_2^{n+1}\rangle &= s_n - \|e_i^n\|_M^2,\\ \langle e_1^{n+1}, -Me_2^{n+1}\rangle &= -s_n.\end{aligned}$$

By construction $s_n > 0$, which excludes the third case, hence $s_n \geq \min\{\|e_i^n\|_M^2, \|e_j^n\|_M^2\} \geq \|M^{-1}\|^{-1}$. On the other hand, denoting $\mathcal{E}(e_0, e_1, e_2) := \|e_0\|_M^2 + \|e_1\|_M^2 + \|e_2\|_M^2$ we obtain

$$\mathcal{E}(e_0^n, e_1^n, e_2^n) - \mathcal{E}(e_0^{n+1}, e_1^{n+1}, e_2^{n+1}) = \|e_i^n + e_j^n\|_M^2 - \|e_i^n - e_j^n\|_M^2 = 4s_n.$$

Summing the previous identity over $\{0, \cdots, n\}$ yields

$$0 \leq \mathcal{E}(e_0^{n+1}, e_1^{n+1}, e_2^{n+1}) = \mathcal{E}(e_0^0, e_1^0, e_2^0) - 4 \sum_{0 \leq i \leq n} s_n \leq 4\|M\| - 4(n+1)\|M^{-1}\|^{-1};$$

hence $n+1 \leq \|M\| \|M^{-1}\| = \kappa^2$ as announced. At the final step N, the continuation criterion in the while loop is false, hence (e_0^N, e_1^N, e_2^N) is M-obtuse. Finally, denoting $\mathcal{F}(e_0, e_1, e_2) := \|e_0\|^2 + \|e_1\|^2 + \|e_2\|^2$ we obtain

$$\|M^{-1}\|^{-1}\mathcal{F}(e_0^N, e_1^N, e_2^N) \leq \mathcal{E}(e_0^N, e_1^N, e_2^N) \leq \mathcal{E}(e_0^0, e_1^0, e_2^0) \leq \|M\|\mathcal{F}(e_0^0, e_1^0, e_2^0) = 4\|M\|,$$

which immediately implies the announced bound on the obtuse superbase elements norm. □

The MA-LBR consistency error is typically smaller than with the WS scheme, for a given stencil V; see Figure 4. Furthermore, while the WS consistency is an asymptotic property, depending on the stencil angular resolution, the MA-LBR has in contrast a consistency set of non-empty interior, and its elements can be identified with a simple test; see Theorem 1.9. Unfortunately, choosing the MA-LBR effective stencil V before a numerical simulation remains at this point a puzzle for the practitioner. The option (1.13) is not practical because: (a) uniform bounds κ^2 on the hessian matrix $\nabla^2 u$ condition number of solutions to (1.1) are seldom available; (b) even if such a bound κ is available, the set (1.13) can be quite large, with cardinality $\gtrsim \kappa^2$. This becomes an issue if the bound is pessimistic or if the solution hessian $\nabla^2 u$ does degenerate in some places, such as along the domain boundary $\partial\Omega$. A third issue (c) is that there is no clear way to a posteriori validate the choice of a given stencil: would the numerical solution be improved with a larger one?

Selling's algorithm, in contrast with the inefficiency of (1.13), adaptively produces an M-obtuse superbase in only a few iterations typically. We present in the next section an adaptive, anisotropic, parameter free and guaranteed stencil refinement algorithm which eliminates the implementation difficulties (a), (b), (c) above. Under the hood, it amounts to an adaptation of Selling's algorithm to non-quadratic functions; see §3.3.

1.2. **Hierarchical stencil refinement.** The previous section fully characterized the consistency set of the MA-LBR operator $\mathcal{D}_V^{\mathrm{LBR}}$, associated to a stencil V. Larger stencils provide consistency on larger collections of matrices, as established in Theorem 1.9, and illustrated on Figure 4. Excessively large stencils are however impractical, since the CPU cost of evaluating the MA-LBR operator (1.7) is proportional to their cardinality. Adapting Selling's obtuse superbase construction, Algorithm 1, we show that one can emulate an MA-LBR with extremely large stencils for a limited numerical cost.

Our adaptive variant of the MA-LBR operator is defined by Algorithm 2 below, which is 6 lines long and only involves elementary operations. Its analysis (and Definition 1.18 of a mild structural constraint on stencils) relies on an arithmetic construction named the Stern-Brocot tree, already used in [2, 17] for the discretization of anisotropic PDEs. Definitions 1.14 and 1.16 introduce this structure. Propositions 1.13 and 1.17 are variants of commonly known facts on the Stern-Brocot tree, whose proof is, for completeness, presented in the appendix.

Proposition 1.13. *The identity $e = f + g$ defines a one to one correspondance between:*

- *Vectors $e = (a, b) \in \mathbb{Z}^2$, such that $\gcd(a, b) = 1$ and $ab \neq 0$.*
- *Direct acute bases (f, g) of $\mathbb{Z}^2$ (i.e. $(f, g) \in (\mathbb{Z}^2)^2$, $\det(f, g) = 1$ and $\langle f, g\rangle \geq 0$).*

Definition 1.14. We emphasize the unique decomposition introduced in Proposition 1.13 by using the notation $e = f \oplus g$. Whenever we write $e = f \oplus g$, we implicitly limit our attention to vectors e satisfying the assumptions of Proposition 1.13.

For instance $(7, 5) = (3, 2) \oplus (4, 3)$, $(3, 2) = (2, 1) \oplus (1, 1)$, and $(1, 1) = (1, 0) \oplus (0, 1)$. If $e = f \oplus g$, then $(e, -f, -g)$ is a superbase of $\mathbb{Z}^2$; all superbases happen to be of that form, up to a permutation of their elements; see Lemma A.3. The next proposition shows how to generate numerous decompositions of the form of Proposition 1.13.

Proposition 1.15. *If $e = f \oplus g$, then $f + e = f \oplus e$ and $e + g = e \oplus g$.*

Proof. We check $\det(f, e) = \det(f, f + g) = \det(f, g) = 1$, and $\langle f, e\rangle = \langle f, f + g\rangle = \|f\|^2 + \langle f, g\rangle \geq 0$. Hence (f, e) is a direct acute basis of $\mathbb{Z}^2$. Likewise for (e, g). □

Definition 1.16. We introduce a graph $\mathbb{T}$, with vertices $\{e \in \mathbb{Z}^2;\ \gcd(e) = 1\}$, and edges $e \to f \oplus e$ and $e \to e \oplus g$ for each $e = f \oplus g$.

We say that an edge $e \to e'$ of $\mathbb{T}$ leaves from e and arrives at e'. We denote by V_8 the eight points stencil illustrated in Figure 1 (left):

$$V_8 := \{(a, b) \in \{-1, 0, 1\}^2;\ ab \neq 0\}. \tag{1.14}$$

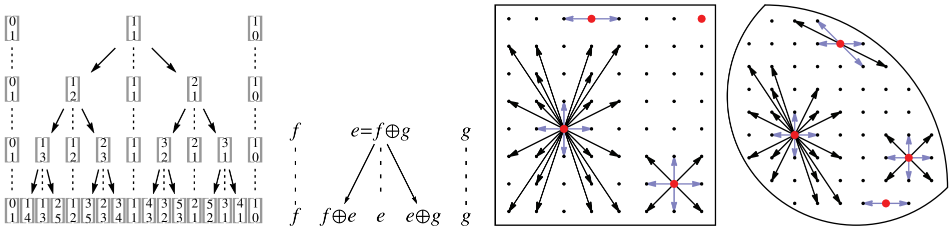

FIGURE 5. Stern-Brocot binary tree (left) and its local structure (center left). Right: some convex domains and the associated discretization grids. Plain black arrows: set $\mathcal{V}_\Omega(x)$, for some $x \in X$. Light blue arrows: vectors $e \in \mathbb{T} \setminus \mathcal{V}_\Omega(x)$ such that $x \pm e \in \Omega$, as in property (Reachability) of stencils.

Proposition 1.17. *The set V_8 has one element in each connected component of $\mathbb{T}$. The four points of the form $(\pm 1, 0)$ or $(0, \pm 1)$ are isolated. The four other points are the root of complete infinite binary trees, each one entirely contained in a quadrant of the plane.*

The Stern-Brocot tree is the subgraph corresponding to the first quadrant, with vertices $\mathbb{T}^+ := \{e = (a, b) \in \mathbb{T};\ a > 0,\ b > 0\}$; see Figure 5. This complete infinite binary tree originates from arithmetic, and in the literature a vertex $e = (a, b) \in \mathbb{T}^+$ is often identified with the positive irreducible fraction a/b.

The MA-LBR adaptive variant, presented below, requires stencils with a special structure. For each $x \in X$ we introduce the set

$$\mathcal{V}_\Omega(x) := \{e = f \oplus g;\ x \pm e, x \pm f, x \pm g \in \Omega\}, \tag{1.15}$$

where "$x \pm e \in \Omega$" stands for "$x + e \in \Omega$ and $x - e \in \Omega$". Note that the continuous domain Ω could in (1.15) be replaced with the discrete one $X := \Omega \cap \mathbb{Z}^2$: indeed for any $z \in \mathbb{Z}^2$, one has $z \in \Omega$ iff $z \in X$. Any subset V of $\mathbb{T}$ is regarded as a subgraph of $\mathbb{T}$, equipped with all edges of $\mathbb{T}$ having their endpoints in V.

Definition 1.18. A family of stencils $\mathcal{V}$ is the data, for each $x \in X$, of a stencil $\mathcal{V}(x)$ satisfying $V_8 \subseteq \mathcal{V}(x) \subseteq \mathbb{T}$ and the additional structural properties:

- (Hierarchy) The set V_8 has an element in each connected component of $\mathcal{V}(x)$.
- (Reachability) $\mathcal{V}(x)$ contains each $e \in \mathbb{T} \setminus \mathcal{V}_\Omega(x)$ such that $x \pm e \in \Omega$.

Aside from these minor constraints, the choice of stencils $\mathcal{V}(x)$, $x \in X$, is left to the user, as in other wide stencil methods [9, 20]. For instance, one may use a position independent stencil $\mathcal{V}(x) = V$, where V is one of the possibilities presented in Figure 1. (Note that when a large stencil is used close to the boundary $\partial\Omega$, the second order difference $\Delta_e u(x)$ is defined by (1.3) instead of (1.2).) A distinguishing feature of our method is that stencils can be adaptively completed at run time, using Algorithm 2, in a parameter free and guaranteed manner; see Theorem 1.21. We rely on this feature in our numerical experiments and define $\mathcal{V}(x) := V_8$; see Figure 1 (leftmost) for all points $x \in X$ except on a layer of four pixels along the boundary $\partial\Omega$, where the stencil completion procedure is less efficient and where $\mathcal{V}(x)$ is chosen as the 48 points stencil illustrated in Figure 1 (rightmost).

These structural constraints are in practice not hard to satisfy. Condition (Hierarchy) is natural in view of the Stern-Brocot tree structure and is satisfied by all stencils illustrated in Figure 1. Condition (Reachability) ensures that the stencils can be extended, in the sense of Proposition 1.20 below. It is vacuous for all $x \in X$ at distance ≥ 1 from $\partial\Omega$ in general, and entirely vacuous in the case of a box domain; see Proposition 3.4 and Corollary 3.5.

The sets $\mathcal{V}_\Omega(x)$, $x \in X$, are small or empty when x is close to $\partial\Omega$, but typically huge when x is far from Ω; see Figure 5. They *do not* constitute a family of stencils, but they can be used to extend an existing family of stencils, as in the next definition.

Definition 1.19. To each family of stencils $\mathcal{V}$, we associate the family of sets $\overline{\mathcal{V}}$ defined by $\overline{\mathcal{V}}(x) := \mathcal{V}(x) \cup \mathcal{V}_\Omega(x)$, $x \in X$.

Proposition 1.20 (Extension of stencils)**.** *If $\mathcal{V}$ is a family of stencils, then so is $\overline{\mathcal{V}}$.*

We next introduce the MA-LBR operator $\mathcal{D}_\mathcal{V}$ associated to a family $\mathcal{V}$ of stencils, as well as a hierarchical variant $\overline{\mathcal{D}}_\mathcal{V}$. (The expected $^{\mathrm{LBR}}$ superscript is omitted for readability.)

$$\mathcal{D}_\mathcal{V} u(x) := \min_{\substack{(e,f,g)\in\mathcal{V}(x)^3\\ \text{superbase}}} h(\Delta_e^+ u(x), \Delta_f^+ u(x), \Delta_g^+ u(x)). \tag{1.16}$$

Algorithm 2 Hierarchical operator $\overline{\mathcal{D}}_\mathcal{V} u(x)$ (the final value of $\mathbb{D}$).

```
Initialize a variable $f \leftarrow (1,0)$, and list $G \leftarrow [(0,1),(-1,0)]$. Set also $\mathbb{D} \leftarrow +\infty$.
While $G$ is non-empty do
    Denote by $g$ the first element of $G$, and set $e := f + g$.
    If $e \in \mathcal{V}(x)$ or $[e \in \mathcal{V}_\Omega(x)$ and $\Delta_e u(x) < \Delta_f u(x) + \Delta_g u(x)]$   (Refinement test)
        then prepend $e$ to $G$, and set $\mathbb{D} \leftarrow \min\{\mathbb{D}, h(\Delta_e^+ u(x), \Delta_f^+ u(x), \Delta_g^+ u(x))\}$
        else remove $g$ from $G$ and set $f \leftarrow g$.
```

Our main result, Theorem 1.21, states that the MA-LBR operator $\mathcal{D}_{\overline{\mathcal{V}}}$ associated to the large stencils $\overline{\mathcal{V}}$ coincides in all cases of interest with the hierarchical, adaptive variant $\overline{\mathcal{D}}_\mathcal{V}$. Algorithm 2 amounts to a depth-first transversal of a finite subtree of the Stern-Brocot tree; see §3.2 and [17], where a similar approach is used for the discretization of Hamilton-Jacobi PDEs. This subtree is characterized by the stopping criterion (Refinement test), allowing us to reject useless branches of $\mathbb{T}$ where the minimum (1.16) defining $\mathcal{D}_{\overline{\mathcal{V}}} u(x)$ cannot be attained. In the case of a quadratic map u_M, $M \in S_2^+$, Algorithm 2 explores a single branch of the Stern-Brocot tree, just as Selling's algorithm; see §3.3.

We say that a property holds "on X" iff it holds at each point of X.

Theorem 1.21 (Adaptive pruning equals extensive sweeping)**.** *Let $\mathcal{V}$ be a family of stencils, and let $u \in \mathbb{U}$. If $\overline{\mathcal{D}}_\mathcal{V} u > 0$ on X or $\mathcal{D}_{\overline{\mathcal{V}}} u > 0$ on X, then we have $\overline{\mathcal{D}}_\mathcal{V} u = \mathcal{D}_{\overline{\mathcal{V}}} u$ on X.*

The identity $\overline{\mathcal{D}}_\mathcal{V} u = \mathcal{D}_{\overline{\mathcal{V}}} u$ may break down when these two operators vanish at some points of X. Theorem 1.21 shows that the adaptive operator $\overline{\mathcal{D}}_\mathcal{V}$ is suitable if the density ρ in (1.1) is everywhere positive, as was assumed. The proposed method, with or without adaptivity, does not work satisfactorily if the density ρ vanishes; see §4.

Consider a smooth function $U : \Omega \to \mathbb{R}$ and a point $x \in \Omega$ such that $\det(\nabla^2 U(x)) > 0$. Then U is convex (or concave) on a neighborhood of x. The next proposition establishes a discrete analog of this property. Consider $u \in \mathbb{U}$ and a family $\mathcal{V}$ of stencils. The discrete counterpart of $\det(\nabla^2 U(x)) > 0$ is $\mathcal{D}_{\mathcal{V}} u(x) > 0$, while the counterpart of the convexity of U locally around x is the positivity of the second order differences centered at x: $\Delta_e u(x) > 0$, $e \in \mathcal{V}(x)$.

Proposition 1.22 (Discrete convexity). *Let $u \in \mathbb{U}$, let $\mathcal{V}$ be a family of stencils on X, and let $x \in X$. Then*

$$\mathcal{D}_{\mathcal{V}} u(x) > 0 \quad \Longleftrightarrow \quad \forall e \in \mathcal{V}(x),\ \Delta_e u(x) > 0.$$

Oberman [21] numerically addressed variational problems posed on the cone of convex functions by imposing the positivity of second order differences, $\Delta_e u(x)$ for all points $x \in X$ and all vectors e within some given stencil V. It is also known (see Appendix A of [18]) that any discrete map $u : X \to \mathbb{R}$ satisfying $\Delta_e u(x) \geq 0$ whenever $x, x \pm e \in X$ needs to coincide $u_{|X'} = U_{|X'}$ with a global convex function $U : \Omega \to \mathbb{R}$ on the subsampled grid $X' := X \cap (2\mathbb{Z}^2)$ of points with even coordinates.

Remark 1.23. Adaptivity in PDE discretizations often refers to the context where a sequence u_n of discrete maps is generated along an iterative procedure, as well as a sequence $\mathcal{D}_n$ of operators, and $\mathcal{D}_{n+1}$ depends on u_n. Our understanding in this paper is different: there is no underlying iteration, but a single operator $\mathcal{D}_{\overline{\mathcal{V}}}$ which is evaluated in a subtle and cheap way as $\overline{\mathcal{D}}_{\mathcal{V}}$.

2. Proof of consistency and minimality

We establish the results announced in §1.1 and related to the MA-LBR consistency and optimal locality. Theorem 1.9 (Consistency) and Remark 1.10 are proved in §2.1, and Theorem 1.11 (Minimality) in §2.2.

2.1. Consistency.

Our first result, Proposition 2.2, preceded by a technical lemma, shows that the MA-LBR operator (1.7) systematically overestimates the hessian determinant of quadratic functions. For any $M \in S_2^+$, defining u_M as in (1.9), one has $\mathcal{D}_V^{\mathrm{LBR}} u_M \geq \det(M)$ on X. Equality holds iff V contains an M-obtuse superbase, which establishes the announced Theorem 1.9 (Consistency). We denote

$$K := \{(a, b, c) \in \mathbb{R}_+^3;\ a \leq b + c,\ b \leq c + a,\ c \leq a + b\}, \tag{2.1}$$

$$h_1(a, b, c) := \frac{1}{2}(ab + bc + ca) - \frac{1}{4}(a^2 + b^2 + c^2). \tag{2.2}$$

Lemma 2.1. *Let $(a, b, c) \in \mathbb{R}_+^3$. Then $h(a, b, c) \geq h_1(a, b, c)$, with equality iff $(a, b, c) \in K$.*

Proof. If $(a, b, c) \in K$, then $h(a, b, c) = h_1(a, b, c)$ by definition (1.8). Otherwise, we may assume without loss of generality that $a > b + c$, so that $h(a, b, c) = bc$ and $h(a, b, c) - h_1(a, b, c) = \frac{1}{4}(a - b - c)^2 > 0$. □

Proposition 2.2. *Let $M \in S_2^+$, let (e_0, e_1, e_2) be a superbase of $\mathbb{Z}^2$, and let $\delta_i := \langle e_i, M e_i \rangle$ for $0 \leq i \leq 2$. Then $h(\delta_0, \delta_1, \delta_2) \geq \det(M)$. Equality holds iff $(\delta_0, \delta_1, \delta_2) \in K$, equivalently iff (e_0, e_1, e_2) is M-obtuse.*

Proof. Given a permutation $\{i,j,k\}$ of $\{0,1,2\}$ we compute

$$\delta_i - \delta_j - \delta_k = \langle e_j + e_k, M(e_j+e_k)\rangle - \langle e_j, Me_j\rangle - \langle e_k, Me_k\rangle = 2\langle e_j, Me_k\rangle.$$

Hence $(\delta_0,\delta_1,\delta_2) \in K$ iff the superbase (e_0,e_1,e_2) is M-obtuse. We prove in the following that $h_1(\delta_0,\delta_1,\delta_2) = \det(M)$, which in view of Lemma 2.1 concludes the proof.

Special case of the superbase $f_0 := (-1,-1)$, $f_1 := (1,0)$, $f_2 := (0,1)$, with $\mu_i := \langle f_i, Mf_i\rangle$. We get $\mu_0 = M_{11}+2M_{12}+M_{22}$, $\mu_1 = M_{11}$, $\mu_2 = M_{22}$. Inserting this into the expression (2.2) yields as announced $h_1(\mu_0,\mu_1,\mu_2) = M_{11}M_{22}-M_{12}^2 = \det(M)$.

General case. Let A be a matrix such that $Af_1 = e_1$ and $Af_2 = e_2$, so that by linearity $Af_0 = e_0$. Note that $|\det(A)| = |\det(e_1,e_2)/\det(f_1,f_2)| = 1$. We obtain $\delta_i = \langle f_i, A^{\mathrm{T}}MAf_i\rangle$, for all $0 \le i \le 2$, so that by the special case $h_1(\delta_0,\delta_1,\delta_2) = \det(A^{\mathrm{T}}MA) = \det(M)$. □

The rest of this subsection is devoted to the proof of Remark 1.10. Let (e_0,e_1,e_2) be a fixed superbase of $\mathbb{Z}^2$. For each $\delta = (\delta_0,\delta_1,\delta_2) \in \mathbb{R}^3_+$ we introduce a polygon $H(\delta)$, defined by linear inequalities, and some of its edges $E_i(\delta)$, $1 \le i \le 3$:

$$\begin{aligned} H(\delta) &:= \{l \in \mathbb{R}^2; \forall\, 1 \le i \le 3,\ |\langle l, e_i\rangle| \le \delta_i\},\\ E_i(\delta) &:= \{l \in H(\delta);\ \langle l, e_i\rangle = \delta_i\}. \end{aligned}$$

The area of $H(\delta)$ is computed in Corollary 2.4, and this polygon (properly scaled and translated) is identified with a subgradient set in Proposition 2.5, concluding the proof of Remark 1.10. The proof unfortunately gives little geometric insight, hence it could be skipped at first reading. Given $A \subseteq \mathbb{R}^n$, $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$ we use the notation $x + \alpha A := \{x + \alpha a; a \in A\}$.

Lemma 2.3. *Let $\delta \in \mathbb{R}^3_+$. Then $E_0(\delta)$ is a segment of length (i) $(\delta_1+\delta_2-\delta_0)\|e_0\|$ if $\delta \in K$, (ii) $2\delta_2\|e_0\|$ if $\delta_1 \ge \delta_0+\delta_2$, (iii) $2\delta_1\|e_0\|$ if $\delta_2 \ge \delta_0+\delta_1$, or (iv) 0 if $\delta_0 > \delta_1+\delta_2$ (a case where $E_0(\delta)$ is in fact empty).*

Proof. Let $x,y \in \mathbb{R}^2$ and let $l := xe_0 + ye_0^\perp$. One has $l \in E_0(\delta)$ iff

$$\langle xe_0 + ye_0^\perp, e_0\rangle = \delta_0 \text{ and } |\langle xe_0+ye_0^\perp, e_i\rangle| \le \delta_i,\ i \in \{1,2\}.$$

The equality is equivalent to $x = \delta_0/\|e_0\|^2$. Recall that $\langle e_0^\perp, e_1\rangle = \det(e_0,e_1) = 1$, and likewise $\langle e_0^\perp, e_2\rangle = \det(e_0,e_2) = -1$. Hence the inequalities respectively hold iff y belongs to the segments

$$S_1 := -\frac{\langle e_0,e_1\rangle\delta_0}{\|e_0\|^2} + [-\delta_1,\delta_1], \qquad S_2 := \frac{\langle e_0,e_2\rangle\delta_0}{\|e_0\|^2} + [-\delta_2,\delta_2].$$

Translating these two segments by $\langle e_0,e_1\rangle\delta_0/\|e_0\|^2$ yields $F_1 = [-\delta_1,\delta_1]$ and $F_2 = [-\delta_2-\delta_0, \delta_2-\delta_0]$. Finally the length of E_δ is

$$\|e_0\| \times \operatorname{length}(F_1 \cap F_2) = \|e_0\|\left(\min\{\delta_1,\delta_2-\delta_0\} + \min\{\delta_1,\delta_2+\delta_0\}\right)_+, \tag{2.3}$$

which coincides with the announced result. □

Corollary 2.4. *For any $\delta = (\delta_0,\delta_1,\delta_2) \in \mathbb{R}_3^+$, one has* $\operatorname{Area}(H(\delta)) = 4h(\delta)$.

Proof. The triangle $\operatorname{Hull}(E_i(\delta) \cup \{0\})$ has area $\frac{1}{2} \times \frac{\delta_i}{\|e_i\|} \times \operatorname{length}(E_i(\delta))$: half the height from the vertex at the origin, times the length of the opposite side. These three (possibly empty) triangles, with their opposites, partition $H(\delta)$. From this point the result follows from Lemma 2.3 and an easy calculation. □

Proposition 2.5. *Let $M \in S_2^+$, and let $\delta_i := \langle e_i, Me_i \rangle$, $1 \leq i \leq 3$. Let $x \in \mathbb{Z}^2$, and let U be the maximal convex map bounded above by u_M at the points x and $x \pm e_i$, $1 \leq i \leq 3$. Then $\partial U(x) = Mx + \frac{1}{2} H(\delta)$.*

Proof. For any $g \in \mathbb{R}^2$, the following are equivalent:

- $g \in \partial U(x)$.
- $U(x) + \langle g, p - x \rangle \leq U(p)$, for all points p of the hexagon of vertices $x \pm e_i$, $1 \leq i \leq 3$.
- $u_M(x) + \langle g, p - x \rangle \leq u_M(p)$, for all $p = x + \varepsilon e_i$, $\varepsilon \in \{-1, 1\}$, $1 \leq i \leq 3$.

In order to further simplify this expression, we write $g = Mx + l/2$, $p = x + e$, where $e = \pm e_i$, $1 \leq i \leq 3$, and insert the expression (1.9) of u_M. The following are then equivalent:

$$\begin{aligned} u_M(x) + \langle g, p - x \rangle &\leq u_M(p), \\ \langle x, Mx \rangle + 2\langle Mx + l/2, (x + e) - x \rangle &\leq \langle x + e, M(x + e) \rangle, \\ \langle l, e \rangle &\leq \langle e, Me \rangle. \end{aligned}$$

We recognize the inequalities defining $H(\delta)$, and the announced result follows. □

Proof of Remark 1.10. Proposition 2.5 and Corollary 2.4 imply as announced that $\operatorname{Area}(\partial U(x)) = \frac{1}{4} \operatorname{Area}(H(\delta)) = h(\delta)$. By Proposition 2.2 one has $h(\delta) = \det M$ iff $\delta \in K$, which by Lemma 2.3 means that $H(\delta)$ is a hexagon: each edge $E_i(\delta)$ has a positive length (we exclude here for simplicity the limit case $\delta \in \partial K$).

The map U is polygonal on a triangulation with vertices x, $x \pm e_i$, $1 \leq i \leq 3$, which is symmetric with respect to x. Only four such triangulations exist, as illustrated in Figure 4, and only the first one leads to a hexagonal subgradient $\partial U(x)$, since the subgradient has one vertex for each triangle containing x. This concludes the proof. □

2.2. **Minimality.** We prove Theorem 1.11 (Minimality), on the optimal locality of the MA-LBR. For that purpose we introduce some definitions and establish in Proposition 2.8 a minimality property of obtuse superbases.

Definition 2.6. We denote by $\operatorname{Cone}(f, g) := \{\alpha f + \beta g;\ \alpha, \beta \in \mathbb{R}_+\}$ the closed convex cone spanned by two elements $f, g \in \mathbb{R}^2$. We say that f, g are *trigonometrically consecutive* elements of a set $V \subseteq \mathbb{R}^2 \setminus \{0\}$ iff they are not collinear and no element of V lies in the interior of $\operatorname{Cone}(f, g)$.

Definition 2.7. A matrix $M \in S_2^+$ is said to be *generic* iff there exists no M-orthogonal basis of $\mathbb{Z}^2$ (i.e. $(f, g) \in (\mathbb{Z}^2)^2$ such that $|\det(f, g)| = 1$ and $\langle f, Mg \rangle = 0$).

Proposition 2.8. *Let $M \in S_2^+$, and let (e_0, e_1, e_2) be an M-obtuse superbase. Then for each $e \in \mathbb{Z}^2 \setminus \{\pm e_0, \pm e_1, \pm e_2\}$ with* $\gcd(e) = 1$ *one has*

$$\|e\|_M \geq \max\{\|e_0\|_M, \|e_1\|_M, \|e_2\|_M\}.$$

The inequality is strict if M is generic.

Proof. Consider the set $S := \{e_0, -e_2, e_1, -e_0, e_2, -e_1\}$, where for convenience elements are ordered trigonometrically, and some $e \in \mathbb{Z}^2 \setminus S$ with $\gcd(e) = 1$. Let f, g be trigonometrically consecutive elements of S such that $e \in \operatorname{Cone}(f, g)$. Since (f, g) is a basis of $\mathbb{Z}^2$, there exists $\alpha, \beta \in \mathbb{Z}$ such that $e = \alpha f + \beta g$. Since $e \in \operatorname{Cone}(f, g)$, we have $\alpha, \beta \geq 0$. Since $\gcd(e) = 1$, one has $\gcd(a, b) = 1$. Since $e \notin S$, one has

$\alpha, \beta \geq 1$. By construction of S one has $\langle f, Mg\rangle \geq 0$, hence $\|e\|_M^2 \geq \|f\|_M^2 + \|g\|_M^2 \geq \max\{\|f\|_M^2, \|g\|_M^2, \|f-g\|_M^2\} = \max\{\|e_0\|_M, \|e_1\|_M, \|e_2\|_M\}^2$ as announced. If M is generic, then $\langle f, Mg\rangle \neq 0$, thus $\langle f, Mg\rangle > 0$, hence inequalities are strict. □

We next study trigonometrically consecutive elements f, g of the stencil V of an operator $\mathcal{D}$ whose consistency set contains a given matrix M. Corollary 2.10, preceded by a technical lemma, identifies the sign of the scalar product $\langle f, Mg\rangle$.

Lemma 2.9. *Let $M \in S_2^+$, and let $f, g \in \mathbb{R}^2$ be non-collinear and such that $\langle f, Mg\rangle < 0$. Then there exists $N \in S_2$ such that (i) $\det(M + \delta N) > \det(M)$ for any sufficiently small $\delta > 0$, and (ii) $\langle e, Ne\rangle \leq 0$ for all $e \in \mathrm{Cone}(f, -g)$.*

Proof. Case $M = \mathrm{Id}$ and $\|f\| = \|g\| = 1$. We define $N \in S_2$ by the (non-definite) quadratic form $\langle x, Nx\rangle = \det(f, x)\det(x, g)$, $x \in X$. It has eigenvectors $f + g$ and $f - g$, by a symmetry argument, with respective eigenvalues $\lambda_0 := \det(f, g)^2/\|f + g\|^2$, and $-\lambda_1$ where $\lambda_1 := \det(f, g)^2/\|f - g\|^2$. Since $\langle f, g\rangle < 0$ we have $\|f + g\| < \|f - g\|$, hence $\lambda_0 > \lambda_1$ and therefore $\det(\mathrm{Id} + \delta N) = 1 + \delta(\lambda_0 - \lambda_1) - \delta^2\lambda_0\lambda_1 > 1$ for small positive δ as announced.

General case. Write $M = A^{\mathrm{T}}A$, for some 2×2 invertible matrix A, and take $N = A^{\mathrm{T}}N'A$ where N' is associated to Id, $Af/\|Af\|$, $Ag/\|Ag\|$. □

Corollary 2.10. *Let $\mathcal{D}$ be a DE2 operator with stencil V and whose consistency set contains the neighborhood of a matrix $M \in S_2^+$. If f, g are trigonometrically consecutive elements of V, then $\langle f, Mg\rangle \geq 0$.*

Proof. Assume for contradiction that $\langle f, Mg\rangle < 0$. Let $N \in S_2$ be given by Lemma 2.9, and let $M_\delta := M + \delta N$ for some small $\delta \geq 0$. An element $e \in V$ cannot belong to the interior of $\mathrm{Cone}(f, g)$ by definition, and neither to the interior of $\mathrm{Cone}(-f, -g)$ by symmetry of V. Hence it belongs to $\mathrm{Cone}(f, -g)$ or $\mathrm{Cone}(-f, g)$, which implies $\langle e, Ne\rangle \leq 0$. We have obtained that $\langle e, M_\delta e\rangle \leq \langle e, Me\rangle$ for all $e \in V$, so that by degenerate ellipticity $\det(M_\delta) = \mathcal{D}(u_{M_\delta}) \leq \mathcal{D}(u_M) = \det(M)$. This contradicts Lemma 2.9, which concludes the proof. □

Our following step, Corollary 2.12 preceded by a technical lemma, shows that without loss of generality one can assume that consecutive elements of a stencil V form bases of $\mathbb{Z}^2$.

Lemma 2.11. *Let $f, g \in \mathbb{Z}^2$, and let T be the triangle of vertices $0, f, g$. If $|\det(f, g)| > 1$, then T contains a point e distinct from its vertices and such that $\gcd(e) = 1$.*

Proof. Since $|\det(f, g)| > 1$ the map $(\alpha, \beta) \in \mathbb{Z}^2 \mapsto \alpha f + \beta g \in \mathbb{Z}^2$ is not surjective. Hence there exists $(\alpha, \beta) \in \mathbb{Q}^2$, at least one of them non-integer, such that $\alpha f + \beta g \in \mathbb{Z}^2$. Up to replacing (α, β) with $(\alpha - m, \beta - n)$, $(m, n) \in \mathbb{Z}^2$, we may assume that $\alpha, \beta \in [0, 1]$. Up to replacing (α, β) with $(1 - \alpha, 1 - \beta)$, we may assume that $\alpha + \beta \leq 1$. The point $e := \alpha f + \beta g \in \mathbb{Z}^2$ belongs to T and is distinct from its vertices. In the case where $\gcd(e) > 1$, we can replace it with $e/\gcd(e)$. □

Corollary 2.12. *Let $\mathcal{D}$ be a DE2 operator with stencil V and with a consistency set of non-empty interior. Then there exists a DE2 operator $\mathcal{D}'$ with stencil V', such that (i) $\mathcal{D}$ and $\mathcal{D}'$ have the same consistency set, (ii) $\mathrm{Hull}(V') \subseteq \mathrm{Hull}(V)$, and (iii) any two trigonometrically consecutive elements $f, g \in V'$ satisfy $|\det(f, g)| = 1$.*

Proof. Let $V' := \{e \in \mathrm{Hull}(V) \cap \mathbb{Z}^2;\, \gcd(e) = 1\}$. For each $e \in V$ one has $e' := e/\gcd(e) \in \mathrm{Hull}(\{-e, e\})$; hence $e' \in V'$ since V is symmetric w.r.t. the origin. Note that $\Delta_e u_M = \langle e, Me\rangle = \gcd(e)^2 \Delta_{e'} u_M$, for any $M \in S_2^+$, using (1.9). Constructing $\mathcal{D}'$ in terms of $\mathcal{D}$ is from this point straightforward.

Let $f, g \in V'$ be trigonometrically consecutive. If $|\det(f, g)| > 1$, then Lemma 2.11 provides a point $e \in [0, f, g] \subseteq \mathrm{Hull}(V)$ with $\gcd(e) = 1$, hence $e \in V'$; this contradicts our assumption on $f, g \in V'$. The case $\det(f, g) = 0$ is excluded by Definition 2.6, hence $|\det(f, g)| = 1$, which concludes the proof. □

Finally, we identify a condition under which a stencil V contains an M-obtuse superbase, and we conclude the proof of the announced Theorem 1.11.

Lemma 2.13. *Let $M \in S_2^+$, and let V be a stencil which contains some non-collinear elements, and such that any two trigonometrically consecutive $f, g \in V$ satisfy $\langle f, Mg\rangle > 0$ and $|\det(f, g)| = 1$. Then V contains an M-obtuse superbase.*

Proof. Let e be an element of an M-obtuse superbase, and let $f, g \in V$ be trigonometrically consecutive and such that $e \in \mathrm{Cone}(f, g)$. Note that f, g exist because V contains some non-collinear elements and is symmetric w.r.t. the origin. Since $|\det(f, g)| = 1$, one has $e = \alpha f + \beta g$ for some $\alpha, \beta \in \mathbb{Z}$. Since $e \in \mathrm{Cone}(f, g)$, we have $\alpha, \beta \geq 0$. Since $\gcd(e) = 1$, one has $\gcd(\alpha, \beta) = 1$. Assuming for contradiction that $\alpha, \beta \geq 1$, we obtain $\|e\|_M^2 > \|f\|_M^2 + \|g\|_M^2 > \max\{\|f\|_M^2, \|g\|_M^2, \|f - g\|_M^2\}$ since $\langle f, Mg\rangle > 0$. This contradicts Proposition 2.8; therefore $\alpha\beta = 0$. But then (α, β) equals $(1, 0)$ or $(0, 1)$, since $\gcd(\alpha, \beta) = 1$. Thus $e \in \{f, g\} \subseteq V$, which concludes the proof. □

Proof of Theorem 1.11. Let $\mathcal{D}$ be a DE2 operator with stencil V and whose consistency set contains the neighborhood of a generic matrix M. Let $\mathcal{D}'$ and V' be as described in Corollary 2.12. Let $f, g \in V'$ be trigonometrically consecutive; note that $|\det(f, g)| = 1$.

Case of a generic matrix $M \in S_2^+$. Corollary 2.10 states that $\langle f, Mg\rangle \geq 0$; hence $\langle f, Mg\rangle > 0$ since M is generic. The consistency assumption implies that V contains non-collinear elements, hence so does V'. Invoking Lemma 2.13 we find that $V' \subseteq \mathrm{Hull}(V)$ contains an M-obtuse superbase, as announced.

Case of a non-generic $M \in S_2^+$. Let $(M_n)_{n \geq 0}$, $M_n \in S_2^+$, be a sequence of generic matrices converging to M. By the previous point, $\mathrm{Hull}(V)$ contains an M_n-obtuse superbase (e_0^n, e_1^n, e_2^n) for all sufficiently large n. By Proposition 1.12 the elements of these superbases are bounded independently of n. Since superbases are discrete objects, infinitely many among this sequence are equal to some fixed $(e_0, e_1, e_2) \in (\mathbb{Z}^2)^3$, also contained in $\mathrm{Hull}(V)$ and which by continuity is an M-obtuse superbase. □

3. Proofs on hierarchical stencil refinement

We establish the results announced in §1.2. Propositions 1.20 (Stencil extension) and 1.22 (Discrete convexity) are proved in §3.1. Algorithm 2 is rephrased in §3.2 as a depth first search within the Stern-Brocot tree. Theorem 1.21 (Adaptive pruning equals extensive sweeping) is established in §3.3 in the quadratic case, and in §3.4 in the general case.

3.1. **Properties of stencils.**

We establish several properties of stencils announced in §1.2, starting with Proposition 1.20: any stencils $\mathcal{V}$ can be extended by union with the sets $\mathcal{V}_\Omega$. This requires two technical lemmas.

Lemma 3.1. *Let $e \in \mathbb{T} \setminus V_8$, $e = f \oplus g$. The graph $\mathbb{T}$ has exactly one edge arriving at e, which is either $f \to e$ or $g \to e$.*

Proof. The existence of a unique edge arriving at e follows from the description of $\mathbb{T}$, Proposition 1.17. Let $e' = f' \oplus g' \in \mathbb{T}$ be such that $e' \to e$. By definition of this graph structure e equals $f' \oplus e'$ or $e' \oplus g'$. By uniqueness of the decomposition $e' \in \{f, g\}$, which concludes the proof. □

Lemma 3.2. *Let $\mathcal{V}$ be a family of stencils, and let $x \in X$. Then any $e \in \mathbb{T}$ such that $x \pm e \in \Omega$ belongs to $\overline{\mathcal{V}}(x) := \mathcal{V}(x) \cup \mathcal{V}_\Omega(x)$.*

Proof. If $e \notin \mathcal{V}_\Omega(x)$, then $e \in \mathcal{V}(x)$ by (Reachability). □

Proof of Proposition 1.20. We consider a family $\mathcal{V}$ of stencils and show that $\overline{\mathcal{V}}$ also is one. The inclusion $\mathcal{V}(x) \subseteq \overline{\mathcal{V}}(x)$ implies (Reachability), as well as $V_8 \subseteq \overline{\mathcal{V}}(x)$. The inclusions $\mathcal{V}(x) \subseteq \mathbb{T}$ and $\mathcal{V}_\Omega(x) \subseteq \mathbb{T}$ imply that $\overline{\mathcal{V}}(x) \subseteq \mathbb{T}$. Only (Hierarchy) is thus left to prove.

Consider an edge $e' \mapsto e$ of $\mathbb{T}$, with $e \in \overline{\mathcal{V}}(x)$. Our objective is to show that $e' \in \overline{\mathcal{V}}(x)$. If $e \in \mathcal{V}(x)$, then this follows from (Hierarchy) for $\mathcal{V}$. Otherwise $e \in \mathcal{V}_\Omega(x)$; hence it admits a decomposition $e = f \oplus g$, and $e' \in \{f, g\}$ by Lemma 3.1. Then $x \pm e' \in \Omega$ by definition (1.15) of $\mathcal{V}_\Omega(x)$, and therefore $e' \in \overline{\mathcal{V}}(x)$ by Lemma 3.2. □

Our next proposition shows as announced in §1.2 that condition (Reachability), required for families of stencils, is vacuous for all points $x \in X$ far from the boundary $\partial\Omega$. Corollary 3.5 shows in addition that it is entirely vacuous if Ω is a box domain. For that purpose we need a technical lemma.

Lemma 3.3. *Let $e = f \oplus g$, and let us assume that e has positive coordinates. Then f, g belong to the triangle $[(1,0), (0,1), e]$.*

Proof. The proof appears in [18], but is reproduced in Appendix A.3 for completeness. □

Proposition 3.4. *Let $x \in X$ be such that $x \pm (1,0), x \pm (0,1) \in \Omega$. Then condition (Reachability) is vacuous for x, in the following sense: any $e \in \mathbb{T} \setminus \mathcal{V}_\Omega(x)$ such that $x \pm e \in \Omega$ must be of the form $(\pm 1, 0)$ or $(0, \pm 1)$, hence automatically $e \in V_8 \subseteq \mathcal{V}(x)$.*

Proof. Consider $x \in X$ and $e \in \mathbb{T}$, distinct from $(\pm 1, 0)$ and $(0, \pm 1)$ and such that $x \pm e \in \Omega$. Without loss of generality, we assume that both coordinates of e are positive and that $x = 0$. By Proposition 1.13 we may introduce the decomposition $e = f \oplus g$. By Lemma 3.3, $f, g \in [(1,0), (0,1), e] \subseteq \Omega$. Likewise $-f, -g \in [(-1,0), (0,-1), -e] \subseteq \Omega$. Thus $e \in \mathcal{V}_\Omega(x)$, which concludes the proof. □

Corollary 3.5. *Assume a box domain $\Omega =]a_1^-, a_1^+[\times]a_2^-, a_2^+[$. Then condition (Reachability) is vacuous for all $x \in X$, in the same sense as in Proposition* 3.4.

Proof. Let $x = (x_1, x_2) \in X$, and let $e = (e_1, e_2) \in \mathbb{T} \setminus \mathcal{V}_\Omega(x)$, distinct from $(\pm 1, 0)$ and $(0, \pm 1)$ and such that $x \pm e \in \Omega$. Since $\gcd(e) = 1$, both e_1 and e_2 are non-zero integers. Since $x \pm e \in \Omega$, we have $a_i^- < x_i - |e_i|$, $a_i^+ > x_i + |e_i|$, for each $i \in \{1,2\}$. Hence $x \pm (1,0), x \pm (0,1) \in \Omega$. Applying Proposition 3.4 we conclude the proof. □

We conclude this section with the proof of Proposition 1.22, in Corollaries 3.7 and 3.9, which ties the positivity of the MA-LBR operator with a local discretization of convexity.

Lemma 3.6. *The map h is non-decreasing in all its variables on $\mathbb{R}_+^3$. For all $a,b,c \in \mathbb{R}_+$ we have $\min\{ab,bc,ca\} \geq h(a,b,c) \geq \frac{3}{4}\min\{a,b,c\}^2$.*

Proof. First point. One easily checks that the piecewise definitions (1.8) of $h(a,b,c)$ agree on the interface ∂K (2.1), i.e. when $a = b + c$ or $b = a + c$ or $c = a + b$; hence h is continuous. We then compute $\nabla h(a,b,c) = (b+c-a,\ a+c-b,\ a+b-c)/2$ for all $(a,b,c) \in K$, and $\nabla h(a,b,c) = (0,c,b)$ when $a \geq b + c$ (resp. likewise permuting the roles of a,b,c). Hence the components of ∇h are non-negative everywhere, and therefore h is non-decreasing in all its variables.

Second point: If $a \geq b + c$, then $h(a,b,c) = bc$ by definition (1.8), and otherwise $h(a,b,c) \leq h(b+c,b,c) = bc$ by the first point. Hence $h(a,b,c) \leq bc$ for all $a,b,c \in \mathbb{R}^+$. Likewise $h(a,b,c) \leq bc$ and $h(a,b,c) \leq ca$, thus $h(a,b,c) \leq \min\{ab,bc,ca\}$ as announced. Finally, denoting $\delta := \min\{a,b,c\} > 0$ we obtain $h(a,b,c) \geq h(\delta,\delta,\delta) = \frac{3}{4}\delta^2$. □

Corollary 3.7. *Let $u \in \mathbb{U}$, let $\mathcal{V}$ be a family of stencils, and let $x \in X$. If $\Delta_e u(x) > 0$ for all $e \in \mathcal{V}(x)$, then $\mathcal{D}_\mathcal{V} u(x) > 0$.*

Proof. The operator value $\mathcal{D}_\mathcal{V} u(x)$ is the minimum (1.16) of a finite collection of terms of the form $h(\Delta_e^+ u(x), \Delta_f^+ u(x), \Delta_g^+ u(x))$, where $e,f,g \in \mathcal{V}(x)$, and which by Lemma 3.6 are positive. □

Lemma 3.8. *Let $\mathcal{V}$ be a family of stencils, and let $x \in X$. If $e \in \mathcal{V}(x)$ and $e = f \oplus g$, then $f,g \in \mathcal{V}(x)$.*

Proof. Denote $V := \mathcal{V}(x)$. We proceed by induction on the integer $\|e\|^2$. If $\|e\|^2 = 1$, then it admits no decomposition of the form $f \oplus g$. If $\|e\|^2 = 2$, then $e \in V_8$, and therefore $f,g \in V_8 \subseteq V$.

If $\|e\|^2 > 2$, then by Proposition 1.17 the graph $\mathbb{T}$ has an edge $e' \to e$. We write $e' = f' \oplus g'$. By (Hierarchy) one has $e' \in V$, and by induction $f', g' \in V$. By definition of $\mathbb{T}$ the vector e is either $f' \oplus e'$ or $e' \oplus g'$. Hence $\{f,g\} \subseteq \{e',f',g'\} \subseteq V$, which concludes the proof. □

Corollary 3.9. *Let $u \in \mathbb{U}$, let $\mathcal{V}$ be a family of stencils, and let $x \in X$. Then*

$$\mathcal{D}_\mathcal{V} u(x) = \min\{h(\Delta_e^+ u(x), \Delta_f^+ u(x), \Delta_g^+ u(x));\ e \in \mathcal{V}(x),\ e = f \oplus g\}. \tag{3.1}$$

Also, if $\mathcal{D}_\mathcal{V} u(x) > 0$, then $\Delta_e u(x) > 0$ for all $e \in \mathcal{V}(x)$.

Proof. First point. Note the symmetries (i) $\Delta_e u(x) = \Delta_{-e} u(x)$ for any $e \in \mathbb{Z}^2$, and (ii) $e \in \mathcal{V}(x)$ iff $-e \in \mathcal{V}(x)$ by Definition 1.2. We denote by D (resp. D') the left (resp. right) hand side of (3.1). If $e \in \mathcal{V}(x)$ and $e = f \oplus g$, then $(e,-f,-g)$ is a superbase of $\mathbb{Z}^2$ and $f,g \in \mathcal{V}(x)$ by Lemma 3.8; hence $D \leq D'$. Conversely let $(e,f,g) \in \mathcal{V}(x)^3$ be a superbase of $\mathbb{Z}^2$. Up to reordering these vectors we may assume that $\|e\| \geq \max\{\|f\|, \|g\|\}$ and $\det(f,g) = 1$. Then $-e = f \oplus g$ by Lemma A.3, which implies $D' \leq D$ and establishes (3.1).

Second point. Let $e \in \mathcal{V}(x)$. If $e = f \oplus g$, then $0 < \mathcal{D}_{\mathcal{V}} u(x) \leq h(\Delta_e^+ u(x), \Delta_f^+ u(x), \Delta_g^+ u(x))$ by (3.1), and therefore $\Delta_e u(x) > 0$ by Lemma 3.6. Otherwise, e is among $(\pm 1, 0)$ or $(0, \pm 1)$; hence we may choose $f, g \in V_8 \subseteq \mathcal{V}(x)$ such that e, f, g is a superbase. Using (1.16) and Lemma 3.6 we again obtain $\Delta_e u(x) > 0$. □

3.2. **Depth-first exploration within the Stern-Brocot tree.** In this section, we interpret the MA-LBR operator $\overline{\mathcal{D}}_{\mathcal{V}}$ defined in Algorithm 2 as a depth-first transversal of a subtree of the Stern-Brocot tree. The concept of depth-first exploration is introduced in Algorithm 3.

Algorithm 3 Depth-first exploration of a finite ordered tree T, with root e_*

```
Initialize a mutable list $L \leftarrow [e_*]$.
While L is non-empty do
    Remove from $L$ its first element $e$.
    Denote by $e_1, \cdots, e_n$ the children of $e$ in the tree $T$.
    Prepend $[e_1, \cdots, e_n]$ to $L$.
```

We introduce in Algorithm 4 a simplified version of the adaptive MA-LBR operator $\overline{\mathcal{D}}_{\mathcal{V}}$. It incorporates a dummy variable L used to emulate Algorithm 3; see Proposition 3.12 below. G_i denotes the i-th element of the mutable list G.

Algorithm 4 Minimization on a subtree of the Stern-Brocot tree.

```
Input: a finite set $V \subseteq \mathbb{Z}^2$ and a map $\varphi : V \to \mathbb{R}$.
Initialize a mutable vertex $f \leftarrow (1,0)$, and a mutable list $G \leftarrow [(0,1)]$.
Initialize also $\mathbb{D} \leftarrow +\infty$.
While $G$ is non-empty do
    Denote by $g := G_1$ the first element of $G$, and set $e := f + g$.
    Denote $n := \text{length}(G)$.
    Introduce the list $L := [f + g,\ G_1 + G_2, \cdots, G_{n-1} + G_n]$.
    If $e \in V$
        then prepend $e$ to $G$, and set $\mathbb{D} \leftarrow \min\{\mathbb{D}, \varphi(e)\}$
        else remove $g$ from $G$ and set $f \leftarrow g$
Return $\mathbb{D}$.
```

Lemma 3.10. *At each iteration of the* While *loop in Algorithm* 4*, one actually has* $L = [f \oplus g, G_1 \oplus G_2, \cdots, G_{n-1} \oplus G_n]$.

Proof. In the first iteration $L = [(1,0) \oplus (0,1)]$. We proceed by induction on the iteration index. Assume that $L = [f \oplus g, G_1 \oplus G_2, \cdots, G_{n-1} \oplus G_n]$. Since $e = f \oplus g$, we have $f + e = f \oplus e$ and $e + g = e \oplus g$ by Proposition 1.15. If $e \in V$, then at the next iteration $L' = [f \oplus e, e \oplus g, G_1 \oplus G_2, \cdots, G_{n-1} \oplus G_n]$. On the other hand, if $e \notin V$, then at the next iteration $L' = [G_1 \oplus G_2, \cdots, G_{n-1} \oplus G_n]$. □

In the following any set $V \subseteq \mathbb{Z}^2$ is *regarded as a graph*, whose edges are those of $\mathbb{T}$ having their endpoints in V; see Definition 1.16. In particular the standard Stern-Brocot tree has vertices $\mathbb{T}^+ := \{(a,b) \in \mathbb{T};\ a > 0, b > 0\}$ and is a complete infinite binary tree of root $(1,1)$. We say that a binary tree is *proper* iff its nodes have either two children (internal nodes) or zero (leaves).

Definition 3.11. Let $V \subseteq \mathbb{Z}^2$ be a finite subtree of $\mathbb{T}^+$ with root $(1,1)$. We denote by $V_* \subseteq \mathbb{Z}^2$ the proper binary subtree of $\mathbb{T}^+$ whose set of internal nodes is V. Note that $\#(V_*) = 2\#(V) + 1$.

Proposition 3.12. *Let $V \subseteq \mathbb{Z}^2$ be a finite subtree of $\mathbb{T}^+$ with root $(1,1)$. Algorithm 4 conducts a depth-first transversal of the tree V_*, and at termination $\mathbb{D} = \min\{\varphi(e);\, e \in V\}$.*

Proof. Let $e = f \oplus g \in V_*$. If $e \in V$, then e has two children in V_*, namely $f \oplus e$ and $e \oplus g$. If $e \in V_* \setminus V$, then e is a leaf of V_*. Inspection of the proof of Lemma 3.10 shows that the list L is updated precisely as expected for a depth-first transversal of V_*; see Algorithm 3. Since the operation $\mathbb{D} \leftarrow \min\{\mathbb{D}, \varphi(e)\}$ is performed only for elements of V, it evaluates the minimum of φ on V. □

We finally introduce a slight generalization of Proposition 3.12 so as to more closely fit the context of Algorithm 2, defining $\overline{\mathcal{D}}_{\mathcal{V}} u(x)$.

Corollary 3.13. *Consider a finite set $V_0 \subseteq \mathbb{Z}^2$ and a map $\varphi : V_0 \to \mathbb{R}$. Applying Algorithm 4 to (V_0, φ) yields at termination $\mathbb{D} = \min\{\varphi(e);\, e \in V\}$, where $V \subseteq \mathbb{Z}^2$ denotes the connected component of $(1,1)$ in V_0.*

Corollary 3.14. *Let $V_0 \subseteq \mathbb{Z}^2$ be a stencil, and let $\varphi : V_0 \to \mathbb{R}$ be an even function. Let $V \subseteq \mathbb{Z}^2$ denote the connected components of $\pm(1,1)$ and $\pm(1,-1)$ in V_0. Apply Algorithm 4 to (V_0, φ) with the modified initialization $G \leftarrow [(0,1), (-1,0)]$. Then at termination $\mathbb{D} = \min\{\varphi(e);\, e \in V\}$.*

Proof. The execution of Algorithm 4 with the modified initialization can be decomposed into two parts. (I) Execution with the standard initialization, which by Corollary 3.13 computes $\mathbb{D}_+ := \min\{\varphi(e);\, e \in V \cap \mathbb{T}^+\}$. (II) Execution with the modified initialization $f \leftarrow (0,1)$, $G \leftarrow [(-1,0)]$, which similarly computes $\mathbb{D}_- := \min\{\varphi(e);\, e \in V \cap \mathbb{T}^-\}$, with $\mathbb{T}^- := \{(a,b) \in \mathbb{T};\, a < 0,\, b > 0\}$. Eventually $\mathbb{D} = \min\{\mathbb{D}_+, \mathbb{D}_-\}$, which is the minimum of φ on V since φ is even and V is symmetric w.r.t. the origin by Definition 1.2. □

3.3. **Increasing functions on trees, and the case of quadratic functions.** The hierarchical MA-LBR operator $\overline{\mathcal{D}}_{\mathcal{V}} u(x)$ can be regarded, essentially (see Corollary 3.14) as a minimization over a subtree of the Stern-Brocot tree. In this section, we identify assumptions under which this pruning procedure is valid; i.e. it only drops useless branches where the minimum would not be found.

Definition 3.15. Let B be a graph, let A be a subset of its vertices, and let $\varphi : B \to \mathbb{R}$. We say that φ is increasing outside of A iff for each edge $a \to b$ of the graph B with $b \in B \setminus A$, one has $\varphi(a) \leq \varphi(b)$.

Proposition 3.16. *Let B be a finite collection of finite trees, and let A be a subset of B containing the root of each tree. If $\varphi : B \to \mathbb{R}$ is increasing outside of A, then $\min_A \varphi = \min_B \varphi$.*

Proof. Let b be a minimizer of φ on B, with minimal (graph) distance from the root of its tree. Assume for contradiction that $b \notin A$. Then b is not the root; hence there exists an edge $a \to b$ in the graph B. Then $\varphi(a) \leq \varphi(b)$ and a is closer to the root, which is a contradiction. □

Given some fixed $u \in \mathbb{U}$, $x \in X$, we introduce the function $\varphi : \mathbb{T} \to \mathbb{R}_+$ defined by

(3.2)
$$\varphi(e) := \begin{cases} h(\Delta_e^+ u(x), \Delta_f^+ u(x), \Delta_g^+ u(x)) & \text{if } e = f \oplus g, \\ +\infty & \text{otherwise, i.e. if } e \in \{(\pm 1, 0), (0, \pm 1)\}. \end{cases}$$

We show in the next proposition, under some assumptions, that φ is increasing in the sense of Definition 3.15 on some subsets of the graph $\mathbb{T}$. From this we deduce the equality of the (non-adaptive) MA-LBR operator $\mathcal{D}_{\mathcal{V}}$ (1.16) associated to some small and large stencils.

Definition 3.17. Let $u \in \mathbb{U}$, let $x \in X$, and let $e = f \oplus g$. We define

$$H_e u(x) := \Delta_e u(x) - \Delta_f u(x) - \Delta_g u(x). \tag{3.3}$$

Proposition 3.18. *Let $u \in \mathbb{U}$, let $x \in X$, and let $\mathcal{U}, \mathcal{V}$ be families of stencils. Assume that $\mathcal{V}(x) \subseteq \mathcal{U}(x) \subseteq \overline{\mathcal{V}}(x)$ and that:*

(A) $\Delta_e u(x) > 0$ for each $e \in \overline{\mathcal{V}}(x)$.

(B) $H_e u(x) \geq 0$ for each $e \in \overline{\mathcal{V}}(x) \setminus \mathcal{U}(x)$.

Then $\mathcal{D}_{\mathcal{U}} u(x) = \mathcal{D}_{\overline{\mathcal{V}}} u(x)$.

Proof. Fix $u \in \mathbb{U}$, $x \in X$, and consider φ defined by (3.2). Denote $V := \mathcal{V}(x)$, $U := \mathcal{U}(x)$, $\overline{V} := \overline{\mathcal{V}}(x)$. We regard $\overline{V}$ as a subgraph of $\mathbb{T}$ by keeping all edges with endpoints in this set.

We claim that the restriction of φ to $\overline{V}$ is increasing outside of U, in the sense of Definition 3.15. Indeed consider an edge $e \to e'$ of $\mathbb{T}$, where $e' \in \overline{V} \setminus U$. Introducing the decomposition $e = f \oplus g$, we note that $e' \in \{f \oplus e, e \oplus g\}$, and also that $e, f, g \in \overline{\mathcal{V}}(x)$ by Lemma 3.8. For each $\omega \in \{e, f, g, e'\}$ let $\delta_\omega := \Delta_\omega u(x)$, which is positive by (A). Assuming without loss of generality that $e' = f \oplus e$ we obtain $\delta_{e'} \geq \delta_e + \delta_f$, by (B). Hence, as announced, using Lemma 3.6,

$$\varphi(e) = h(\delta_e, \delta_f, \delta_g) \leq \delta_e \delta_f = h(\delta_{e'}, \delta_e, \delta_f) = \varphi(e'). \tag{3.4}$$

Thus $\varphi : \overline{V} \to \mathbb{R}$ is increasing outside of U, and therefore $\mathcal{D}_{\mathcal{U}} u(x) = \min\{\varphi(e);\ e \in U\} = \min\{\varphi(e);\ e \in \overline{V}\} = \mathcal{D}_{\overline{\mathcal{V}}} u(x)$ by Proposition 3.16 and Corollary 3.9. This concludes the proof. □

We focus in the rest of this section on the case of a quadratic function u_M, where $M \in S_2^+$ is fixed. We link the adaptive MA-LBR operator $\overline{\mathcal{D}}_{\mathcal{V}} u_M(x)$ with Selling's algorithm page (Algorithm 1). Since $\Delta_e u_M(x) = \langle e, Me \rangle > 0$ for any $x \in X$, $e \neq 0$, assumption (A) of Proposition 3.18 is automatically satisfied. Regarding (B) we observe the simplification: if $e = f \oplus g$,

$$H_e u_M(x) = \langle f, Mg \rangle. \tag{3.5}$$

We thus introduce

$$\mathbb{T}_M := \{e \in \mathbb{T};\ e = f \oplus g,\ \langle f, Mg \rangle < 0\}. \tag{3.6}$$

We shall use the identity: for any $f, g \in \mathbb{R}^2$,

$$\det(f, g)^2 + \langle f, g \rangle^2 = \|f\|^2 \|g\|^2. \tag{3.7}$$

Lemma 3.19. *If M is diagonal, then $\mathbb{T}_M = \emptyset$. Otherwise $\mathbb{T}_M = \{e_0, \cdots, e_n, -e_0, \cdots, -e_n\}$, for some finite branch $e_0 \to e_1 \to \cdots \to e_n$ of $\mathbb{T}$, with $e_0 \in \{(1, 1), (-1, 1)\}$.*

Proof. Claim (symmetry): One has $e \in \mathbb{T}_M$ iff $-e \in \mathbb{T}_M$. Indeed if $e = f \oplus g$, then $-e = (-f) \oplus (-g)$, and $\langle f, Mg\rangle = \langle (-f), M(-g)\rangle$. Claim (tree structure): For any edge $e \to e'$ of $\mathbb{T}$, one has $e' \in \mathbb{T}_M \Rightarrow e \in \mathbb{T}_M$. Indeed write $e = f \oplus g$, so that $e' = f \oplus e$ (resp. or $e' = e \oplus g$). Then $\langle f, Me\rangle = \langle f, Mg\rangle + \langle f, Mf\rangle \geq \langle f, Mg\rangle$ (resp. likewise $\langle e, Mg\rangle \geq \langle f, Mg\rangle$) as announced. Claim (single branch): If $e = f \oplus g$, then at most one of $f \oplus e$ and $e \oplus g$ belongs to $\mathbb{T}_M$. Indeed $\langle f, Me\rangle + \langle e, Mg\rangle = \langle e, Me\rangle \geq 0$, hence at most one of these scalar products is negative.

In order to conclude the proof, it suffices to establish the finiteness of $\mathbb{T}_M$. Let $e = f \oplus g$, let λ denote the smallest eigenvalue of M, and let $\kappa(M) := \sqrt{\|M\|\|M^{-1}\|}$. Finiteness follows from the claim: if $\|e\| \geq 1 + \kappa(M)$, then $e \notin \mathbb{T}_M$. Indeed

$$\begin{aligned}\langle f, Mg\rangle^2 &= \|f\|_M^2 \|g\|_M^2 - \det(M)\det(f,g)^2\\ &\geq \lambda^2 \|f\|^2\|g\|^2 - \det(M)\\ &> \lambda^2\left[(\|e\|-1)^2 - \kappa(M)^2\right].\end{aligned}$$

We applied (3.7) to $M^{\frac{1}{2}}f$ and $M^{\frac{1}{2}}g$ for the first identity, and used that $\det(f,g) = 1$ for the following inequality. The last inequality used $\|f\|+\|g\| > \|e\|$, $\min\{\|f\|, \|g\|\} \geq 1$, hence $\|f\|\|g\| > \|e\| - 1$, and $\det(M)/\lambda^2 = \kappa(M)^2$. Since $\|e\| \geq 1 + \kappa(M)$ we have shown $\langle f, Mg\rangle \neq 0$. Apply this observation to the family of matrices $M_t := (1-t)\,\mathrm{Id} + tM$, $t \in [0,1]$, which satisfy $\kappa(M_t) \leq \kappa(M)$. Proposition 1.13 states that $\langle f, M_0 g\rangle = \langle f, g\rangle \geq 0$, hence $\langle f, M_t g\rangle > 0$ for all $t \in [0,1]$, thus $\langle f, Mg\rangle > 0$ and therefore $e \notin \mathbb{T}_M$ as announced. □

In the following corollary, a superbase (e_0, e_1, e_2) of $\mathbb{Z}^2$ is said to be equivalent to the superbases $(\varepsilon e_i, \varepsilon e_j, \varepsilon e_k)$, for any permutation $\{i, j, k\}$ of $\{0, 1, 2\}$ and any sign $\varepsilon \in \{-1, 1\}$.

Corollary 3.20. *Let $M \in S_2^+$. If M is diagonal, then Selling's algorithm stops at the first iteration. Otherwise let $e_0 \to \cdots \to e_n$ be as in Lemma* 3.19, *write $e_i = f_i \oplus g_i$. Then Selling's algorithm, initialized with the superbase $(e_0, -f_0, -g_0)$, generates in its successive iterations superbases equivalent to $(e_i, -f_i, -g_i)$. It terminates at the n-th iteration.*

Proof. Claim: The superbase $(e_n, -f_n, -g_n)$ is M-obtuse. Indeed, $\langle (-f_n), M(-g_n)\rangle = \langle f_n, Mg_n\rangle < 0$ since $e_n \in \mathbb{T}_n$. On the other hand, $f_n \oplus e_n \notin \mathbb{T}_M$ and $e_n \oplus g_n \notin \mathbb{T}_M$, by Lemma 3.19 and the structure of $\mathbb{T}$; see Definition 1.16. Hence $\langle (-f_n), Me_n\rangle \leq 0$ and $\langle -g_n, Me_n\rangle \leq 0$.

Proof by induction on the iteration count i, $0 \leq i \leq n$. Case $i = 0$ holds by the choice of initialization. Induction: Consider the superbase $(e_i, -f_i, -g_i)$ of the i-th iteration, for some $0 \leq i < n$. Assume that $e_{i+1} = f_i \oplus e_i$ (the case $e_{i+1} = e_i \oplus g_i$ is similar), which means that $\langle f_i, Me_i\rangle < 0$. One has $\langle (-f_i), M(-g_i)\rangle < 0$ since $e_i \in \mathbb{T}_M$, $\langle (-f_i), Me_i\rangle > 0$, and $\langle (-g_i), Me_i\rangle = -\langle e_i, Me_i\rangle + \langle f_i, Me_i\rangle < 0$. Hence Selling's algorithm constructs for the next iteration the superbase $(e_i - (-f_i), -f_i, -e_i) = (e_i \oplus f_i, -f_i, -e_i) = (e_{i+1}, -f_{i+1}, -g_{i+1})$ as announced. □

The next proposition establishes our main result, Theorem 1.21, in the special case of quadratic functions. It also shows that the pruning procedure defining the adaptive operator $\overline{\mathcal{D}}_{\mathcal{V}}$ is extremely well behaved, since it only explores (in addition to the basic stencil $\mathcal{V}(x)$) a single branch of the Stern-Brocot tree, within $\mathbb{T}_M$ as in Selling's algorithm.

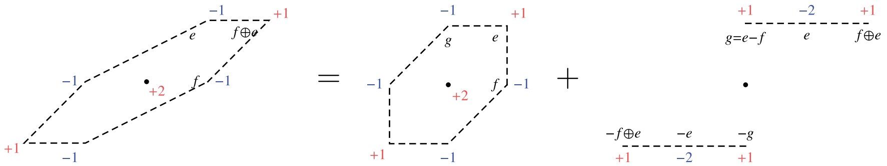

Figure 6. Illustration of Lemma 3.22

Proposition 3.21. *Let $M \in S_2^+$, and let $u := u_M$. Let $x \in X$, let $\mathcal{V}$ be a family of stencils, and let*

$$\mathcal{U}(x) := \mathcal{V}(x) \cup \{e \in \mathcal{V}_\Omega(x);\ H_e u(x) < 0\} \\ = \mathcal{V}(x) \cup (\mathbb{T}_M \cap \mathcal{V}_\Omega(x)). \tag{3.8}$$

Then $\overline{\mathcal{D}}_\mathcal{V} u(x) = \mathcal{D}_\mathcal{U} u(x) = \mathcal{D}_{\overline{\mathcal{V}}} u(x)$. In addition, when computing $\overline{\mathcal{D}}_\mathcal{V} u(x)$ through Algorithm 2, the evaluation of (3.2) is performed only when $e \in \mathcal{U}(x)$.

Proof. Let $U := \mathcal{U}(x)$, $V := \mathcal{V}(x)$, $\overline{V} := \overline{\mathcal{V}}(x)$, so that $U = V \cup (\mathbb{T}_M \cap \overline{V})$ by (3.5). Note that the collection of subtrees of the same tree, with the same root, is stable by unions and intersections. Hence U satisfies property (Hierarchy) of stencils, by Lemma 3.19 and Proposition 1.20.

We recognize in (3.8) the (Refinement test) appearing in Algorithm 2. Corollary 3.14 applied to $\mathcal{U}(x)$ and (3.2) states that $\overline{\mathcal{D}}_\mathcal{V} u(x) = \mathcal{D}_\mathcal{U} u(x)$. On the other hand, $\mathcal{D}_\mathcal{U} u(x) = \mathcal{D}_{\overline{\mathcal{V}}}(x)$ by Proposition 3.18, which concludes the proof. □

3.4. **Equality of the adaptive and the extensive MA-LBR operator.** We prove Theorem 1.21, stating under mild assumptions the equality of the adaptive MA-LBR operator $\overline{\mathcal{D}}_\mathcal{V}$ and the brute-force one $\mathcal{D}_{\overline{\mathcal{V}}}$, which extensively sweeps through the extended stencils. For that purpose, and similarly to the quadratic case, we use through Proposition 3.18 the fact that the minimized function (3.2) is increasing on some portion of the Stern-Brocot tree.

The key to the proof is the next proposition, preceded by a technical lemma, which weakens the assumptions of Proposition 3.18. Strikingly, the stencils at each $x \in X$ cannot be dealt with independently. A simultaneous and global argument is used instead, inspired by [18].

Lemma 3.22. *Let $u \in \mathbb{U}$, $x \in X$, and $e = f \oplus g$. If $e, e+f \in \mathcal{V}_\Omega(x)$, then $(x+e) \pm f \in \Omega$, $(x-e) \pm f \in \Omega$, and*

$$H_{e+f} u(x) = H_e u(x) + \Delta_f u(x+e) + \Delta_f u(x-e). \tag{3.9}$$

Likewise if $e, e+g \in \mathcal{V}_\Omega(x)$, exchanging the roles of f and g.

Proof. Since $e \in \mathcal{V}_\Omega(x)$ we have $x \pm e, x \pm f, x \pm g \in \Omega$. Since in addition $e+f \in \mathcal{V}_\Omega(x)$ we have $x \pm (e+f) \in \Omega$. Note that $(x+e) - f = x+g$ and $(x-e)+f = x-g$. Expanding the expressions on both sides of (3.9), using that $e+f = f \oplus e$ for the left side, we find that they only involve the values of $u \in \mathbb{U}$ at points of $X = \Omega \cap \mathbb{Z}^2$, and not on the boundary $\partial\Omega$ (as in (1.2) and not (1.3)). A cancellation occurs, as illustrated on Figure 6, and the result is proved. □

Proposition 3.23. *Let $\mathcal{U},\mathcal{V}$ be families of stencils, and let $u \in \mathbb{U}$. For each $x \in X$, assume that $\mathcal{V}(x) \subseteq \mathcal{U}(x) \subseteq \overline{\mathcal{V}}(x)$ and that:*

(a) $\Delta_e u(x) > 0$ for each $e \in \mathcal{U}(x)$.

(b) $H_e u(x) \geq 0$ for each $e \in \overline{\mathcal{V}}(x) \setminus \mathcal{U}(x)$ for which there exists $e' \in \mathcal{U}(x)$ such that $e' \to e$.

Then $u, \mathcal{U}, \mathcal{V}$ satisfy the assumptions of Proposition 3.18*, for each $x \in X$.*

Proof. Fix the stencils $\mathcal{V}$, and proceed by decreasing induction on the cardinality $\#(\mathcal{U}) := \sum_{x\in X} \#(\mathcal{U}(x))$. If $\#(\mathcal{U}) = \#(\overline{\mathcal{V}})$, then $\mathcal{U} = \overline{\mathcal{V}}$ and there is nothing to prove.

Assume that $\#(\mathcal{U}) < \#(\mathcal{V})$, and consider a point $x \in X$ and a vector $e \in \overline{\mathcal{V}}(x) \setminus \mathcal{U}(x)$, such that $\|e\|$ is minimal. Let us introduce the sets $\mathcal{U}'(x) := \mathcal{U}(x) \cup \{\pm e\}$, and $\mathcal{U}'(y) := \mathcal{U}(y)$ for all $y \neq x$, and note that $\#(\mathcal{U}') = \#(\mathcal{U}) + 2$. We prove in the following that $\mathcal{U}'$ is a family of stencils satisfying the assumptions (a) and (b). Hence by induction $\mathcal{U}'$ satisfies (a) and (b), which immediately implies the same properties for $\mathcal{U}$ and concludes the proof.

Proof that $\mathcal{U}'$ is a family of stencils. Only (Hierarchy) needs to be checked. Since $e \in \overline{\mathcal{V}}(x) \setminus \mathcal{U}(x)$ we have $e \in \mathbb{T} \setminus V_8$; hence we may introduce the decomposition $e = f \oplus g$. By Lemma 3.1, either $f \to e$ or $g \to e$ is an edge of the graph $\mathbb{T}$. By Lemma 3.8 we have $f, g \in \overline{\mathcal{V}}(x)$. By minimality of $\|e\|$ we have $f, g \in \mathcal{U}(x)$. By (Hierarchy) for $\mathcal{U}$ the set V_8 has an element in the connected component of f and g in $\mathcal{U}(x)$, hence in the connected component of e in $\mathcal{U}'(x)$. This establishes (Hierarchy) for $\mathcal{U}'$.

Proof that $\mathcal{U}'$ satisfies (a). It suffices to check this property for the additional elements $\pm e$. Using (a) for $\mathcal{U}$ we obtain $\Delta_f u(x) > 0$, $\Delta_g u(x) > 0$. Using (b) for $\mathcal{U}$ we get $H_e u(x) \geq 0$. Therefore

$$\Delta_e u(x) = H_e u(x) + \Delta_f u(x) + \Delta_g u(x) > 0. \tag{3.10}$$

Proof that $\mathcal{U}'$ satisfies (b). The two edges originating from e in the graph $\mathbb{T}$ are $e \to f \oplus e$ and $e \to e \oplus g$; see Definition 1.16. Let us assume that $e+f \in \overline{\mathcal{V}}(x) \setminus \mathcal{U}'(x)$ and establish that $H_{e+f} u(x) \geq 0$. Note that $\overline{\mathcal{V}}(x) \setminus \mathcal{U}(x) \subseteq (\mathcal{V}(x) \cup \mathcal{V}_\Omega(x)) \setminus \mathcal{V}(x) \subseteq \mathcal{V}_\Omega(x)$, hence $e, e+f \in \mathcal{V}_\Omega(x)$. Applying Lemma 3.2 we obtain $(x+e) \pm f \in \Omega$, hence $f \in \overline{\mathcal{V}}(x+e)$ by Lemma 3.2, thus $f \in \mathcal{U}(x+e)$ by minimality of $\|e\|$, and therefore $\Delta_f u(x+e) > 0$ by (a) for the stencils $\mathcal{U}$. Likewise $\Delta_f u(x-e) > 0$. Using (3.9) yields as announced $H_{e+f} u(x) > 0$. Likewise $H_{e+g} u(x) > 0$ if $e+g \in \overline{\mathcal{V}}(x) \setminus \mathcal{U}'(x)$. This establishes (b) for $\mathcal{U}'$ and concludes the proof. □

Our last proposition immediately implies the announced Theorem 1.21.

Proposition 3.24. *Let $\mathcal{V}$ be a family of stencils, and let $u \in \mathbb{U}$. If $\overline{\mathcal{D}}_{\mathcal{V}} u > 0$ on X, then $\overline{\mathcal{D}}_{\mathcal{V}} u = \mathcal{D}_{\overline{\mathcal{V}}} u$ on X. In all cases $\mathcal{D}_{\overline{\mathcal{V}}} u \leq \overline{\mathcal{D}}_{\mathcal{V}} u$ on X.*

Proof. Let $\mathcal{V}$ be a family of stencils, and let $u \in \mathbb{U}$. We introduce, for each $x \in X$, the set

$$\mathcal{U}_0(x) := \mathcal{V}(x) \cup \{e \in \mathcal{V}_\Omega(x);\ H_e u(x) < 0\}. \tag{3.11}$$

By construction $\mathcal{V}(x) \subseteq \mathcal{U}_0(x) \subseteq \overline{\mathcal{V}}(x)$. We regard $\mathcal{U}_0(x)$ as a subgraph of $\mathbb{T}$, keeping all edges whose endpoints are both in $\mathcal{U}_0(x)$. Denote by $\mathcal{U}(x)$ the union of connected components intersecting V_8 in $\mathcal{U}_0(x)$. By construction, $\mathcal{V}(x) \subseteq \mathcal{U}(x) \subseteq \overline{\mathcal{V}}(x)$, and $\mathcal{U} = (\mathcal{U}(x))_{x\in X}$ is a family of stencils.

We recognize in the definition (3.11) of $\mathcal{U}_0(x)$ the (Refinement test) appearing in the computation of $\overline{\mathcal{D}}_{\mathcal{V}}(x)$, Algorithm 2. Corollary 3.14 applied to $\mathcal{U}_0(x)$ and the map (3.2) thus states that: for any $x \in X$,

$$\overline{\mathcal{D}}_{\mathcal{V}} u(x) = \mathcal{D}_{\mathcal{U}} u(x). \tag{3.12}$$

Recalling that $\mathcal{U}(x) \subseteq \overline{\mathcal{V}}(x)$, for all $x \in X$, we obtain $\mathcal{D}_{\overline{\mathcal{V}}} u \le \mathcal{D}_{\mathcal{U}} u = \overline{\mathcal{D}}_{\mathcal{V}} u$ on X as announced.

The stencils $\mathcal{U}$ satisfy by construction assumption (b) of Proposition 3.23. Introducing the assumption that $\overline{\mathcal{D}}_{\mathcal{V}} = \mathcal{D}_{\mathcal{U}}$ is positive on X, and using Proposition 1.22, we find that $\mathcal{U}$ also satisfies assumption (a) of Proposition 3.23. Thus $\mathcal{D}_{\mathcal{U}} u = \mathcal{D}_{\overline{\mathcal{V}}} u$ on X, by Proposition 3.18, which concludes the proof. □

4. Numerical experiments

We compare the introduced MA-LBR (Monge-Ampère using Lattice Basis Reduction), with two alternative solvers of Monge-Ampère equations. The Finite Differences scheme $\mathcal{D}^{\rm FD}$ (see (1.4) and [15]) is consistent but lacks the convergence guarantees associated to degenerate elliptic schemes. The Wide Stencil scheme $\mathcal{D}^{\rm WS}_{\mathcal{V}}$ (see (1.6) and [10]) provides these guarantees, but at the price of a difficult compromise between consistency error and scheme locality, governed by the chosen stencil angular resolution; see Figures 1 and 3. Our numerical scheme, the MA-LBR, aims to combine the qualities of these two methods: consistency and monotony, with a comparable numerical cost. We use the MA-LBR adaptive implementation $\overline{\mathcal{D}}_{\mathcal{V}}$ of Algorithm 2, with an 8 point stencil $\mathcal{V}(x) = V_8$, except on a layer of 4 pixels along the domain boundary (where hierarchical refinement is mostly ineffective), where we use the 48 points stencil of Figure 1 (right). (Recall that the second order differences $\Delta_e u(x)$, used to construct the discrete operators, are defined by (1.3) when $x+e, x-e \notin \Omega$, using the provided boundary data $u_{|\partial\Omega} = \sigma$.) The filtered scheme introduced in [10] also attempts to combine the strengths of the Wide Stencil scheme $\mathcal{D}^{\rm WS}_{\mathcal{V}}$ and the Finite Differences scheme $\mathcal{D}^{\rm FD}$; this scheme is omitted in our experiments because it depends on several parameters which make benchmarks and comparisons difficult.

We limit our attention to synthetic test cases, posed on the unit square $\Omega :=]0,1[^2$. A known convex function $U : \overline{\Omega} \to \mathbb{R}$ is numerically recovered from its hessian determinant $\rho := \det(\nabla^2 U)$ and its boundary values $\sigma := U_{|\partial\Omega}$. The tests are (supposedly) ordered by increasing difficulty, starting from a simple quadratic function and ending with a non-differentiable function (on a domain corner).

- (Quadratic) $U(x) := \frac{1}{2}\langle x, Mx\rangle$, where $M = M(\kappa,\theta)$ is as in (1.12) with $\kappa := 10$, $\theta := \pi/3$.
- (Smoothed cone) $U(x) := \sqrt{\delta^2 + \|x - x_0\|^2}$, with $\delta := 0.1$ and $x_0 := (1/2, 1/2)$.
- (Flat, [10]) $U(x) := (\|x - x_0\|_+ - r_0)^2 + \frac{\varepsilon}{2}\|x - x_0\|^2$, with $r_0 := 0.2$ and $\varepsilon = 10^{-6}$.
- (Singular, [10]) $U(x) := -\sqrt{2 - \|x\|^2}$.

An iterative solver is applied to the discrete system (1.5), starting from a strictly convex seed; see Remark 4.1. Although the convergence guarantees of DE schemes only encompass Euler iterative solvers, we used without trouble a damped[2] Newton

[2] Precisely, the iteration at a point $u \in \mathbb{U}$ takes the form $u' = u + \delta^k v$, where v is Newton's descent direction, $\delta := 0.7$, and $k \ge 0$ is the smallest integer such that: $\mathcal{D}(u + \delta^k v)$ is positive on

solver. This may come as a surprise to those who regard Newton methods as local and excessively sensitive to initialization. The Monge-Ampère PDE fortunately benefits from a more favorable situation, since a suitably damped Newton method has been shown [15] to converge globally: in the continuous setting, with periodic boundary conditions, and a Holder smooth positive right hand side. Discrete MA schemes which preserve the operator ellipticity may heuristically be expected to inherit this good behavior.

Note that Algorithm 2 for stencil refinement is applied independently at each step of the iterative numerical method, in contrast with e.g. hierarchical mesh refinement methods in finite element discretizations of PDEs. There is no attempt to reuse or refine stencils from earlier iterations of the Newton solver. Theorem 1.21 indeed guarantees that the operator $\overline{\mathcal{D}}_V$, associated to the refined stencils, is well defined and consistent with the costly $\mathcal{D}_{\overline{\mathcal{V}}}$, which is independent of the stencil refinement procedure. On this topic of adaptivity, let us mention that the evaluation of $\overline{\overline{\mathcal{D}}}_{\mathcal{V}}$ takes about 0.2 second in the smoothed cone test case, on a 100 × 100 discretization grid, whereas the non-adaptive variant with extended stencils $\mathcal{D}_{\overline{\mathcal{V}}}$ requires 55 seconds. This speed up by a factor $\approx$ 270 is remarkable in view of Theorem 1.21, which guarantees the exact identity $\overline{\mathcal{D}}_{\mathcal{V}} u = \mathcal{D}_{\overline{\mathcal{V}}} u$.

All four test functions are strictly convex, so that the corresponding density $\rho = \det(\nabla^2 U)$ is strictly positive as was assumed in (1.1). If in contrast ρ vanishes, then the Jacobian of the system (1.5) vanishes as well at the problem solution, which makes Newton's method non-applicable. Some tentative workarounds have been proposed, such as the variational characterization of degenerate equations in [11].

Quadratic test case. The MA-LBR recovers this solution exactly, up to floating point errors, thanks to the adaptivity of Algorithm 2, which refines the initial 8 point stencil until the vector $(2, 3)$ is included, and thus also the M-obtuse superbase $(2, 3), (-1, -1), (-1, -2)$. Scheme FD also recovers the exact solution for a range of resolutions, but afterwards the discrete iterative solver switches to some erroneous alternative solution; see Figure 7. Scheme WS produces a substantial L^∞ error, which does not decrease with the grid scale. Indeed, it reflects a consistency error, and not a discretization error. Scheme FD could presumably recover the exact solution at all resolutions if its iterative solver was initialized more sensibly, for instance using the output of Scheme WS or using a filtered combination of the two [10].

Smoothed cone test case. The recovered function is C^∞, yet its hessian is simultaneously (i) almost degenerate (rank one) close to the domain boundary, and (ii) strongly peaked in a small region around the center. Scheme FD entirely fails this test. Choosing the best stencil for scheme WS is non-trivial, since point (i) suggests using a large stencil for better angular resolution, but point (ii) mandates a scheme as local as possible. As a result, the best stencil, in terms of resulting L^∞ error, successively has 8, 16, 24 and 48 points for grid sizes $n \times n$ with n in the interval $[5, 30]$, $[30, 80]$, $[80, 160]$, $[160, \infty]$. The MA-LBR avoids the need for such manual parameter optimization and produces numerical errors often one order of magnitude smaller. It also needs the least damped Newton iterations to reach convergence.

X (except for scheme FD), and $\|f - \mathcal{D}(u + \delta^k v)\|_{L^\infty(X)}$ is a local minimum in k. Convergence is numerically observed but not claimed in general.

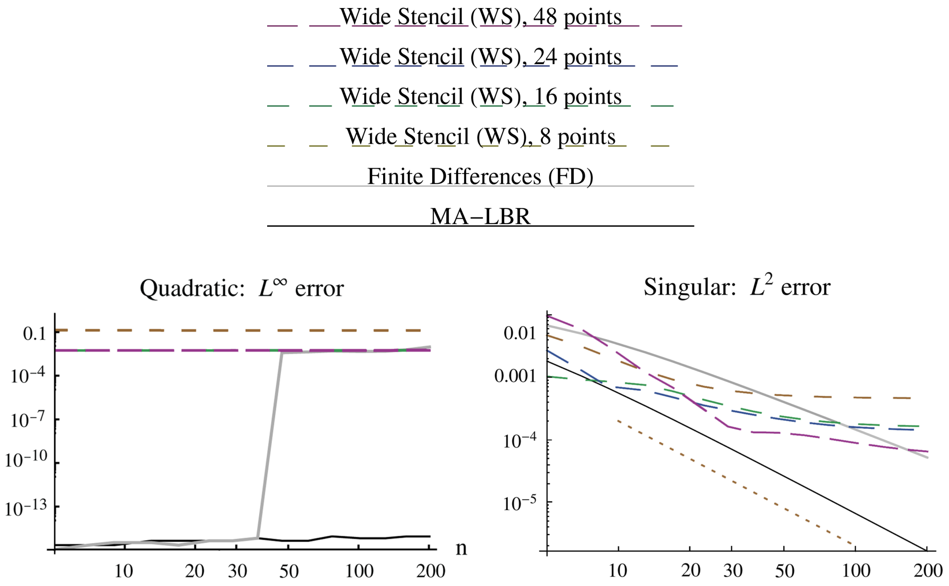

FIGURE 7. Numerical error in the Quadratic and Singular cases. Logarithmic scale on all axes. The slope of the dotted line represents second order convergence, i.e. error$(n) \approx \lambda n^{-2}$.

However, none of the studied methods can handle the limit case $U(x) = \|x - x_0\|$ of a pointed cone, corresponding to $\delta = 0$ instead of the previous small value 0.1. The Monge-Ampère operator must then be regarded as the non-negative measure $\det(\nabla^2 U) = \pi \delta_{x_0}$ in Pogorelov's sense, which violates both the positivity and the boundedness of the density ρ assumed in (1.1). The framework of viscosity solutions does not encompass these more general solutions, and the machinery of degenerate elliptic schemes fails; geometric approaches must be considered instead [22].

Flat test case. The recovered function is C^1, has a Lipschitz gradient, but is not C^2. It is also (almost) identically 0 on a disk, up to a quadratic perturbation introduced to help the Newton solver. The effect of this perturbation on the numerical solution is negligible in comparison with the discretization error. The best stencil for Scheme WS is the largest one, with 48 points, for all resolutions $n \times n$ with $n \geq 17$. Despite the lack of C^2 regularity, scheme FD performs well in this test, better in fact than WS. The MA-LBR again outperforms the tested alternatives and seems to provide a (slightly) improved asymptotic convergence rate in comparison with FD.

Singular test case. The recovered function is non-differentiable at the domain corner $(1, 1)$, where its gradient is formally $(+\infty, +\infty)$. Scheme FD fails this test, even if helped by initializing the iterative solver with a sampling of the known exact solution [9]. Regarding scheme WS, the L^∞ error curves and the number of Newton iterations exhibit a puzzling erratic behavior: despite the scheme degenerate ellipticity, nasty things seem to occur close to singular point $(1, 1)$. The L^2 error curve is smoother (see Figure 7) and suggests that the optimal stencil size is successively 16, 24, 48 at resolutions $n \times n$ with n in the respective intervals $[0, 8]$, $[8, 23]$ and $[23, \infty]$ (note that an even larger stencil would be preferable at resolutions ≥ 100).

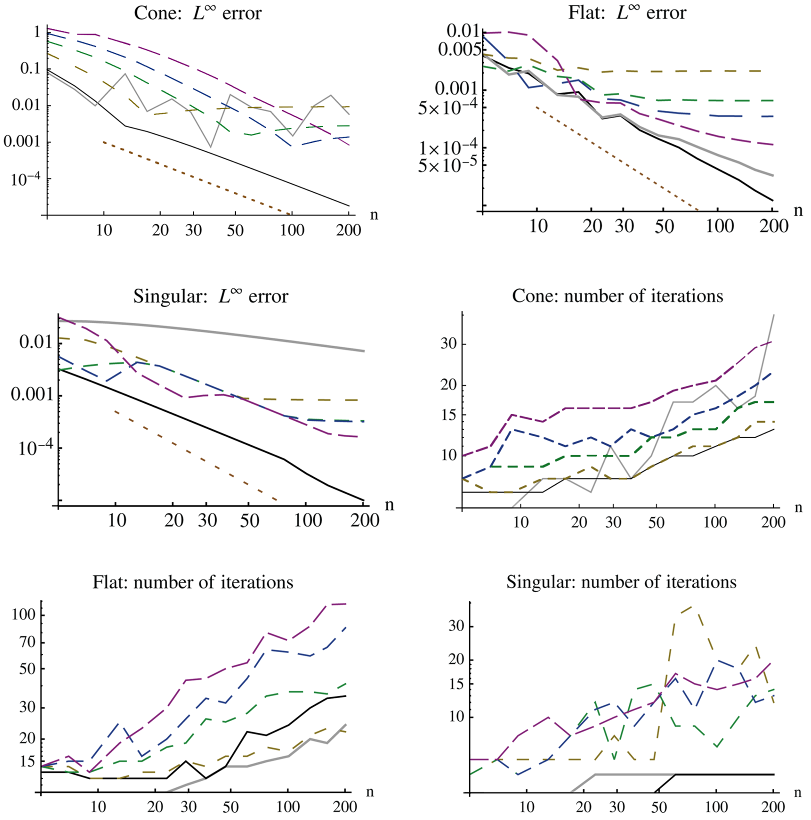

Figure 8. Numerical error and convergence speed in various test cases. Legend in Figure 7.

The MA-LBR avoids this difficult choice of stencil and improves numerical error often by an order of magnitude. Our discretization handles local singularities well and offers second order accuracy in smooth regions. The MA-LBR good balance is confirmed by the fast convergence of the damped Newton solver, which here never needs more than 5 iterations.

Remark 4.1 (Initialization). We initialize the damped Newton iterative solver with the restriction $u = V_{|X\cup\partial\Omega} \in \mathbb{U}$ of a strictly convex function $V \in C^0(\overline{\Omega})$, built using solely the prescribed boundary conditions σ on $\partial\Omega$. The construction is as follows: (i) Find $\varepsilon > 0$ such that $\sigma_\varepsilon(x) := \sigma(x) - \varepsilon\|x\|^2$ is convex on any segment of $\partial\Omega$. (ii) Find the maximal convex extension $\Sigma_\varepsilon : \overline{\Omega} \to X$ of σ_ε. This step requires the computation of a three dimensional convex hull, which is a classical problem of discrete geometry for which efficient procedures are available [4]. (iii) Initialize with the strictly convex $V(x) := \Sigma_\varepsilon(x) + \varepsilon\|x\|^2$.

5. Conclusion

The MA-LBR introduced in this paper is a new numerical scheme for two dimensional Monge-Ampère PDEs which combines consistency and degenerate ellipticity. In our numerical experiments, these properties become accuracy and robustness.

Our scheme is not strictly local and may involve long range stencils, but they are built in a sparse, adaptive, and anisotropic manner using a guaranteed and parameter free refinement algorithm. Our construction is also shown to be as local as it can be, among symmetric, consistent and degenerate elliptic schemes for the Monge-Ampère PDE. The analysis of our algorithm involves tools seldom used in the context of numerical analysis, including elements of lattice geometry [5] and the arithmetic of the Stern-Brocot tree.

Future research will be devoted to some natural questions that the present method cannot directly address. In particular (i) the computation of solutions of the weaker Alexandroff type, (ii) the additional difficulties tied to the discretization of optimal transport problems, instead of boundary value problems, (iii) Monge-Ampère problems posed on three dimensional domains, and (iv) local adaptation and refinement of the discretization grid.

Appendix A. Structure of the Stern-Brocot tree

A.1. Unique decomposition $e = f \oplus g$.

The two following propositions together establish Proposition 1.13.

Proposition A.1. *Let f, g be a direct acute basis of $\mathbb{Z}^2$. Then $e := f + g$ has co-prime coordinates, both non-zero.*

Proof. One has $\det(f, e) = \det(f, f+g) = 1$; hence the coordinates of e are co-prime as announced. Also $\|e\|^2 = \|f\|^2 + 2\langle f, g\rangle + \|g\|^2 \geq 2$, hence $\|e\| > 1$. Assuming for contradiction that a coordinate of e is zero, we find that the other one can only be ± 1, since they are co-prime. But then $\|e\| = 1$, which is a contradiction. This concludes the proof. □

In the following, a quadrant of the plane is a set of the form: for some $\alpha, \beta \in \{-1, 1\}$,

$$Q_{\alpha,\beta} := \{(a, b) \in \mathbb{R}^2;\ \alpha a \geq 0,\ \beta b \geq 0\}.$$

Proposition A.2. *Let $e = (a, b) \in \mathbb{Z}^2$ be such that* $\gcd(a, b) = 1$ *and $ab \neq 0$. Then there exists a unique direct basis (f, g) of $\mathbb{Z}^2$ such that $e = f + g$. Furthermore f and g belong to the same (closed) quadrant of the plane as e.*

Proof. Let R be the rotation of $\pi/2$. The image (Rf, Rg) of a direct acute basis of $\mathbb{Z}^2$ still is one. Also, R cyclically permutes the four quadrants of the plane. Without loss of generality, we may thus assume that a and b are positive.

Existence. Consider a Bezout relation: $u, v \in \mathbb{Z}^2$ such that $av - bu = 1$. For any $k \in \mathbb{Z}$, one also has the relation $a(v + kb) - b(u + ka) = 1$. By euclidean division and up to such a transformation, we may therefore assume that $0 \leq u < a$. Then $av = 1 + bu \leq 1 + b(a-1) \leq ab$, thus $0 < v \leq b$. The vectors $f := (a-u, b-v)$ and $g := (u, v)$ have non-negative entries. Hence they belong to the same quadrant as e and satisfy $\langle f, g\rangle \geq 0$. Also $\det(f, g) = (a-u)v - (b-v)u = av - bu = 1$. This concludes the proof of existence.

Uniqueness. Let (f', g') be another direct acute basis such that $e = f' + g'$. We introduce the coordinates (u', v') of g' and observe that $f' = (a - u', b - v')$. Then $\det(f', g') = (a - u')v' - (b - v')u' = av' - bu'$. We recognize another Bezout relation between the co-prime integers a, b. Hence $u' = u + ka$ and $v' = v + kb$ for some $k \in \mathbb{Z}$. Recall that $0 \leq u < a$ and $0 < v \leq b$. If $k < 0$, then the coordinates of $g' = (u', v') = g + ke$ satisfy $u \leq u - a' < 0$, $v' \leq v - b \leq 0$, while both coordinates of $f' = f - ke$ are positive; this contradicts the assumption $\langle f', g'\rangle \geq 0$. The case $k > 0$ is excluded by a similar argument, exchanging the roles of f' and g'. Hence $k = 0$, which concludes the proof of uniqueness. □

A.2. **Connected components of the graph $\mathbb{T}$.** We identify the structure of the graph $\mathbb{T}$, as announced in Proposition 1.17.

Lemma A.3. *All edges of $\mathbb{T}$ have both their endpoints in the interior of the same quadrant.*

Proof. Any edge of $\mathbb{T}$ has the form $e \to f \oplus e$ or $e \to e \oplus g$, where $e = f \oplus g$. By Proposition A.2, f, g belong to the same quadrant as e. Since both coordinates of e are non-zero, it belongs to the interior of its quadrant. Since this quadrant is a convex cone, the edge joins, as announced, two points of its interior. □

Lemma A.4. *Let $e = f \oplus g$. If $\|f\| > \|g\|$, then $f = (f - g) \oplus g$. If $\|g\| > \|f\|$, then $g = f \oplus (g - f)$. If $\|f\| = \|g\|$, then $\|e\|^2 = 2$.*

Proof. Since (f, g) is a direct basis, one has $\det(f, g) = 1$. Hence $\langle f, g\rangle^2 + 1 = \|f\|^2\|g\|^2$ by (3.7).

If $\|f\| > \|g\|$, then $\langle f, g\rangle^2 + 1 > \|g\|^2\|g\|^2$, thus $\langle f, g\rangle \geq \|g\|^2$, and therefore $\langle f - g, g\rangle \geq 0$. Remarking in addition that $\det(f - g, f) = \det(f, g) = 1$, we obtain, as announced, that $f = (f - g) \oplus g$. The case $\|g\| > \|f\|$ is similar.

If $\|f\| = \|g\|$, then $\langle f, g\rangle^2$ and $\|f\|^2\|g\|^2$ are consecutive perfect squares, hence equal to 0 and 1. Thus $\langle f, g\rangle = 0$, $\|f\|^2 = \|g\|^2 = 1$, and therefore $\|e\|^2 = \|f+g\|^2 = 2$, as announced. □

Lemma A.5. *Let $e = f \oplus g$. If $\|e\|^2 = 2$, then no edge of $\mathbb{T}$ arrives at e. If $\|e\|^2 > 2$, then exactly one edge of $\mathbb{T}$ arrives at e, and it must be either $f \to e$ or $g \to e$.*

Proof. Edges of $\mathbb{T}$ have the form $e' \to f' \oplus e'$ (resp. $e' \to e' \oplus g'$) where $e' = f' \oplus g'$. If such an edge arrives at e, then by uniqueness of the decomposition, one must have $f' = f$ and $e' = g$, thus $g' = g - f$ (resp. $e' = f$ and $g' = g$, thus $f' = f - g$). This corresponds to the two announced cases $f \to e$ or $g \to e$.

If the first case is realized, then $\langle f, g - f\rangle = \langle f', g'\rangle \geq 0$ (resp. second case, $\langle f - g, g\rangle \geq 0$). Assuming for contradiction that the two cases are realized, we obtain by addition $-\|f - g\|^2 \geq 0$, and therefore $f = g$. This contradicts the assumption that (f, g) is a basis of $\mathbb{Z}^2$. □

Let us summarize the properties of the graph $\mathbb{T}$. By Lemma A.3 all edges of $\mathbb{T}$ have their endpoints within the interior of the same quadrant. Also, for any $e \in \mathbb{T}$:

- If $\|e\|^2 = 1$, then no edge arrives at e or leaves from e.
- If $\|e\|^2 = 2$, then no edge arrives at e, but two edges leave from e.
- If $\|e\|^2 > 2$, then one edge arrives at e, and two edges leave from e.

Furthermore the graph $\mathbb{T}$ is well founded, in the sense that there is no infinite sequence $e_0 \leftarrow e_1 \leftarrow \cdots$. Indeed the presence of an edge $e \to e'$ between two points implies a strict inequality $\|e\|^2 < \|e'\|^2$ on their squared norms, which are positive integers. These properties together characterize a graph of the form described in Proposition 1.17.

A.3. Two lemmas from the preprint [18].

Lemma (Lemma 2.3 in [18]). *Let* (e_0, e_1, e_2) *be a superbase of* $\mathbb{Z}^2$*, ordered so that* $\|e_0\| \geq \max\{\|e_1\|, \|e_2\|\}$ *and* $\det(e_1, e_2) = 1$*. Then* $-e_0 = e_1 \oplus e_2$*.*

Proof. Observing that $\det(e_0, e_1) = 1$, we find that the coordinates of e_0 are co-prime. Since e_0, e_1, e_2 are pairwise non-collinear, at least one of them is not in the set $V_4 := \{(\pm 1, 0), (0, \pm 1)\}$. Since e_0 has the largest norm, $e_0 \notin V_4$. By Proposition 1.13, there exists a direct acute basis (e'_1, e'_2) such that $-e_0 = e'_1 \oplus e'_2$.

Since $\det(e_0, e'_1 - e_1) = 1 - 1 = 0$, there exists $k \in \mathbb{R}$ such that $e'_1 = e_1 + ke_0$. Since e_0 has co-prime coordinates, $k \in \mathbb{Z}$. If $k > 0$, then we observe that $\langle e_1, -e_0 \rangle = \langle e'_1 - ke_0, -e_0 \rangle = \|e'_1\|^2 + \langle e'_1, e'_2 \rangle + k\|e_0\|^2 > \|e_0\|^2$, which implies the contradiction $\|e_1\| > \|e_0\|$. If $k < 0$, then observing that $e'_2 = -e_0 - e'_1 = e_2 - ke_0$ we reach a similar contradiction $\|e_2\| > \|e_0\|$. Thus $k = 0$, and therefore $e'_1 = e_1$, $e'_2 = e_2$, which concludes the proof. □

For any $f, g \in \mathbb{R}^2$, we denote $\overset{\circ}{\text{Cone}}(f, g) := \{\lambda f + \mu g;\ \lambda, \mu > 0\}$ (the interior of $\text{Cone}(f, g)$).

Lemma (Lemma 3.2 in [18], here Lemma 3.3). *Let* $e = f' \oplus g'$ *and let* (f, g) *be a direct acute basis of* $\mathbb{Z}^2$ *such that* $e \in \overset{\circ}{\text{Cone}}(f, g)$*. Then* $f + g$*,* f'*,* g' *belong to the triangle* $T := [e, f, g]$*.*

Proof. Let α, β denote the coordinates of e in the basis (f', g'), which are positive integers by construction. Observing that $1e + (\beta - 1)f + (\alpha - 1)g = (\alpha + \beta - 1)(f + g)$ we obtain as announced that $f + g \in T$.

We fix e and prove that $f', g' \in T := [e, f, g]$, for any direct acute basis (f, g) such that $e \in \overset{\circ}{\text{Cone}}(f, g)$, by *decreasing* induction on the integer $k := \langle f, g \rangle$. Initialization. Assuming that $k \geq \frac{1}{2}\|e\|^2$, we obtain the impossibility $\|e\|^2 = \|\alpha f + \beta g\|^2 > 2\alpha\beta\langle f, g \rangle \geq 2\langle f, g \rangle \geq \|e\|^2$. This case is vacuous, hence true.

Induction. If $e = f + g$, then $e = f \oplus g$ and therefore $f = f'$, $g = g'$; the result follows. Otherwise, we have either $e \in \overset{\circ}{\text{Cone}}(f, f + g)$ or $e \in \overset{\circ}{\text{Cone}}(f + g, g)$. By induction, since $\langle f, f + g \rangle > \langle f, g \rangle$ and $\langle f + g, g \rangle > \langle f, g \rangle$, we obtain that f', g' belong to $T_1 := [e, f, f + g]$ or $T_2 := [e, g, f + g]$. Recalling that $f + g \in T$ we obtain $T_1 \cup T_2 \subseteq T$ which concludes the proof. □

Acknowledgements

The authors gratefully acknowledge the support of the ANR, through the project ISOTACE (ANR-12-MONU-0013), the grant NS-LBR ANR-13-JS01-0003-01, and INRIA through the "action exploratoire" MOKAPLAN.

References

[1] J.-D. Benamou, B. D. Froese, and A. M. Oberman, *Numerical solution of the optimal transportation problem using the Monge-Ampère equation*, J. Comput. Phys. **260** (2014), 107–126, DOI 10.1016/j.jcp.2013.12.015. MR3151832

[2] J. F. Bonnans, É. Ottenwaelter, and H. Zidani, *A fast algorithm for the two dimensional HJB equation of stochastic control*, M2AN Math. Model. Numer. Anal. **38** (2004), no. 4, 723–735, DOI 10.1051/m2an:2004034. MR2087732 (2005e:93165)

[3] S. C. Brenner and M. Neilan, *Finite element approximations of the three dimensional Monge-Ampère equation*, ESAIM Math. Model. Numer. Anal. **46** (2012), no. 5, 979–1001, DOI 10.1051/m2an/2011067. MR2916369

[4] B. Chazelle, *An optimal convex hull algorithm in any fixed dimension*, Discrete Comput. Geom. **10** (1993), no. 4, 377–409, DOI 10.1007/BF02573985. MR1243335 (94h:52026)

[5] J. H. Conway and N. J. A. Sloane, *Low-dimensional lattices. VI. Voronoĭ reduction of three-dimensional lattices*, Proc. Roy. Soc. London Ser. A **436** (1992), no. 1896, 55–68, DOI 10.1098/rspa.1992.0004. MR1177121 (93h:11074)

[6] M. G. Crandall, H. Ishii, and P.-L. Lions, *User's guide to viscosity solutions of second order partial differential equations*, Bull. Amer. Math. Soc. (N.S.) **27** (1992), no. 1, 1–67, DOI 10.1090/S0273-0979-1992-00266-5. MR1118699 (92j:35050)

[7] J. Fehrenbach and J.-M. Mirebeau, *Sparse non-negative stencils for anisotropic diffusion*, J. Math. Imaging Vision **49** (2014), no. 1, 123–147, DOI 10.1007/s10851-013-0446-3. MR3180960

[8] X. Feng, R. Glowinski, and M. Neilan, *Recent developments in numerical methods for fully nonlinear second order partial differential equations*, SIAM Rev. **55** (2013), no. 2, 205–267, DOI 10.1137/110825960. MR3049920

[9] B. D. Froese and A. M. Oberman, *Convergent finite difference solvers for viscosity solutions of the elliptic Monge-Ampère equation in dimensions two and higher*, SIAM J. Numer. Anal. **49** (2011), no. 4, 1692–1714, DOI 10.1137/100803092. MR2831067 (2012i:65230)

[10] B. D. Froese and A. M. Oberman, *Convergent filtered schemes for the Monge-Ampère partial differential equation*, SIAM J. Numer. Anal. **51** (2013), no. 1, 423–444, DOI 10.1137/120875065. MR3033017

[11] B. D. Froese, *Numerical Methods for the Elliptic Monge-Ampere Equation and Optimal Transport*, ProQuest LLC, Ann Arbor, MI. Thesis (Ph.D.)–Simon Fraser University (Canada), 2012. MR3218235

[12] C. E. Gutiérrez, *The Monge-Ampère Equation*, Progress in Nonlinear Differential Equations and their Applications, 44, Birkhäuser Boston, Inc., Boston, MA, 2001. MR1829162 (2002e:35075)

[13] M. Kocan, *Approximation of viscosity solutions of elliptic partial differential equations on minimal grids*, Numer. Math. **72** (1995), no. 1, 73–92, DOI 10.1007/s002110050160. MR1359708 (97k:65239)

[14] H. J. Kuo and N. S. Trudinger, *Discrete methods for fully nonlinear elliptic equations*, SIAM J. Numer. Anal. **29** (1992), no. 1, 123–135, DOI 10.1137/0729008. MR1149088 (93e:65129)

[15] G. Loeper and F. Rapetti, *Numerical solution of the Monge-Ampère equation by a Newton's algorithm*, C. R. Math. Acad. Sci. Paris **340** (2005), no. 4, 319–324, DOI 10.1016/j.crma.2004.12.018. MR2121899

[16] J. J. Manfredi, A. M. Oberman, and A. P. Sviridov, *Nonlinear elliptic partial differential equations and p-harmonic functions on graphs*, Differential Integral Equations **28** (2015), no. 1-2, 79–102. MR3299118

[17] J.-M. Mirebeau, *Efficient fast marching with Finsler metrics*, Numer. Math. **126** (2014), no. 3, 515–557, DOI 10.1007/s00211-013-0571-3. MR3164145

[18] Jean-Marie Mirebeau, *Adaptive, Anisotropic and Hierarchical Cones of Convex functions*, preprint (2014).

[19] J.-M. Mirebeau, *Anisotropic fast-marching on Cartesian grids using lattice basis reduction*, SIAM J. Numer. Anal. **52** (2014), no. 4, 1573–1599, DOI 10.1137/120861667. MR3229657

[20] A. M. Oberman, *Convergent difference schemes for degenerate elliptic and parabolic equations: Hamilton-Jacobi equations and free boundary problems*, SIAM J. Numer. Anal. **44** (2006), no. 2, 879–895 (electronic), DOI 10.1137/S0036142903435235. MR2218974 (2007a:65173)

[21] A. M. Oberman, *A numerical method for variational problems with convexity constraints*, SIAM J. Sci. Comput. **35** (2013), no. 1, A378–A396, DOI 10.1137/120869973. MR3033053

[22] V. I. Oliker and L. D. Prussner, *On the numerical solution of the equation $(\partial^2 z/\partial x^2)(\partial^2 z/\partial y^2)-((\partial^2 z/\partial x\partial y))^2 = f$ and its discretizations. I*, Numer. Math. **54** (1988), no. 3, 271–293, DOI 10.1007/BF01396762. MR971703 (90h:65164)

[23] E. Selling, *Ueber die binären und ternären quadratischen Formen* (German), J. Reine Angew. Math. **77** (1874), 143–229, DOI 10.1515/crll.1874.77.143. MR1579602

[24] J. Urbas, *On the second boundary value problem for equations of Monge-Ampère type*, J. Reine Angew. Math. **487** (1997), 115–124, DOI 10.1515/crll.1997.487.115. MR1454261 (98f:35057)

Mokaplan, INRIA, Domaine de Voluceau, BP 105 78153, Le Chesnay Cedex, France
E-mail address: `jean-david.benamou@inria.fr`

Mokaplan, INRIA, Domaine de Voluceau BP 105 78153, Le Chesnay Cedex, France

Laboratoire de Mathématiques d'Orsay, University Paris-Sud, CNRS, University Paris-Saclay, 91405 Orsay, France
E-mail address: `jean-marie.mirebeau@math.u-psud.fr`