Classification of the equilibria for the semi-discrete Perona-Malik equation

G. Bellettini [*] , M. Novaga [†] , M. Paolini [‡] , C. Tornese [§]

Abstract

We give a complete classification of the stability properties of the equilibria for the semi-discrete one-dimensional Perona-Malik equation, with Dirichlet boundary conditions. We also give the Γ-expansion of the corresponding discretized functionals up to the order two, as the discretization parameter goes to zero.

Key words and phrases. Nonconvex functionals, semi-discrete schemes, stability of equilibria, Γ-expansion.

1 Introduction

In this paper we are interested in the analysis of the nonconvex functional

$$F(u) := \int_I \phi(u_x)\ dx, \tag{1.1}$$

where the smooth function ϕ is defined as

$$\phi(p) := \frac{1}{2}\log(1+p^2), \qquad p \in \mathbb{R},$$

and $I := (0,\ell) \subset \mathbb{R}$ is an interval. Note that ϕ has sublinear growth at infinity, it is strictly convex for $|p| < 1$ and *strictly concave* for $|p| > 1$. The formal gradient flow of F was considered by Perona and Malik in [19] for the study of certain problems in image segmentation [1], [20], [18] and turns out to give the interesting forward-backward parabolic equation

$$\frac{\partial u}{\partial t} = (\phi'(u_x))_x. \tag{1.2}$$

[*]Dipartimento di Matematica, Università di Roma "Tor Vergata", via della Ricerca Scientifica, 00133 Roma, Italy, and INFN Laboratori Nazionali di Frascati, via E. Fermi 40, 00044 Frascati (Roma), Italy. E-mail: Giovanni.Bellettini@lnf.infn.it

[†]Dipartimento di Matematica, Università di Pisa, Largo B. Pontecorvo 5, 56127 Pisa, Italy. E-mail: novaga@dm.unipi.it

[‡]Dipartimento di Matematica e Fisica, Università Cattolica di Brescia, via Trieste 41, 25121 Brescia, Italy. E-mail: paolini@dmf.unicatt.it

[§]Dipartimento di Matematica, Università di Roma "Tor Vergata", via della Ricerca Scientifica, 00133 Roma, Italy. E-mail: tornese@mat.uniroma2.it

We refer to [8], [16], [15], [11], [22], [3], [23], [14], [12], [13] for a discussion on the subject; here we just mention that the mathematical analysis of the Cauchy problem for (1.2) presents a lot of aspects still to be understood.
From a certain point of view, the functional F could be considered to have a rather poor structure: indeed (provided u_x stands for the absolutely continuous part of the derivative) it vanishes on piecewise constant functions, which constitute a dense subset of $L^2(I)$. However, the situation radically changes if we consider the finite element approximation F_h of F. Namely, consider the restriction

$$F_h(u) = h\sum_{i=1}^{N} \phi\Big(\frac{u_i - u_{i-1}}{h}\Big), \qquad u \in PL_h(I) \tag{1.3}$$

of F to $PL_h(I)$, where $PL_h(I)$ is the space of all piecewise linear functions on the mesh of I consisting of N subintervals of equal length $h := \ell/N$. The analysis of F_h may be relevant for applications and numerical analysis, and it turns out that the structure of discrete minimizers and critical points is rather rich. The aim of this paper is to explore such a structure.
Our main result is a complete classification of the equilibrium solutions of the system of ODEs

$$\dot{u}_i^h = \frac{1}{h}\left\{\phi'\Big(\frac{u_{i+1}^h - u_i^h}{h}\Big) - \phi'\Big(\frac{u_i^h - u_{i-1}^h}{h}\Big)\right\}, \qquad i \in \{1,\dots,N-1\}, \tag{1.4}$$

expressing the space discretization of (1.2), which we couple with the Dirichlet boundary conditions

$$u_0^h = \overline{u}_l, \qquad u_N^h = \overline{u}_r \tag{1.5}$$

(see Section 2 for the details on the notation). In Sections 2.2-2.4 and 3 (see in particular our main result, Theorem 3.2) we determine precisely the shape of each equilibrium of (1.4) (1.5) and, by a careful linear stability analysis, we count the number of its unstable directions. In particular, assume for simplicity that $\mathcal{H} := \overline{u}_r - \overline{u}_l > 0$, and that the equilibrium u has slope not identically equal to $+1$ or to -1. Then we show that u is a Dirichlet discrete local minimizer of F_h (i.e., u is stable) *if and only if* one of the following two conditions hold:

- u is linear with slope less than one,
- u has exactly two slopes $\alpha \in (0,1)$ and $\alpha^* = 1/\alpha > 1$; the "unstable" region of u where its slope is α^* consists in a single subinterval of the mesh (this, roughly speaking, would correspond to only one "jump" in the limit $h \to 0^+$, arbitrarily located in I), and

$$\alpha^2 = -\frac{\phi''(\alpha^*)}{\phi(\alpha)} < \frac{1}{N-1}. \tag{1.6}$$

Some consequences of the above result, in particular of (1.6), are discussed at the end of Section 3. The last section of the paper is devoted to identify the asymptotic Γ-expansion of F_h for small h up to the order 2. In the proof of the Γ-expansion result (Theorem 4.4) we make use on some informations on stable critical points obtained in the previous sections. The expression of the Γ-limit (4.2) should be compared with the one found in [2], where a fourth order regularization of (1.2) was considered, together with a rescaling of times.
The presence of many discrete critical points with different stability properties, coupled with a (non-trivial) passage to the limit (possibly along subsequences) as $h \to 0^+$ in (1.4), hopely allows to proceed further in the direction of defining a meaningful notion of solution to the original equation (1.2). This will be the object of future work. See also [9], [4], [10], [5] for related results.
We conclude this introduction observing that our model is similar to the finite-differences model studied by Schaeffer, Shearer and Witelski in [21], for the one-dimensional version of an ill-posed nonlinear parabolic problem. Their analysis has also implications to other ill-posed nonlinear PDEs such as (1.2) and some models for clustering instabilities in granular materials.

2 Notation and preliminary results

We set $I := (0, \ell)$ for some $\ell > 0$. Given a positive integer N we divide I in N closed subintervals of equal length $h := \ell/N$ (spatial mesh-size) with nodes $x_i = ih$, so that the i-th subinterval is $[(i-1)h, ih]$. We let $x_0 = 0$ and $x_N = \ell$. The ij-th entry of a matrix A will be denoted by a_{ij}, or $a_{i,j}$ if necessary.
$PL_h(I)$ (resp. $PC_h(I)$) is the $(N+1)$-dimensional vector subspace of $\mathrm{Lip}(I)$ (resp. N-dimensional vector subspace of $L^2(I)$) of all piecewise linear (resp. left-continuous piecewise constant) functions on the mesh.
Given $u \in PL_h(I)$ we denote with $u_1, \dots, u_{N-1}$ the coordinates of u at the interior nodes with respect to the basis of the hat functions, and $u_0 = \overline{u}_l$ and $u_N = \overline{u}_r$, where $\overline{u}_l, \overline{u}_r \in \mathbb{R}$ are given (Dirichlet data).
$PL_h^D(I)$ is the corresponding $(N-1)$-dimensional affine subspace defined as

$$PL_h^D(I) := \{u \in PL_h(I) : u(0) = \overline{u}_l, u(\ell) = \overline{u}_r\}.$$

Any function is identified with the vector of its nodal values. Given $w \in PC_h(I)$ we denote with $w_1, \dots, w_N$ the coordinates of w with respect to the basis of the flat functions.
We define the linear map $D_h^- : PL_h(I) \to PC_h(I)$ and its adjoint $D_h^+ : PC_h(I) \to PL_h(I)$ as

$$(D_h^- u)_i = \frac{1}{h}(u_i - u_{i-1}), \qquad i \in \{1, \dots, N\}, \tag{2.1}$$

$$(D_h^+ w)_i = \frac{1}{h}(w_{i+1} - w_i), \qquad i \in \{1, \dots, N-1\},$$

where in (2.1) the Dirichlet values are taken into account.
We will use indifferently the notation $\phi'((D_h^- u)_i)$ or $\phi'((D_h^- u))_i$.
The equation (1.2) is discretized in space in the standard way, obtaining

$$\frac{\partial u}{\partial t} = D_h^+ \phi'(D_h^- u)$$

namely the well posed system of ODEs for the unknown $u = u^h$

$$\dot{u}_i^h = \frac{1}{h}\left\{\phi'\left(\frac{u_{i+1}^h - u_i^h}{h}\right) - \phi'\left(\frac{u_i^h - u_{i-1}^h}{h}\right)\right\}, \qquad i \in \{1, \dots, N-1\}, \tag{2.2}$$

with the boundary conditions

$$u_{i-1}^h = \overline{u}_l \quad \text{when } i = 1, \qquad\qquad u_{i+1}^h = \overline{u}_r \quad \text{when } i = N-1. \tag{2.3}$$

Note that the restriction to $PL_h(I)$ (see (1.3)) of the functional F is a Liapunov functional for (2.2); in particular, u^h is an equilibrium solution of (2.2) if and only if it is a critical point of F_h. We say that u is a Dirichlet discrete local (resp. global) minimizer of F_h if u is a local (resp. global) minimizer of F_h in $PL_h(I) \subset L^1(I)$ under the conditions (2.3).
Finally, in what follows we use the notation

$$\psi := \phi'.$$

2.1 Spectrum of tridiagonal, real symmetric matrices

Theorem 2.1. (Gershgorin). *Let $A = (a_{ij})$ be an $(N \times N)$ complex matrix and define*

$$r_i := \sum_{j \neq i} |a_{ij}|, \quad C_i := \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\}, \quad C := \bigcup_i C_i.$$

Then

(i) the spectrum of A is contained in C,

(ii) if K is a connected component of C, and K is the union of k of the circles C_i, then K contains exactly k eigenvalues (the circles are counted with their multiplicity and the eigenvalues are counted with their algebraic multiplicity);

(iii) if K is the closure of a connected component of the interior of C, and K is the union of k of the circles C_i, then K contains at least k eigenvalues.

Proof. If λ is an eigenvalue of A and x a corresponding eigenvector we have $\|x\|_\infty = |x_k|$ for some k, and

$$\sum_{j \neq k} a_{kj} x_j = (\lambda - a_{kk}) x_k \quad \Rightarrow |\lambda - a_{kk}| \|x\|_\infty \leq \sum_{j \neq k} |a_{kj}| |x_j| \leq r_k \|x\|_\infty,$$

hence $\lambda \in C_k$. See also [17]. Consider now $A_s := sA + (1 - s)D$, where $D := \mathrm{diag}(a_{11}, \ldots, a_{NN})$ and $s \in [0, 1]$. Let $\lambda_i(s)$, $i \in \{1, \ldots, N\}$, be the eigenvalues of A_s ordered in such a way that $\lambda_i(s)$ is a continuous function with respect to s. Then $A_0 = D$, $\lambda_i(0) = a_{ii}$ and $A_1 = A$, $\lambda_i(1) = \lambda_i$. The Gerschgorin's circles of A_s have the same centers of the circles C_i, and radius sr_i, hence they are contained in the interior of C_i for $s \in [0, 1)$. From the continuity with respect to s it follows that the k eigenvalues that belong to K for $s = 0$ cannot leave the connected component K. □

2.1.1 Triple recursion relations for the determinants

Let W be a tridiagonal matrix with elements w_i in the diagonal, b_i in the subdiagonal, $i \in \{1, \ldots, N - 1\}$, c_i in the superdiagonal, $i \in \{2, \ldots, N\}$. The index i always refers to the number of the column of W. Let us denote by W_k the minor with order k made up of the first k rows and k columns and with q_k the characteristic polynomial of W_k.
The q_k are related by a triple recursion relation which is obtained calculating the determinant $W_k - \lambda I_k$ with respect to the last row and the last column:

$$q_{k+1}(\lambda) = \det(W_{k+1} - \lambda I_{k+1}) = (w_{k+1} - \lambda) q_k(\lambda) - b_k c_{k+1} q_{k-1}(\lambda). \tag{2.4}$$

Assuming by convention $q_0(\lambda) \equiv 1$ and recalling that $q_1(\lambda) = (w_1 - \lambda)$, the relation (2.4) remains valid even for the computation of $q_2(\lambda) = (w_2 - \lambda)(w_1 - \lambda) - b_1 c_2$.
If $b_i c_{i+1} > 0$ for every $i \in \{1, \ldots, N - 1\}$, it is known that the sequence $q_0, \ldots, q_N$ becomes a *Sturm sequence* [17, Chapter 3, Section 4.2], i.e. (besides the fact that q_0 does not change sign) it verifies the properties

$$q_{k-1}(\lambda) q_{k+1}(\lambda) < 0 \qquad \text{if } \lambda \text{ is such that } q_k(\lambda) = 0, \quad k = 1, \ldots, N - 1,$$

$$q_N'(\lambda) q_{N-1}(\lambda) < 0 \qquad \text{if } \lambda \text{ is such that } q_N(\lambda) = 0.$$

Theorem 2.2. (Givens). *Let the tridiagonal, real symmetric matrix W be defined as above with $b_i = c_{i+1} \neq 0$, $i \in \{1, \ldots, N - 1\}$. Then*

(i) the zeros of each q_k, $k = 2, \ldots, N$, are distinct and are separated by the zeros of q_{k-1};

(ii) if λ is such that $q_N(\lambda) \neq 0$, the number of eigenvalues of W that are strictly larger than λ is equal to the number of sign variations in the sequence

$$1, -q_1(\lambda), \ldots, (-1)^{N-1} q_{N-1}(\lambda), (-1)^N q_N(\lambda).$$

Proof. See for instance [17, Chapter 4, Section 3]. □

2.2 The system for the slopes: the matrix T

If we set, for $u \in PL_h^D(I)$,

$$v_i^h = v_i := (D_h^- u)_i \qquad i \in \{1, \dots, N\}, \qquad u = u^h,$$

the discrete Dirichlet problem (2.2) can be equivalently written (recalling also that $\dot{u}_0 = \dot{u}_N = 0$) as

$$\begin{cases} h^2 \dot{v}_i = \phi'(v_{i+1}) - 2\phi'(v_i) + \phi'(v_{i-1}), \quad i \in \{2, \dots, N-1\} \\ h^2 \dot{v}_1 = \phi'(v_2) - \phi'(v_1), \\ h^2 \dot{v}_N = -\phi'(v_N) + \phi'(v_{N-1}) \\ h \sum_{i=1}^{N} v_i = \overline{u}_r - \overline{u}_l =: \mathcal{H} \end{cases} \tag{2.5}$$

The global constraint expressed by the last equation in (2.5) replaces the boundary conditions in (2.3), and turns out to be automatically satisfied after the initial time, as can be proved by summing up the first three relations in (2.5).

Remark 2.3. Note that u is a critical point of F_h with Dirichlet boundary conditions if and only if $v := D_h^- u$ is an equilibrium solution of (2.2).

If $\phi'(v)$ is the vector defined as $\phi'(v)_i := \phi'(v_i)$, the system in (2.5) becomes

$$h^2 \dot{v} = -T\phi'(v), \tag{2.6}$$

where T is the real $(N \times N)$ tridiagonal symmetric positive semidefinite matrix defined as

$$T := \begin{pmatrix} 1 & -1 & & & & \\ -1 & 2 & -1 & & & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix}. \tag{2.7}$$

The properties of T used in the sequel are the following.

Lemma 2.4. *The matrix T has the following properties:*

(i) The spectrum of T is contained in the interval $[0,4]$; the eigenvalues $\lambda_1, \dots, \lambda_N$ of T are simple, $\lambda_1, \dots, \lambda_{N-1}$ are strictly positive and $\lambda_N = 0$;

(ii) there exists a lower triangular matrix $P = (p_{ij})$ such that

$$T = PP^t \tag{2.8}$$

(Cholesky factorization) which is singular with $p_{NN} = 0$.

Proof. The symmetric matrix T is singular, with eigenvector $\mathbf{v}_N = [1, \dots, 1]^t$ corresponding to the eigenvalue $\lambda_N = 0$. Gerschgorin's Theorem implies that the spectrum of T is contained in the interval $[0, 4]$, and $\mathbf{v}_N$ can be completed with an orthonormal basis $\{\mathbf{v}_i : i \in \{1, \dots, N\}\}$ with $\mathbf{v}_i \in \{x \in \mathbb{R}^N : \sum_{i=1}^N x_i = 0\}$ for any $1 \le i \le N-1$. From Givens' Theorem we have that all the eigenvalues of T are simple and so $\lambda_1, \dots, \lambda_{N-1}$ are strictly positive. In addition all principal minors T_k of T formed by the first k rows and k columns are positive definite with the only exception of $T = T_N$. This allows to use the Gauss elimination method without the need of pivotal strategies having non-null pivotal elements [17] . In the N-th step with Gaussian elimination we obtain the upper triangular matrix

$$\begin{pmatrix} 1 & -1 & & & & & & \\ 0 & 1 & -1 & & & & & \\ & \ddots & \ddots & \ddots & & & & \\ & & 0 & 1 & -1 & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & 0 & 1 & -1 \\ & & & & & 0 & 0 \end{pmatrix} \tag{2.9}$$

From the Gauss elimination method we obtain the factorization $T = LU$ with $L = (l_{ij})$ lower triangular, $l_{ii} = 1$, $U = (u_{ij})$ an upper triangular matrix, $u_{ii} > 0$ for $i \in \{1, \dots, N-1\}$ and $u_{NN} = 0$. Setting $G := \operatorname{diag}(\sqrt{u_{11}}, \dots, \sqrt{u_{NN}})$, we obtain the Cholesky factorization (2.8) with $P = LG$ and $P = (p_{ij})$ lower triangular, tridiagonal and singular, with $p_{NN} = 0$ (in this case its expression is equal to the transposed of (2.9)). □

2.3 First results on the nature of the equilibria

Definition 2.5. *Given $\alpha \in [-1, 1]$ we set $\psi^{-1}(\psi(\alpha)) = \{\alpha, \alpha^*\}$.*

Note that $\psi^{-1}(\psi(0)) = \{0\}$ and if $\psi(\alpha) \in \psi(\mathbb{R}) \setminus \{0\} = [-\frac{1}{2}, 0) \cup (0, \frac{1}{2}]$ then

$$\alpha^* = 1/\alpha. \tag{2.10}$$

Proposition 2.6. *A vector $v = (v_1, \dots, v_N) \in \mathbb{R}^N$ is an equilibrium solution of system (2.6) if and only if*

$$h \sum_{i=1}^{N} v_i = \mathcal{H} \tag{2.11}$$

and there exists $\alpha = \alpha(v) \in [-1, 1]$ such that

$$v_i \in \{\alpha, \alpha^*\} \qquad \forall i \in \{1, \dots, N\}. \tag{2.12}$$

Proof. The vector $v = (v_1, \dots, v_N) \in \mathbb{R}^N$ is an equilibrium solution of (2.5) if and only if v satisfies (2.11) and $\phi'(v) \in \operatorname{Ker}(T) = \operatorname{span}\{[1, \dots, 1]^t\}$, i.e., if there exists $L \in \phi'(\mathbb{R}) = [-1/2, 1/2]$ independent of i such that

$$L = \phi'(v_i) \qquad \forall i \in \{1, \dots, N\}.$$

□

Definition 2.7. *Let $v^{\text{eq}} = (v_1^{\text{eq}}, \dots, v_N^{\text{eq}})$ be an equilibrium solution of (2.5). We define the unstable and the stable region of v^{eq} as the following subsets of I:* $\operatorname{unstab}(v^{\text{eq}})$ *is the union of the subintervals $[(i-1)h, ih]$ of the mesh where $|v_i^{\text{eq}}| > 1$, and* $\operatorname{stab}(v^{\text{eq}}) := I \setminus \operatorname{unstab}(v^{\text{eq}})$.

We also define

$$\kappa(v^{\text{eq}}) \text{ the number of the subintervals of the mesh composing } \operatorname{unstab}(v^{\text{eq}}). \tag{2.13}$$

Remark 2.8. Once the Dirichlet value $\mathcal{H} = u(\ell) - u(0) = \overline{u}_r - \overline{u}_l > 0$ and $\kappa \in \mathbb{N} \cup \{0\}$ are given, we have a constraint for the choice of the slope $\alpha = \alpha_\kappa \in (0,1)$ such that

$$h\Big(\kappa\alpha^* + (N-\kappa)\alpha\Big) = \mathcal{H}. \tag{2.14}$$

Namely, using (2.10),

$$h(N-\kappa)\alpha^2 - \mathcal{H}\alpha + h\kappa = 0, \tag{2.15}$$

which is a parabola with concavity upwards, positive for $\alpha = 0$ and with a value $hN - \mathcal{H} = \ell - \mathcal{H}$ in $\alpha = 1$. If $\mathcal{H} > \ell$ (i.e. mean slope greater than ℓ) and since $\frac{\kappa}{N} \in [0,1]$ implies $1 \geq 4\frac{\kappa}{N}\left(1 - \frac{\kappa}{N}\right)$ we have $\mathcal{H}^2 > 4h^2\kappa(N-\kappa)$, hence there exists a solution

$$\alpha_\kappa = \frac{\mathcal{H} - \sqrt{\mathcal{H}^2 - 4h^2\kappa(N-\kappa)}}{2h(N-\kappa)} \tag{2.16}$$

of (2.15) with $0 < \alpha_\kappa < 1$. Moreover, it is not difficult to prove that $\frac{\mathcal{H}+\sqrt{\mathcal{H}^2-4h^2\kappa(N-\kappa)}}{2h(N-\kappa)} > 1$, and we conclude that if $\mathcal{H} > \ell$ there exists a unique solution $\alpha_\kappa \in (0,1)$ of (2.15).

The following result eliminates most candidates in the search for stable equilibria.

Proposition 2.9. *Let u^{cr} be a critical point of F_h with Dirichlet boundary conditions* (2.3) *and let v^{eq} be the equilibrium solution of* (2.5) *associated with u^{cr}. If $\kappa = \kappa(v^{\mathrm{eq}}) \geq 2$ then u^{cr} is not a local minimum of F_h.*

Proof. Let $\kappa \geq 2$ and let α_κ be given by (2.14). To show that u is not a local minimum we differentiate F_N along an appropriate curve $\sigma \in (-\delta,\delta) \to u(\sigma) \in \mathbb{R}^{N+1}$ through $u^{\mathrm{cr}} = u(0)$. We construct this one-parameter family of near-equilibrium states by perturbing two of the κ values where the slope is $v_i = \frac{(u^{\mathrm{cr}})_i - (u^{\mathrm{cr}})_{i-1}}{h} = \alpha_\kappa^*$, corresponding to the indices $i = i_1$ and $i = i_2$, as follows:

$$\sigma \to v_i(\sigma) := \begin{cases} \alpha_\kappa^* + \sigma & i = i_1 \\ \alpha_\kappa^* - \sigma & i = i_2 \\ \alpha_\kappa^* & i = i_j \quad \text{for } \; j \in \{3,4,\ldots,\kappa\} \;\; (\text{when } \;\kappa \geq 3) \\ \alpha_\kappa & \text{otherwise,} \end{cases}$$

where $\alpha_\kappa^* := (\alpha_\kappa)^*$ and we note that for $\sigma = 0$ we recover the equilibrium u^{cr}, while the constraint (2.11) is satisfied for all σ. Write $f(\sigma) := hF_h\big(u(\sigma)\big)$ and observe that

$$f(\sigma) = (\kappa-2)\phi(\alpha_\kappa^*) + (N-\kappa)\phi(\alpha_\kappa) + \phi(\alpha_\kappa^* + \sigma) + \phi(\alpha_\kappa^* - \sigma).$$

The conclusion then follows from

$$f'(0) = 0, \qquad f''(0) = 2\phi''(\alpha_\kappa^*) < 0.$$

□

A sort of converse of Proposition 2.9 will be given in Corollary 3.10, as a consequence of Theorem 3.2.

2.4 The linearized systems: the matrices A and $A_{u^{\rm cr}}$

In this section we begin the analysis of the linear stability of the equilibria of (2.5).

Definition 2.10. *Let $v = v^{\rm eq}$ be an equilibrium solution of* (2.5)*, and let α be the corresponding number given by Proposition 2.6. If $\alpha \in [0,1)$, we define*

$$r = r(v^{\rm eq}) := \phi''(\alpha) > 0, \qquad r^* = r^*(v^{\rm eq}) := \phi''(\alpha^*) < 0.$$

Since $\phi''(\alpha^*) = \frac{1-\alpha^{-2}}{(1+\alpha^{-2})^2} = \frac{\alpha^{-2}(\alpha^2-1)}{\alpha^{-4}(\alpha^2+1)^2} = -\alpha^2\phi''(\alpha)$, we have $\phi''(\alpha^*) > -\phi''(\alpha)$, and

$$\frac{|r^*|}{r} = \alpha^2 < 1.$$

To study the behaviour near an equilibrium solution $v^{\rm eq} = (v_1^{\rm eq}, \ldots, v_N^{\rm eq})$ we linearize (2.5) around $v^{\rm eq}$ and obtain the linearized system

$$h^2\dot{w} = -Aw \tag{2.17}$$

for the vector $w = (w_1, \ldots, w_N)$. Here $A = A(v^{\rm eq})$ is the $(N \times N)$ tridiagonal matrix defined as

$$A := \begin{pmatrix} \varphi_1'' & -\varphi_2'' & & & & & \\ -\varphi_1'' & 2\varphi_2'' & -\varphi_3'' & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & -\varphi_{i-1}'' & 2\varphi_i'' & -\varphi_{i+1}'' & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -\varphi_{N-2}'' & 2\varphi_{N-1}'' & -\varphi_N'' \\ & & & & & -\varphi_{N-1}'' & \varphi_N'' \end{pmatrix}$$

with

$$\varphi_i'' := \phi''(v_i^{\rm eq}) \in \{r, r^*\} \qquad \text{for any } i \in \{1, \ldots, N\}, \tag{2.18}$$

where the inclusion follows from Proposition 2.6.

Remark 2.11. If there exists $i \in \{1, \ldots, N\}$ such that $\varphi_i'' = 0$, then $v_i^{\rm eq} = \alpha = \alpha^* = 1$, and therefore $\varphi_j'' = 0$ for any $j \in \{1, \ldots, N\}$, and A is the null matrix. In the following we will always suppose that

$$r^* < 0 < r, \tag{2.19}$$

since the case $\alpha = \alpha^* = 1$ corresponds to $A \equiv 0$, and therefore the study of the linearized system (2.17) does not give information on the original problem.

The sum of the elements in each column of A is equal to 0, hence A is degenerate. If $\varphi_i'' \neq 0$ for any i, the minor of order $(N-1)$ obtained by eliminating the first row and the last column is nondegenerate, and so the rank of A is $N-1$.

Definition 2.12. *Let $v^{\rm eq}$ be an equilibrium solution of* (2.5)*. We define*

$$M = M(v^{\rm eq}) := \text{diag}(m_{11}, \ldots, m_{NN}), \tag{2.20}$$

$$m_{ii} = m_{ii}(v^{\rm eq}) := \begin{cases} r^*/r & \text{if } [(i-1)h, ih] \subseteq \text{unstab}(v^{\rm eq}), \\ 1 & \text{if } ((i-1)h, ih) \subseteq \text{stab}(v^{\rm eq}), \end{cases} \qquad i \in \{1, \ldots, N\}.$$

Recalling the expression of T in (2.7)*, we also define*

$$\widehat{A} = \widehat{A}(v^{\rm eq}) := TM. \tag{2.21}$$

Note that

$$A = r\widehat{A}.$$

Remark 2.13. Linearizing system (2.2), (2.3) around the corresponding solution $u^{\rm cr}$ yields the system

$$-h^2\dot{\omega}_i = -\phi''(v_i^{\rm eq})\omega_{i-1} + [\phi''(v_{i+1}^{\rm eq}) + \phi''(v_i^{\rm eq})]\omega_i - \phi''(v_{i+1}^{\rm eq})\omega_{i+1}, \qquad 0 < i < N,$$

in the unknown $\omega = (\omega_1, \dots, \omega_{N-1})$. In this case we have that

$$h^2\dot{\omega} = -A_{u^{\rm cr}}\ \omega \tag{2.22}$$

where $A_{u^{\rm cr}}$ is the $(N-1)\times(N-1)$ symmetric tridiagonal matrix defined as

$$A_{u^{\rm cr}} := \begin{pmatrix} \varphi_1'' + \varphi_2'' & -\varphi_2'' & & & & & \\ -\varphi_2'' & \varphi_2'' + \varphi_3'' & -\varphi_3'' & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & -\varphi_i'' & \varphi_i'' + \varphi_{i+1}'' & -\varphi_{i+1}'' & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & -\varphi_{N-2}'' & \varphi_{N-2}'' + \varphi_{N-1}'' & -\varphi_{N-1}'' \\ & & & & & -\varphi_{N-1}'' & \varphi_{N-1}'' + \varphi_N'' \end{pmatrix} \tag{2.23}$$

The possible presence of negative eigenvalues of A obviously indicates instability of $v^{\rm eq}$ with respect to small perturbations of the initial datum. The following is a first result on the nature of equilibria of (2.2). Recall that $\kappa(v^{\rm eq})$ is defined in (2.13).

Proposition 2.14. *Let $v^{\rm eq}$ be an equilibrium solution of* (2.6) *not identically* 1 *or* -1 *(see Remark 2.11). At least $\kappa(v^{\rm eq})$ eigenvalues of $\widehat{A}$ are nonpositive and at least $N - \kappa(v^{\rm eq})$ are nonnegative.*

Proof. The rows of the transposed matrix $\widehat{A}^t$ are multiplied by 1 (resp. by r^*/r). For any $i \in \{2, \dots, N-1\}$ the Gerschgorin circle of $\widehat{A}^t$ is centered at 2 (resp. $2r^*/r$) and has radius 2 (resp. $2|r^*/r|$); for $i = 1$ or $i = N$ the circle is centered at 1 (resp. r^*/r) and has radius 1 (resp. $|r^*/r|$). The positive and the negative circles are *tangent* at the origin. The thesis then follows from Theorem 2.1 (iii). □

3 Main result

In the following, given an equilibrium solution $v^{\rm eq}$ of (2.6), we write $\kappa = \kappa(v^{\rm eq})$, $r = r(v^{\rm eq})$, $r^* = r^*(v^{\rm eq})$, and $\widehat{A} = \widehat{A}(v^{\rm eq})$. We also set

$$\gamma = \gamma(v^{\rm eq}) := \frac{r^*}{r}.$$

Moreover, $u^{\rm cr}$ will always indicate the critical point of F_h associated with $v^{\rm eq}$.

Definition 3.1. *Let $v^{\rm eq}$ be an equilibrium solution of* (2.6)*. We define*

- $\mu^- = \mu^-(\kappa, \gamma, N)$ *the number of negative eigenvalues of $\widehat{A}$,*
- $\eta_0 = \eta_0(\kappa, \gamma, N)$ *the algebraic multiplicity of the eigenvalue* 0 *of $\widehat{A}$,*
- $\tilde{\eta}_0 = \tilde{\eta}_0(\kappa, \gamma, N)$ *the algebraic multiplicity of the eigenvalue* 0 *of $A_{u^{\rm cr}}$.*

From (2.21) and (2.19) it follows that the number of negative eigenvalues of A is equal to $\mu^-(\kappa,\gamma,N)$ and the algebraic multiplicity of the eigenvalue 0 of A is $\eta_0(\kappa,\gamma,N)$.
The main result of the paper is the following.

Theorem 3.2. *Let v^{eq} be an equilibrium solution of (2.6) not identically 1 or -1 and let u^{cr} be the corresponding critical point of F_h. The null eigenvalue of $\widehat{A}(v^{\text{eq}})$ (resp. of $A_{u^{\text{cr}}}$) has algebraic multiplicity either 1 or 2 (resp. either 0 or 1), while all remaining eigenvalues are real and simple. Moreover*

$$\begin{cases} \gamma > \gamma_{\text{crit}}(\kappa) & \Rightarrow \mu^- = \kappa - 1 \quad \text{and} \quad \eta_0 = 1 \text{ (resp. } \tilde{\eta}_0 = 0) \\ \gamma < \gamma_{\text{crit}}(\kappa) & \Rightarrow \mu^- = \kappa \quad \text{and} \quad \eta_0 = 1 \text{ (resp. } \tilde{\eta}_0 = 0) \\ \gamma = \gamma_{\text{crit}}(\kappa) & \Rightarrow \mu^- = \kappa - 1 \quad \text{and} \quad \eta_0 = 2 \text{ (resp. } \tilde{\eta}_0 = 1), \end{cases} \tag{3.1}$$

where

$$\gamma_{\text{crit}}(n) := \begin{cases} -\dfrac{n}{N-n} & \text{if} \quad n \in \{0,1,\dots,N-1\}, \\ -\infty & \text{if} \quad n = N. \end{cases}$$

As a consequence, we obtain the follow stability criterion.

Corollary 3.3. *Let v^{eq} be an equilibrium solution of (2.6) not identically 1 or -1. Then v^{eq} is a stable equilibrium if and only if one of the following two cases holds:*

(i) $\kappa = 0$*;*

(ii) $\kappa = 1$ *and* $\gamma > -\frac{1}{N-1}$.

Proof. If v^{eq} is a stable equilibrium, then $\widehat{A}(v^{\text{eq}})$ has no negative eigenvalues, i.e. $\mu^- = 0$. Therefore $\kappa \geq 2$ is not allowed by Theorem 3.2. Conversely, assume that $\kappa = 0$. Then $\gamma_{\text{crit}}(0) = 0$, and therefore $\gamma < 0$. Hence from Theorem 3.2 it follows that $\mu^- = 0$. Similarly, assume that $\kappa = 1$ and $\gamma > \gamma_{\text{crit}}(1) = -\frac{1}{N-1}$. Again, from Theorem 3.2 it follows $\mu^- = 0$. □

3.1 Proof of Theorem 3.2

To compute the eigenvalues of $\widehat{A} = \widehat{A}(v^{\text{eq}})$ and $A_{u^{\text{cr}}}$ defined in (2.21), (2.23) respectively, we will use similarity transformations.

Lemma 3.4. *The following properties hold:*

(i) the eigenvalues of $\widehat{A}$ are real;

(ii) there exists an invertible matrix $\widetilde{P}$ such that

$$B = B(v^{\text{eq}}) := \widetilde{P}^{-1}\widehat{A}\widetilde{P}$$

is tridiagonal with the last row null. Moreover the $(N-1)\times(N-1)$ principal minor $B_{N-1} = B_{N-1}(v^{\text{eq}})$ of B satisfies

$$rB_{N-1} = A_{u^{\text{cr}}}; \tag{3.2}$$

(iii) B has at least a null eigenvalue. All other eigenvalues of B, which correspond to those of B_{N-1}, are real and simple.

Proof. For $\epsilon > 0$ we define the matrix P_ϵ substituting the null element with ϵ on the diagonal of P, defined in the proof of Lemma 2.4, in position p_{NN}. It turns out that $T_\epsilon = P_\epsilon P_\epsilon^t$ (compare (2.8)) differs from T only for the element in position (N,N) which becomes $1+\epsilon^2$ instead of 1. T_ϵ is a symmetric positive definite matrix (and all its eigenvalues are distinct).

Set $\widehat{A}_\epsilon := T_\epsilon M$ (compare (2.21)); then $\lim_{\epsilon\to 0}\widehat{A}_\epsilon = \widehat{A}$. Moreover, the matrix

$$B_\epsilon := P_\epsilon^{-1}\widehat{A}_\epsilon P_\epsilon$$

is symmetric, since

$$B_\epsilon^t = P_\epsilon^t M T_\epsilon (P_\epsilon^{-1})^t = P_\epsilon^t M P_\epsilon = P_\epsilon^{-1} P_\epsilon P_\epsilon^t M P_\epsilon = P_\epsilon^{-1} T_\epsilon M P_\epsilon = P_\epsilon^{-1}\widehat{A}_\epsilon P_\epsilon = B_\epsilon.$$

So B_ϵ has real eigenvalues, which coincide with those of $\widehat{A}_\epsilon$ and passing to the limit as $\epsilon \to 0$ also the eigenvalues of $\widehat{A}$ are real.
Since the matrix P is singular, it can not be directly used for a similarity transformation of $\widehat{A}$. We then use $\widetilde{P} := P_1$ ($= P_\epsilon$ for $\epsilon = 1$). It immediately follows that $\widetilde{P}$ is lower triangular and tridiagonal, with 1 on the diagonal and with -1 on the subdiagonal. The inverse matrix of $\widetilde{P}$ can be written as

$$(\widetilde{P}^{-1})_{ij} = \begin{cases} 1 & \text{if } j \le i \\ 0 & \text{if } j > i \end{cases} \tag{3.3}$$

The matrix

$$B := \widetilde{P}^{-1}\widehat{A}\widetilde{P}$$

is tridiagonal, and one can check that

$$b_{ij} = \sum_k \sum_l (\widetilde{P}^{-1})_{ik}\hat{a}_{kl}(\widetilde{P})_{lj} = m_{jj}\sum_{k\le i} t_{kj} - m_{(j+1),(j+1)}\sum_{k\le i} t_{k,(j+1)} \tag{3.4}$$

where the second term is missing if $j = N$ and we have set $T = (t_{ij})$ and $\hat{A} = (\hat{a}_{ij})$. If $1 < j < N-1$ we have

$$b_{ij} = \begin{cases} 0 & \text{if } |i-j| > 1 \\ -m_{jj} & \text{if } i = j-1 \\ m_{jj} + m_{(j+1),(j+1)} & \text{if } i = j \\ -m_{(j+1),(j+1)} & \text{if } i = j+1 \end{cases} \tag{3.5}$$

which proves to be valid also for $j = 1$ and for $j = N-1$ if $i < N$. A direct test eventually leads to:

$$b_{(N-1),N} = m_{NN}, \qquad b_{N,(N-1)} = b_{NN} = 0.$$

Therefore the $(N-1)\times(N-1)$ principal minor B_{N-1} is symmetric and tridiagonal and, from (2.23) and (2.20), we obtain the equality (3.2).
It remains to prove (iii). The null eigenvalue proves to be evident from the last row, while all other eigenvalues of B correspond to those of the principal minor B_{N-1} which, from Givens' Theorem 2.2, are real and simple. □

Corollary 3.5. *The eigenvalues of $\widehat{A}$ coincide with those of B. Moreover, the null eigenvalue of $\widehat{A}$ has algebraic multiplicity either* 1 *or* 2.

Proof. It is a consequence of Lemma 3.4. □

Remark 3.6. From Lemma 3.4 it follows that the number of negative eigenvalues of both B_{N-1} and $A_{u^{cr}}$ is equal to $\mu^-(\kappa,\gamma,N)$ and the algebraic multiplicity of the null eigenvalue of B_{N-1} is $\tilde{\eta}_0(\kappa,\gamma,N)$.

We want to understand in what circumstances the minor B_{N-1} has some negative eigenvalue.

Proposition 3.7 (Sign of the eigenvalues of B_{N-1}). *The following properties hold:*

- *if $\kappa = 0$ then B_{N-1} has no negative eigenvalues;*
- *if $\kappa = 1$ then B_{N-1} is positive definite if $\gamma > -\frac{1}{N-1}$ and has a negative eigenvalue if $\gamma < -\frac{1}{N-1}$. If $\gamma = -\frac{1}{N-1}$ then B_{N-1} has one zero eigenvalue and all remaining eigenvalues are positive;*
- *if $\kappa \ge 2$ then B_{N-1} has at least one negative eigenvalue.*

Proof. Let d_j be the determinant of the principal minor of $B = (b_{nm})$ with order $j \in \{1, \dots, N\}$ and assume $d_0 = 1$ by convention. Recalling (2.4) we have for $2 \le j < N$:

$$d_j = b_{jj} d_{j-1} - (b_{j,j-1})^2 d_{j-2} = (m_{jj} + m_{(j+1),(j+1)}) d_{j-1} - (m_{jj})^2 d_{j-2}, \tag{3.6}$$

where we have used the symmetry of B_{N-1} and formula (3.5). Theorem 2.2 (applied with N replaced by $N-1$, $W = -B_{N-1}$ and $\lambda = 0$) implies that the number of negative eigenvalues of B_{N-1} corresponds to the number of sign variations in the sequence $d_0, d_1, \dots, d_{N-1}$.

Assume $\kappa = 0$. Then $m_{ii} = 1$ for every i. Therefore (3.6) becomes

$$d_j = 2d_{j-1} - d_{j-2}, \tag{3.7}$$

coupled with the initial conditions

$$d_0 = 1, \quad d_1 = 2. \tag{3.8}$$

The general solution of (3.7) is $d_j = p_0(j)$ where p_0 is a suitable first order polynomial. By imposing (3.8) we obtain $p_0(j) = j + 1 > 0$. Therefore there are no changes of sign in the sequence $d_0, d_1, d_2, \dots, d_{N-1}$, hence B_{N-1} has no negative eigenvalues.

Assume $\kappa = 1$. Write $\text{unstab}(v^{\text{eq}}) = [(k_0 - 1)h, k_0 h]$ with $k_0 > 1$ (the case $k_0 = 1$ gives the same result recalling that $d_0 = 1$). Then $m_{(k_0-1),(k_0-1)} = 1$, $m_{k_0,k_0} = \gamma$ and using the previous case (hence $d_{k_0-2} = k_0 - 1$, $d_{k_0-3} = k_0 - 2$)

$$d_{k_0-1} = (1+\gamma)(k_0 - 1) - (k_0 - 2) = 1 + (k_0 - 1)\gamma. \tag{3.9}$$

If (3.9) is negative there would be a change in the sign and therefore a negative eigenvalue. Therefore we can suppose $1 + (k_0 - 1)\gamma \ge 0$, i.e.

$$\gamma \ge -\frac{1}{k_0 - 1}. \tag{3.10}$$

A direct computation based on (3.6) and (3.9) shows that

$$d_{k_0} = (1+\gamma)\left((1 + (k_0 - 1)\gamma\right) - \gamma^2 (k_0 - 1) = 1 + k_0\gamma \tag{3.11}$$

and again we can assume (3.11) to be positive, i.e.

$$\gamma \ge -\frac{1}{k_0}, \tag{3.12}$$

which is a condition more restrictive than (3.10).
From now on the identity $d_j = p_1(j)$ is restored, where $p_1(j)$ is a first order polynomial that must be chosen so as to obtain the expressions (3.9), (3.11) for the indices $j = k_0 - 1$ e $j = k_0$, i.e.:

$$p_1(k_0 - 1) = 1 + (k_0 - 1)\gamma, \qquad p_1(k_0) = 1 + k_0\gamma. \tag{3.13}$$

Subtracting the two relations in (3.13) we obtain

$$p_1(k_0) - p_1(k_0 - 1) = \gamma$$

It follows that the value d_j strictly decreases by the amount $|\gamma|$ as j increases, until it reaches the value $d_{N-1} = d_{k_0} + (N - 1 - k_0)\gamma = 1 + N\gamma - \gamma$ when $j = N - 1$. The problem reaches a stable condition (no negative eigenvalues) when

$$\gamma > -\frac{1}{N-1} \tag{3.14}$$

Since the sequence $\{d_j\}$ is monotone decreasing for $k_0 - 1 \leq j \leq N-1$, only a single sign change can occur, if $d_{N-1} < 0$. Since $d_{N-1}(\gamma) = 0$ has a simple zero for $\gamma = -\frac{1}{N-1}$, the matrix B_{N-1} has a single negative eigenvalue for $\gamma < -\frac{1}{N-1}$.
Note that the null eigenvalue of B has multiplicity 2 in case of equality in (3.14).

Assume $\kappa = 2$. Write $\text{unstab}(v^{\text{eq}}) = [(k_0-1)h, k_0 h] \cup [(k_1-1)h, k_1 h]$ and suppose, without loss of generality, that $k_1 > k_0$. Proceeding as in the case $\kappa = 1$ we obtain $d_j = 1 + j\gamma$ for $k_0 - 1 \leq j < k_1 - 1$ where the inequality $d_{k_1-2} > 0$ implies

$$\gamma > -\frac{1}{k_1 - 2}.$$

From (3.6) we get

$$\begin{aligned} d_{k_1-1} &= (1+\gamma)\Big(1 + (k_1-2)\gamma\Big) - \Big(1 + (k_1-3)\gamma\Big) = (k_1-2)\gamma^2 + 2\gamma, \\ d_{k_1} &= (1+\gamma)\Big(\gamma^2(k_1-2) + 2\gamma\Big) - \gamma^2\Big(1 + \gamma(k_1 - 2\gamma)\Big) = (k_1-1)\gamma^2 + 2\gamma. \end{aligned}$$

We can use these relations to obtain

$$d_j = (j-1)\gamma^2 + 2\gamma, \qquad j \geq k_1 - 1.$$

This is an increasing sequence in j; we are interested in the value for $j = k_1 - 1$,

$$d_{k_1-1} = \gamma\Big(1 + 1 + (k_1-2)\gamma\Big) = \gamma(1 + d_{k_1-2}).$$

The latter necessarily becomes negative if d_{k_1-2} is nonnegative. It follows that there exists at least one negative eigenvalue for B_{N-1}. There are two negative eigenvalues if and only if d_{N-1} is positive which would necessarily imply two changes of sign in the sequence $\{d_j\}$.

Assume $\kappa > 2$. Write $\text{unstab}(v^{\text{eq}}) = \bigcup_{i \in \mathcal{I}} [(k_{i-1}-1)h, k_{i-1}h]$ where $\mathcal{I} = \{1, \ldots, \kappa\}$. If $k_{i-1} < k_i$ are two consecutive indices in $\mathcal{I}$, then for $k_{i-1} - 1 \leq j < k_i - 1$, $d_j = p_i(j)$ where p_i is an appropriate first order polynomial $p_i(j) = \alpha_i j + \beta_i$. We want to determine p_i on the basis of p_{i-1}, setting $p_0(j) = j + 1$. Then

$$\begin{aligned} d_{k_i-1} &= (1+\gamma)p_{i-1}(k_i-2) - p_{i-1}(k_i-3) \\ &= \gamma p_{i-1}(k_i-2) + \alpha_{i-1} \\ &= \gamma\alpha_{i-1}(k_i-2) + \gamma\beta_{i-1} + \alpha_{i-1}, \\ d_{k_i} &= (1+\gamma)d_{k_i-1} - \gamma^2 p_{i-1}(k_i-2) \end{aligned}$$

and so

$$\begin{aligned} \alpha_i &= d_{k_i} - d_{k_i-1} = \gamma(d_{k_i-1} - \gamma p_{i-1}(k_i-2)) \\ &= \gamma(p_{i-1}(k_i-2) - p_{i-1}(k_i-3)) = \gamma\alpha_{i-1}, \\ \beta_i &= d_{k_i-1} - \alpha_i(k_i-1) \\ &= \gamma\beta_{i-1} + \alpha_{i-1} - \gamma\alpha_{i-1}. \end{aligned}$$

It follows that $\alpha_i = \gamma^i$ and $\beta_i = \gamma\beta_{i-1} + (1-\gamma)\gamma^{i-1} = \gamma^{i-1}(\gamma + i - i\gamma)$. In particular the determinant of the minor $(N-1) \times (N-1)$ can be calculated with

$$d_{N-1} = \alpha_\kappa(N - \kappa + \kappa\gamma^{-1}) \tag{3.15}$$

and it cancels out only for the value $\gamma_{\text{crit}}(\kappa) = -\frac{\kappa}{N-\kappa}$ for $\kappa \in \{0, 1, \ldots, N-1\}$. For $\kappa = N$, $d_{N-1} \neq 0$. Besides, it is clear that d_{N-1} changes its sign when γ is greater than this critical value, in fact the term in brackets in (3.15) is decreasing with respect to γ and therefore it is negative when $\gamma_{\text{crit}}(\kappa) < \gamma < 0$, whereas α_κ has sign $(-1)^\kappa$. ☐

Remark 3.8. In order to calculate μ^- we should observe that if $d_{N-1} \neq 0$ than the eigenvalue 0 is simple and therefore μ^- can assume the values κ or $\kappa - 1$ only, in fact Lemma 2.14 implies that at least $N - \kappa$ eigenvalues are nonnegative and therefore the negative ones are at most κ. In other words, since 0 is an eigenvalue with multiplicity 1, it follows that the negative eigenvalues are at least $\kappa - 1$. On the other hand the sign of d_{N-1} gives the parity of the negative eigenvalues; $d_{N-1}(-1)^{\mu^-} > 0$. Thus if $\gamma_{\text{crit}}(\kappa) < \gamma < 0$, μ^- has opposite parity with respect to the parity κ and therefore $\mu^- = \kappa - 1$. Hence we obtain

$$\begin{cases} \gamma > \gamma_{\text{crit}}(\kappa) & \Rightarrow \mu^- = \kappa - 1, \\ \gamma < \gamma_{\text{crit}}(\kappa) & \Rightarrow \mu^- = \kappa. \end{cases}$$

Eventually, the proof of Theorem 3.2. is a consequence of (i) of Lemma 3.4, Proposition 3.7 and Remarks 3.6 and 3.8. □

Remark 3.9. Let $\kappa \in \{1, \dots, N\}$. If $\mathcal{H} > \ell$, from Remark 2.8 we have one and only one satisfactory solution $\alpha_\kappa \in (0,1)$ given by (2.16). With the approximation $N >> \kappa$ we have from (2.16) that

$$\alpha_\kappa \approx \frac{h\kappa}{\mathcal{H}}.$$

The ratio γ is given by $\gamma = \frac{r^*}{r} = -\alpha^2 \approx -\frac{h^2\kappa^2}{\mathcal{H}^2}$ which entails $|\gamma| < |\gamma_{\text{crit}}(\kappa)|$, i.e. $\mu^- = \kappa - 1$.

We have seen in Proposition 2.9 that

$$\mathcal{H} > 0, \quad u^{\text{cr}} \text{ Dirichlet discrete local minimizer of } F_h \quad \Longrightarrow \quad \kappa \in \{0, 1\}. \tag{3.16}$$

The following corollary gives a sort of converse of (3.16).

Corollary 3.10. *Assume that $\mathcal{H} > 0$. Let v^{eq} be a critical point of* (2.6) *and let $\kappa = \kappa(v^{\text{eq}})$. Then*

$$\kappa \in \{0, 1\}, \quad N \text{ sufficiently large} \quad \Longrightarrow \quad u^{\text{cr}} \text{ Dirichlet discrete local minimizer of } F_h.$$

Proof. Recalling Corollary 3.3, it is enough to consider the case $\kappa = 1$ and to show that (1.6) is satisfied for N sufficiently large. From (2.16) we have

$$\alpha = \frac{\mathcal{H} - \sqrt{\mathcal{H}^2 - 4h^2(N-1)}}{2h(N-1)} = \frac{h}{\mathcal{H}} + o(h) = \frac{\ell}{N\mathcal{H}} + o\left(\frac{1}{N}\right). \tag{3.17}$$

Hence $\alpha^2 = \frac{\ell^2}{N^2\mathcal{H}^2} + o\left(\frac{1}{N^2}\right)$, and therefore (1.6) is valid provided

$$\frac{1}{N}\frac{\ell^2}{\mathcal{H}^2} + o\left(\frac{1}{N}\right) < 1.$$

□

Remark 3.11. Assume that $\mathcal{H} > 0$. Let v^{eq} be a critical point of (2.6) and assume that $\kappa = \kappa(v^{\text{eq}}) = 0$. If N is sufficiently large then u^{cr} is a Dirichlet discrete local but not global minimizer of F_h. Indeed, we already know that if $N >> 1$ then u^{cr} is a local minimizer. A direct computation shows that $F_h(u^{\text{cr}}) = O(1)$. On the other hand, if we take a stable equilibrium v_h^* of (2.6) with $\kappa(v_h^*) = 1$, one checks that the corresponding u_h^* is such that $F_h(u_h^*) = o(1)$ for $h > 0$ small enough. We conclude that Dirichlet discrete global minimizers of F_h necessarily have $\kappa = 1$ for N large enough, independently on the value of $\mathcal{H}$.

4 Γ-expansion of the discrete functionals F_h

We recall [7] that a family of functionals $G_h : L^1(I) \to [-\infty, +\infty]$ Γ-converges to $G^{(0)} := \Gamma - \lim_{h\to 0^+} G_h : L^1(I) \to [-\infty, +\infty]$ as $h \to 0^+$ if

- (Γ-liminf inequality) for all $u \in L^1(I)$ and for every sequence $\{u_h\}$ converging to u in $L^1(I)$ we have $G^{(0)}(u) \le \liminf_{h\to 0^+} G_h(u_h)$;
- (Γ-limsup inequality) for all $u \in L^1(I)$ there exists a sequence $\{u_h\}$ converging to u in $L^1(I)$ (called recovery sequence) such that $G^{(0)}(u) = \lim_{h\to 0^+} G_h(u_h)$.

We now give the definition of asymptotic expansion by Γ-convergence, following [6].

Definition 4.1. *Let $G_h : L^1(I) \to [0, +\infty]$ be a family of functionals, let $G^{(0)} := \Gamma - \lim_{h\to 0^+} G_h$ and let $w^{(0)}$ be a local minimizer of $G^{(0)}$. Let $f_1, f_2 : (0, 1) \to [0, +\infty]$ with $\lim_{h\to 0^+} f_1(h) = 0$ and $f_2(h) \in o(f_1(h))$ as $h \to 0^+$. Finally let $w^{(1)} \in L^1(I)$. We write*

$$G_h \stackrel{\Gamma}{=} G^{(0)} + f_1(h)G^{(1)} + f_2(h)G^{(2)} + o(f_2(h)), \tag{4.1}$$

and we say that (4.1) is the Γ-expansion of G_h up to the order 2, if

- $G^{(1)} = \Gamma - \lim_{h\to 0^+} \dfrac{G_h - G^{(0)}(w^{(0)})}{f_1(h)}$;
- *$w^{(1)}$ is a local minimizer of $G^{(1)}$;*
- $G^{(2)} = \Gamma - \lim_{h\to 0^+} \dfrac{G_h - G^{(0)}(w^{(0)}) - f_1(h)G^{(1)}(w^{(1)})}{f_2(h)}$.

Remark 4.2. The authors in [6] consider global minimizers instead of local ones, hence their Γ-expansion does not depend on the choice of $w^{(0)}$ and $w^{(1)}$.

We extend the definition of F_h to the whole of $L^1(I)$ by setting $F_h \equiv +\infty$ in $L^1(I) \setminus PL_h(I)$.
Denoting by $PC(I)$ the space of piecewise constant functions defined on I with a finite number of jump points, we define $F^{(0)}, F^{(1)} : L^1(I) \to [0, +\infty]$ as

$$F^{(0)} \equiv 0, \qquad F^{(1)}(u) := \begin{cases} \#J_u & \text{if } u \in PC(I) \\ +\infty & \text{if } u \in L^1(I) \setminus PC(I), \end{cases}$$

where $J_u \subset I$ is the set of jump points of u and $\#$ its cardinality.

Remark 4.3. Any $\overline{u} \in PC(I)$ is a local minimizer of $F^{(1)}$.

Finally, given $\overline{u} \in PC(I)$ we define $F^{(2)}_{\overline{u}} : L^1(I) \to [-\infty, +\infty]$ as

$$F^{(2)}_{\overline{u}}(u) := \begin{cases} -\infty & \text{if } u \in PC(I) \text{ and } \#J_u < \#J_{\overline{u}} \\ \sum_{x\in J_u} \log|u(x_+) - u(x_-)| & \text{if } u \in PC(I) \text{ and } \#J_u = \#J_{\overline{u}} \\ +\infty & \text{if } u \in L^1(I) \setminus PC(I), \end{cases} \tag{4.2}$$

where $u(x_\pm)$ are the right and left limits of $u \in PC(I)$ at $x \in J_u$.

Theorem 4.4. *Fix $\overline{u} \in PC(I)$. The following Γ-expansion of F_h up to the order 2 holds:*

$$F_h \stackrel{\Gamma}{=} F^{(0)} + h\phi\left(\frac{1}{h}\right) F^{(1)} + hF^{(2)}_{\overline{u}} + o(h).$$

Proof. It is easy to check that $\Gamma-\lim_h F_h \equiv 0 = F^{(0)}$ using the property of sublinear growth at infinity of ϕ together with the $L^1(I)$-density of the space $PC(I)$ in $L^1(I)$, and the lower semicontinuity of Γ-limits.
We now show that

$$F^{(1)} = \Gamma - \lim_{h\to 0^+} \frac{F_h}{h\phi(1/h)}.$$

In order to check the Γ-limsup inequality, since the argument is local, it is enough to consider the case of a function u having a single jump. In particular, if we let

$$u(x) := \begin{cases} a_0 & \text{if } x \in (0,\overline{x}), \\ a_1 & \text{if } x \in (\overline{x},\ell), \end{cases} \tag{4.3}$$

where $\overline{x} \in (0,\ell)$, a recovery sequence $\{u_h\} \subset PL_h(I)$ can be constructed as follows:

$$u_h(x) := \begin{cases} a_0 & \text{if } 0 < x < h\lfloor \overline{x}/h \rfloor \\ \left(\left\lfloor \frac{\overline{x}}{h} \right\rfloor + 1 - \frac{x}{h}\right) a_0 + \left(\frac{x}{h} - \left\lfloor \frac{\overline{x}}{h} \right\rfloor\right) a_1 & \text{if } h\lfloor \overline{x}/h \rfloor < x < h\left(\lfloor \overline{x}/h \rfloor + 1\right) \\ a_1 & \text{if } h\left(\lfloor \overline{x}/h \rfloor + 1\right) < x < \ell, \end{cases} \tag{4.4}$$

where $\lfloor \cdot \rfloor$ denotes the integer part. Indeed, $u_h \to u$ in $L^1(I)$, and letting $\overline{i}_h := \left\lfloor \frac{\overline{x}}{h} \right\rfloor + 1$, we have

$$\begin{aligned} \lim_{h\to 0^+} \frac{F_h(u_h)}{h\phi(1/h)} &= \lim_{h\to 0^+} \frac{1}{\phi(1/h)} \sum_{i=1}^{\ell/h} \phi\left(\frac{(u_h)_i - (u_h)_{i-1}}{h}\right) \\ = \quad & \lim_{h\to 0^+} \frac{1}{\phi(1/h)} \log\left(\frac{\left|(u_h)_{\overline{i}_h} - (u_h)_{\overline{i}_h - 1}\right|}{h}\right) \\ = \quad & \lim_{h\to 0^+} \left(\frac{\log\left|(u_h)_{\overline{i}} - (u_h)_{\overline{i}-1}\right|}{|\log h|} + 1\right) = 1 = \#J_u. \end{aligned} \tag{4.5}$$

To prove the Γ-liminf inequality, let $u \in L^1(I)$ and let $u_h \to u$ in $L^1(I)$ as $h \to 0^+$. We have to show that

$$F^{(1)}(u) \le \liminf_{h\to 0^+} \frac{F_h(u_h)}{h\phi(1/h)}. \tag{4.6}$$

We can assume that $\liminf_{h\to 0^+} \frac{F_h(u_h)}{h\phi(1/h)} < +\infty$ (so that $\{u_h\} \subset PL_h(I)$), otherwise there is nothing to prove, and that u is not identically constant in I. Possibly passing to a (not relabelled) subsequence, we can assume that the $\liminf_{h\to 0^+} \frac{F_h(u_h)}{h\phi(1/h)}$ is a limit, that $c := \sup_h \dfrac{F_h(u_h)}{h\phi(1/h)} < +\infty$, and that $u_h \to u$ almost everywhere as $h \to 0^+$. As a consequence, we can fix points

$$0 < \overline{x}_1 < \cdots < \overline{x}_{n+1} < \ell, \tag{4.7}$$

with $n \ge 1$, such that each $\overline{x}_i$ is a Lebesgue point of u, $\lim_{h\to 0^+} u_h(\overline{x}_i) = u(\overline{x}_i)$, and $u(\overline{x}_{i+1}) \ne u(\overline{x}_i)$. Moreover, without loss of generality, we can suppose $u_h(\overline{x}_{i+1}) - u_h(\overline{x}_i) > 0$. The thesis follows if we prove that

$$\liminf_{h\to 0^+} \frac{F_h(u_h)}{h\phi(1/h)} \ge n. \tag{4.8}$$

Note that the points $\overline{x}_i$ may not belong to the mesh. In order to prove (4.8) we use the characterization of Dirichlet discrete minimizers of F_h established in Section 3, in particular Remark 3.11. For any $h > 0$ and $i \in \{1,\dots,n\}$ we let $x_i^h := h\left\lfloor \frac{\overline{x}_i}{h} \right\rfloor$ and we define

$$I_i^h := (x_i^h, x_{i+1}^h + h).$$

Note that $I_i^h \supseteq (\overline{x}_i, \overline{x}_{i+1})$, and the boundary points of I_h^i belong to the mesh. Let us now consider the minimum problem

$$\min\left\{F_h(w, I_i^h) : w \in PL_h(I_i^h),\ w(\overline{x}_i) = u_h(\overline{x}_i),\ w(\overline{x}_{i+1}) = u_h(\overline{x}_{i+1})\right\} \tag{4.9}$$

where we have set

$$F_h(w, I_i^h) := h \sum_{j=\frac{x_i^h}{h}+1}^{\frac{x_{i+1}^h}{h}+1} \phi\left((D_h^- w)_j\right)$$

the localization of the functional F_h on the inteval I_i^h. Let w_h be a solution of (4.9). Then, by Corollary 3.3, we have $\kappa(w_h) \in \{0, 1\}$. From Remark 3.11 it follows that for $h > 0$ small enough, necessarily

$$\kappa(w_h) = 1.$$

We now estimate the slope $\alpha^* = \alpha^*(w_h)$ of w_h in its unstable region (which consists of one subinterval of the mesh). Applying formula (2.15) for $\kappa = 1$, when I is replaced by I_i^h and $\mathcal{H}$ is replaced by $w_h(x_{i+1}^h + h) - w_h(x_i^h)$, it follows

$$h(\alpha^*)^2 - \left(w_h(x_{i+1}^h + h) - w_h(x_i^h)\right)\alpha^* + (x_{i+1}^h - x_i^h) = 0. \tag{4.10}$$

Hence

$$\begin{aligned}\alpha^* = \quad & \frac{w_h(x_{i+1}^h + h) - w_h(x_i^h)}{h} + o\left(\frac{1}{h}\right) = \frac{u_h(\overline{x}_{i+1}) - u_h(\overline{x}_i) + o(1)}{h} + o\left(\frac{1}{h}\right) \\ = \quad & \frac{u_h(\overline{x}_{i+1}) - u_h(\overline{x}_i)}{h} + o\left(\frac{1}{h}\right).\end{aligned}$$

We now estimate

$$F_h(w_h, I_i^h) \geq F_h(w_h, \text{unstab}(w_h)) = h\phi(\alpha^*) = h\phi\left(\frac{u_h(\overline{x}_{i+1}) - u_h(\overline{x}_i)}{h}\right) + o(h|\log h|).$$

We then have

$$\begin{aligned}\frac{F_h(u_h)}{h\phi(1/h)} \quad & \geq \quad \frac{1}{h\phi(1/h)} \sum_{i=1}^n F_h(u_h, I_i^h) + o(1) \geq \frac{1}{h\phi(1/h)} \sum_{i=1}^n F_h(w_h, I_i^h) + o(1) \\ & \geq \quad \frac{1}{\phi(1/h)} \sum_{i=1}^n \phi\left(\frac{u_h(\overline{x}_{i+1}) - u_h(\overline{x}_i)}{h}\right) + o(1) \geq n + o(1).\end{aligned}$$

Passing to the limit as $h \to 0$ we obtain inequality (4.6) and $u \in PC(I)$.
Fix now $\overline{u} \in PC(I)$. It remains to prove that

$$F_{\overline{u}}^{(2)} = \Gamma - \lim_{h \to 0^+} \frac{F_h - h\phi(1/h)F^{(1)}(\overline{u})}{h}. \tag{4.11}$$

Let $u \in PC(I)$. The Γ-limsup inequality can be proved choosing the sequence $\{u_h\} \subset PL_h(I)$ as in

(4.4) (performing the construction for any jump point of u). Indeed, from (4.5) we get

$$\lim_{h\to 0^+}\left(\frac{F_h(u_h)}{h}-\phi(1/h)F^{(1)}(\overline{u})\right)=\lim_{h\to 0^+}\left(\sum_{i=1}^{\ell/h}\phi\left((D_h^- u_h)_i\right)-\phi(1/h)\#J_{\overline{u}}\right)$$

$$=\lim_{h\to 0^+}\left(\sum_{i=1}^{\ell/h}\log\left|\frac{(u_h)_i-(u_h)_{i-1}}{h}\right|+\log h\#J_{\overline{u}}\right)$$

$$=\lim_{h\to 0^+}\left(\sum_{x\in J_u}\log\left|\frac{(u_h)_{\lfloor\frac{x}{h}\rfloor+1}-(u_h)_{\lfloor\frac{x}{h}\rfloor}}{h}\right|+\log h\#J_{\overline{u}}\right)$$

$$=\sum_{x\in J_u}\log|u(x_+)-u(x_-)|-\lim_{h\to 0^+}\log h\,(\#J_u-\#J_{\overline{u}})=F^{(2)}_{\overline{u}}(u).$$

The Γ-liminf inequality can be also obtained reasoning as before. Let $u\in L^1(I)$ and let $\{u_h\}\subset PL_h(I)$ be a sequence converging to u in $L^1(I)$ and almost everywhere, and let $\overline{x}_1,\dots\overline{x}_{n+1}$ be as in (4.7). We have

$$\lim_{h\to 0^+}\left(\frac{F_h(u_h)}{h}-\phi(1/h)\#J_{\overline{u}}\right)$$

$$\geq\lim_{h\to 0^+}\left(\sum_{i=1}^{n}\phi\left(\frac{u_h(\overline{x}_{i+1})-u_h(\overline{x}_i)}{h}\right)+o(|\log h|)+\log h\,\#J_{\overline{u}}\right)$$

$$\geq\sum_{i=1}^{n}\log|u(\overline{x}_{i+1})-u(\overline{x}_i)|+\lim_{h\to 0^+}\log h\left(\#J_{\overline{u}}-n+\frac{o(|\log h|)}{\log h}\right).$$

This implies, in particular, that the Γ-limit is $+\infty$ if $u\notin PC(I)$. Moreover, the expression (4.2) follows from the fact that, if $u\in PC(I)$, we can choose $n=\#J_u$. □

Remark 4.5. For all $\overline{u}\in PC(I)$, we can consider the functionals F_h in a suitable ball of $L^2(I)$ (or of $L^1(I)$), centered at $\overline{u}$, whose radius is chosen in such a way that $\overline{u}$ is a global minimizer for $F^{(1)}$ in such a ball. In this case, we have a Γ-expansion for F_h, in the sense of [6], of the type

$$F_h=h\phi\left(\frac{1}{h}\right)F^{(1)}+hF^{(2)}+o\left(h\right),$$

where the functional

$$F^{(2)}(u)=\begin{cases}\sum_{x\in J_u}\log|u(x_+)-u(x_-)| & u\in PC(I)\\ +\infty & u\in L^1(I)\setminus PC(I),\end{cases}\tag{4.12}$$

does not depend on the choice of $\overline{u}\in PC(I)$.
We observe that the restriction of F_h to a small ball of $L^2(I)$ is a natural assumption, when considering the gradient flow of F_h for small times.

References

[1] L. Alvarez, F. Guichard, P.L. Lions, J.M. Morel, *Axioms and fundamental equations of image processing*, Arch. Rational Mech. Anal. **123** (1993), 199–257.

[2] G. Bellettini, G. Fusco: *T*he Γ-limit and the related gradient flow for singular perturbation functionals of Perona-Malik type, Transactions Amer. Math. Soc., to appear.

[3] G. Bellettini, G. Fusco, N. Guglielmi, *A concept of solution for forward-backward equations of the form* $u_t = \frac{1}{2}(\phi'(u_x))_x$ *and numerical experiments for the singular perturbation* $u_t = -\epsilon^2 u_{xxxx} + \frac{1}{2}(\phi'(u_x))_x$, Discrete Contin. Dyn. Syst., **16** (4), 2006, 783-842.

[4] G. Bellettini, M. Novaga, E. Paolini, *Global solutions to the gradient flow equation of a nonconvex functional*, SIAM J. Math. Anal., **37** (5) (2006), 1657–1687.

[5] G. Bellettini, M. Novaga, M. Paolini, C. Tornese: Convergence of discrete schemes for the Perona-Malik equation, *preprint Centro De Giorgi*, 2006, submitted.

[6] A. Braides, L. Truskinovsky, *Asymptotic expansions by Γ-convergence*, preprint (2007).

[7] G. Dal Maso, An Introduction to Γ-convergence, *Birkhäuser, Boston*, (1993).

[8] E. De Giorgi, *Su alcuni problemi instabili legati alla teoria della visione*, in Atti del Convegno in onore di Carlo Ciliberto (Napoli, 1995), T. Bruno, P. Buonocore, L. Carbone, V. Esposito eds., 91–98, La Città del Sole, Napoli, 1997.

[9] S. Esedoglu, *An analysis of the Perona-Malik scheme*, Comm. Pure Appl. Math. **LIV** (2001), 1442–1487.

[10] S. Esedoglu, *Stability properties of the Perona-Malik scheme*, SIAM J. Numer. Anal. **44** (2006), 1297–1313.

[11] F. Fierro, R. Goglione, M. Paolini, *Numerical simulations of mean curvature flow in presence of a nonconvex anisotropy*, Math. Models Methods Appl. Sci. **8** (1998), 573–601.

[12] M. Ghisi, M. Gobbino, *Gradient estimates for the Perona-Malik equation*, Math. Ann., to appear.

[13] M. Ghisi, M. Gobbino, *A class of local classical solutions for the one-dimensional Perona-Malik equation*, CVGMT-preprint (2007).

[14] M. Gobbino, *Entire solutions to the one-dimensional Perona-Malik equation*, Comm. Partial Differential Equations, to appear.

[15] B. Kawhol, N. Kutev, *Maximum and comparison principle for the one-dimensional anisotropic diffusion* , Math. Ann., **311** (1998), 107-123.

[16] S. Kichenassamy, *The Perona-Malik paradox* , SIAM J. Appl. Math., **57** (1997), 1382-1342.

[17] E. Isaacson, H.B. Keller, Analysis of Numerical Methods, Dover, New York, 1994

[18] J.M. Morel and S. Solimini, Variational Methods in Image Segmentation, Vol. 14 of *Progress in Nonlinear Differential Equations and Their Applications*, Birkhäuser, Basel, 1995.

[19] P. Perona, M. Malik, *Scale space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell. **12** (1990), 629–639.

[20] M. Nitzberg, D. Mumford and T. Shiota, Filtering, Segmentation and Depth, *Lecture Notes in Computer Science* 662, Springer-Verlag, Berlin (1993).

[21] D.G. Schaeffer, M. Shearer, T.P Witelski, *A discrete model for an ill-posed non linear parabolic PDE*, Physica D **160** (2001), 189–221.

[22] S. Taheri, Q. Tang, K. Zhang, *Young measure solutions and instability of the one-dimensional Perona-Malik equation*, J. Math. Anal. Appl. **308** (2005), 467–490.

[23] K. Zhang, *Existence of infinitely many solutions for the one-dimensional Perona-Malik problem*, Calc. Var. Partial Differential Equations **326** (2006), 171–199.