CONVERGENCE OF A DAMPED NEWTON'S METHOD FOR DISCRETE MONGE-AMPÈRE FUNCTIONS WITH A PRESCRIBED ASYMPTOTIC CONE

GERARD AWANOU

ABSTRACT. We prove the convergence of a damped Newton's method for the nonlinear system resulting from a discretization of the second boundary value problem for the Monge-Ampère equation. The boundary condition is enforced through the use of the notion of asymptotic cone. The differential operator is discretized based on a discrete analogue of the subdifferential.

1. INTRODUCTION

In this paper we prove the convergence of a damped Newton's method for a finite difference approximation of the second boundary value problem for Monge-Ampère type equations introduced in [1]. Our result leads to the existence of a solution to the discretization. The damped Newton's method allows the use of an initial guess which may be far from the solution of the discrete problem. We establish the global convergence of the algorithm for the discretization proposed.

Monge-Ampère type equations with the second boundary value condition arise in geometric optics and optimal transport. The approach in [1] is to interpret the boundary condition as the prescription of the asymptotic cone of the epigraph of the convex solution of the Monge-Ampère equation. The convergence analysis of the damped Newton's method given here generalizes the ones given in [7, 5].

Let Ω and Ω^* be bounded convex polygonal domains of $\mathbb{R}^d, d \geq 1$ and let R be a locally integrable function on $\mathbb{R}^d$ such that $R > 0$ on Ω^* and $R = 0$ on $\mathbb{R}^d \setminus \Omega^*$. Let $f \geq 0$ be an integrable function on Ω. We assume that the compatibility condition

$$\int_\Omega f(x) dx = \int_{\Omega^*} R(p) dp, \tag{1.1}$$

holds. Recall that for a function u on Ω the subdifferential of u at $x \in \Omega$ is defined by

$$\partial u(x) = \{ \, p \in \mathbb{R}^d : u(y) \geq u(x) + p \cdot (y - x), \text{ for all } y \in \Omega \, \}. \tag{1.2}$$

We are interested in approximating a convex function u which solves

$$\begin{aligned} R(Du(x)) \det D^2 u(x) &= f(x) \text{ in } \Omega \\ \partial u(\Omega) &= \Omega^*. \end{aligned} \tag{1.3}$$

The first equation in (1.3) is to be interpreted in the sense of Aleksandrov, i.e. for a Borel set $E \subset \Omega$, it is required that $\omega(R, u, E) := \int_{\partial u(E)} R(p) dp = \int_E f(x) dx$, where

$\omega(R, u, .)$ is the R-Monge-Ampère measure associated to u. The second equation in (1.3) is the second boundary condition.

Let k_{Ω^*} denote the support function of Ω^*, i.e. for $x \in \mathbb{R}^d, k_{\Omega^*}(x) = \sup_{p \in \overline{\Omega^*}} p \cdot x$. It is shown in [1] that if one defines for $x \notin \Omega$

$$u(x) = \inf_{y \in \partial \Omega} u(y) + k_{\Omega^*}(x-y), \tag{1.4}$$

one obtains a convex extension of u to $\mathbb{R}^d$. Moreover, (1.3) is equivalent to finding a convex function u on $\mathbb{R}^d$ solving (1.4) such that in the sense of Aleksandrov

$$R(Du(x)) \det D^2 u(x) = f(x) \text{ in } \mathbb{R}^d. \tag{1.5}$$

We refer to solutions of (1.5) as Monge-Ampère functions, anticipating applications to a more general setting [4].

Solutions of (1.5) are unique up to an additive constant. For $x^1 \in \Omega$, we may require that $u(x^1) = 0$. In view of the compatibility condition (1.1), (1.5) is formally equivalent to finding a convex function u on $\mathbb{R}^d$ solving (1.4) such that in the sense of Aleksandrov

$$R(Du(x)) \det D^2 u(x) = f(x) + w\, u(x^1) \text{ in } \mathbb{R}^d, \tag{1.6}$$

where w is a constant which we assume to be non positive. For approximating solutions of (1.6), it is proposed in [2] to use piecewise linear convex functions. However, if v is a piecewise linear convex function, $\partial v(\Omega)$ is a polygon [1, Lemma 10]. We also approximate the domain Ω^* by a polygon $Y \subset \Omega^*$.

Let h be a small positive parameter and let $\mathbb{Z}^d_h = a + \{ mh, m \in \mathbb{Z}^d \}$ denote the orthogonal lattice with mesh length h and with an offset $a \in \mathbb{Z}^d$. The offset a may make it easier to choose the decomposition of the domain used for the discrete Monge-Ampère equation (1.8) below. Put $\Omega_h = \Omega \cap \mathbb{Z}^d_h$ and denote by $(r_1, \ldots, r_d)$ the canonical basis of $\mathbb{R}^d$. Let

$$\partial \Omega_h = \{ x \in \Omega_h \text{ such that for some } i = 1, \ldots, d, x + hr_i \notin \Omega_h \text{ or } x - hr_i \notin \Omega_h \}.$$

We note that with our notation $\partial \Omega_h \subset \Omega_h$. For $x \in \Omega_h$, let $V(x)$ be a finite subset of $\mathbb{Z}^d \setminus \{0\}$. We will refer to functions defined on $\mathbb{Z}^d_h$ as mesh functions.

We consider the following discrete analogue of the subdifferential of a function. For $x \in \mathbb{Z}^d_h$ we define for a mesh function v_h

$$\partial_V v_h(x) = \{ p \in \mathbb{R}^d, p \cdot (he) \geq v_h(x) - v_h(x - he) \, \forall e \in V(x) \},$$

and consider the following discrete version of the R-Monge-Ampère measure

$$\omega_V(R, v_h, E) := \int_{\partial_V v_h(E)} R(p) dp.$$

We are interested in mesh functions v_h which are V-discrete convex in the sense that for all $x \in \Omega_h$ and $e \in V(x)$

$$\Delta_{he} v_h(x) := v_h(x + he) - 2 v_h(x) + v_h(x - he) \geq 0.$$

We also require a discrete analogue of (1.7), namely, for $x \in \mathbb{Z}^d_h \setminus \Omega_h$

$$v_h(x) = \min_{y \in \partial \Omega_h} v_h(y) + k_Y(x-y), \tag{1.7}$$

and we recall that $Y \subset \Omega^*$ is a polygonal domain approximating Ω^*.

We denote by $\mathcal{C}_h^{Y,V}$ the set of V-discrete convex mesh functions which satisfy (1.7). We can now describe the discretization of the second boundary value problem we consider in this paper: find $u_h \in \mathcal{C}_h^{Y,V}$ such that

$$\omega_V(R, u_h, \{ x \}) = \int_{C_x} f(t) dt + w\, u_h(x^1), x \in \Omega_h, \tag{1.8}$$

where $x^1 \in \Omega_h$ and $(C_x)_{x \in \Omega_h}$ form a partition of Ω, i.e. $C_x \cap \Omega_h = \{ x \}, \cup_{x \in \Omega_h} C_x = \Omega$, and $C_x \cap C_y$ is a set of measure 0 for $x \neq y$. In the interior of Ω one may choose as $C_x = x + [-h/2, h/2]^d$ the cube centered at x with $C_x \cap \Omega_h = \{ x \}$. We make an abuse of notation by not making explicit the dependence of x^1 on h. The requirement that the sets C_x form a partition is essential for the discretization of the measure with density f.

The unknowns in (1.8) are the mesh values $u_h(x), x \in \Omega_h$. For $z \notin \Omega_h$, the value $u_h(z)$ needed for the evaluation of $\partial_V v_h(x)$ is obtained from the discrete extension formula (1.7).

We show that a subsequence of the damped Newton iterations converges to a solution of (1.8) for $w < 0$. This gives existence of a solution to (1.8). Taking limits of solutions to (1.8) as $w \to 0$, we obtain existence of a solution to the problem: find $u_h \in \mathcal{C}_h^{Y,V}$ such that

$$\omega_V(R, u_h, \{ x \}) = \int_{C_x} f(t) dt, x \in \Omega_h. \tag{1.9}$$

While adding a constant to a solution of (1.9) results in another solution, this is not the case for (1.8).

We organize the paper as follows. In the next section, we give additional preliminaries. The damped Newton's method is introduced in section 3 in a general setting. In section 4 we give its convergence analysis for (1.8) and discuss the extension to (1.9).

2. Preliminaries

We will use the notation C for a generic constant and $|| \cdot ||$ for the Euclidean norm. We first describe the extended mesh needed for the evaluation of $\partial_V v_h(x)$ in (1.8). A stencil V is a set valued mapping from Ω_h to the set of finite subsets of $\mathbb{Z}^d \setminus \{0\}$. Recall that a subset W of $\mathbb{Z}^d$ is symmetric with respect to the origin if $\forall y \in W, -y \in W$. We define V_{min} to be a finite subset of $\mathbb{Z}^d \setminus \{0\}$ which is symmetric with respect to the origin, contains the elements of the canonical basis of $\mathbb{R}^d$, and contains a vector parallel to a normal to each facet of the polygonal domain Y. Put

$$\Omega_{ext} = \Omega_h \cup \{ x + he : x \in \Omega_h, e \in V_{min} \}.$$
$$V_{max}(x) = \{ e \in \mathbb{Z}^d \setminus \{0\}, \exists y \in \Omega_{ext}, y = x + he \}.$$

We assume that

$$V_{min} \subset V(x) \subset V_{max}(x), x \in \Omega_h,$$

and $V(x) \subset \mathbb{Z}^d \setminus \{ 0 \}$ is symmetric with respect to the origin for $x \in \Omega_h$. We furthermore assume that if e and f are in $V(x)$, $e = rf$ for a scalar r if and only if $r = -1$.

We will need below the following results from [1].

Lemma 2.1. [1, Theorem 4] *There exists a unique solution to* (1.8) *for* $V = V_{max}$ *and for which* $u_h(x^1) = 0$.

The above theorem asserts the existence and uniqueness of a solution to (1.8) when the stencil V is maximal. We prove below that a solution exists when V is not necessarily equal to V_{max}.

Lemma 2.2. [1, Lemma 2] *There exists a constant* C *independent of* $v_h \in \mathcal{C}_h^{Y,V}$ *and* h *such that for all* $x, y \in \Omega_h$

$$|v_h(x) - v_h(y)| \leq C||x - y||.$$

Using $V_{\min} \subset V(x), x \in \Omega_h$, the next lemma follows from [1, Lemma 2].

Lemma 2.3. *Let* $v_h \in \mathcal{C}_h^{Y,V}$. *Then for* $x \in \Omega_h$, $\partial_V v_h(x) \subset Y \subset \Omega^*$.

3. The damped Newton's method

We first give a general convergence result, the assumptions of which are then verified in the next section for our discretization. Let $\mathbb{U}$ denote an open subset of $\mathbb{R}^M$. In section 4 we will identify a mesh function with a vector consisting in its values at grid points. Here, we then denote by v_h a generic element of $\mathbb{U}$.

We are interested in the zeros of a mapping $G : \mathbb{U} \to \mathbb{R}^M$ with $G(v_h) = (G_l(v_h))_{l=1,\dots,M}$. We recall that $||.||$ denotes the Euclidean norm in $\mathbb{R}^M$ and put

$$\mathbb{U}_\epsilon = \{\, v_h \in \mathbb{U}, G_l(v_h) > -\epsilon, l = 1, \dots, M \,\},$$

for a parameter $\epsilon > 0$. This allows us to study the differentiability of G on an open set. We assume that $G \in C^1(\mathbb{U}_\epsilon, \mathbb{R}^M)$ for all $\epsilon > 0$. The current iterate of the damped Newton's method is denoted v_h^k and the following iterate is sought along the path

$$p^k(\tau) = v_h^k - \tau\big(G'(v_h^k) - \nu I\big)^{-1} G(v_h^k),$$

where ν is a non negative constant and I the $M \times M$ identity matrix. The term $-\nu I$ is introduced to deal with situations where $G'(v_h^k)$ is not known to be invertible but $G'(v_h^k) - \nu I$ can be shown to be invertible. When the term $-\nu I$ is used in the computations, it is desirable to choose ν small, as suggested by the rate in Theorem 3.1 below. For the general setting considered in this section, we make the assumption that when $v_h^k \in \mathbb{U}$, there exists $\tau_k \in (0, 1]$ such that $p^k(\tau) \in \mathbb{U}$ for all $\tau \in [0, \tau_k]$. Let $\delta \in (0, 1)$ and choose $\rho \in (0, 1)$, e.g. $\rho = 1/2$.

Algorithm 1 A damped Newton's method

1: Choose $v_h^0 \in \mathbb{U}_\epsilon$ and set $k = 0$
2: **If** $G(v_h^k) = 0$ **stop**
3: Let i_k be the smallest non-negative integer i such that $p^k(\rho^i) \in \mathbb{U}_\epsilon$ and

$$||G(p^k(\rho^i))|| \leq (1 - \delta\rho^i)||G(v_h^k)||.$$

Set $v_h^{k+1} = p^k(\rho^{i_k})$
4: $k \leftarrow k + 1$ and **go to** 2

The general convergence result for damped Newton's methods is analogous to [5, Proposition 6.1] where maps with probability measures as values are considered. Therein, the map G is assumed to be in $C^{1,\alpha}, 0 < \alpha \leq 1$. For G to be merely C^1, as in certain geometric optics problems, and with $\det G'(x) \neq 0$ for all $x \in \mathbb{U}$, one has convergence [3]. For completeness we adapt the proof of [3] to the case where the domain of the mapping G is an open set of $\mathbb{R}^M$. As with [7, Proposition 2.10] we will assume that the mapping G is proper, i.e. the preimage of any compact set is a compact set.

Theorem 3.1. *Let $G \in C^1(\mathbb{U}, \mathbb{R}^M)$ be a proper map with $\det(G'(x) - \nu I) \neq 0$ for some $\nu \geq 0$ and for all $x \in \mathbb{U}$. Assume that when $v_h^k \in \mathbb{U}$, there exists τ_k in $(0,1]$ such that for all $0 < \tau \leq \tau_k, p_k(\tau) \in \mathbb{U}$. With an initial guess v_h^0 in $\mathbb{U}_\epsilon$, the iterates v_h^k of the damped Newton's method are well defined and there is a subsequence k_l such that $v_h^{k_l}$ converges to a zero u_h of G. For $k_l \geq k_0$ and k_0 sufficiently large, if $i_{k_l} = 0$ in step 3 of Algorithm 1*

$$||v_h^{k_l+1} - u_h|| \leq \nu ||v_h^{k_l} - u_h|| + C||v_h^{k_l} - u_h|| \, ||\psi(v_h^{k_l} - u_h)||,$$

where $\psi : \mathbb{R}^M \to \mathbb{R}^M$ with $||\psi(v_h^{k_l} - u_h)|| \to 0$ when $||v_h^{k_l} - u_h|| \to 0$. If G has a unique zero u_h in $\mathbb{U}_\epsilon$, the whole sequence converges to u_h.

Proof. The proof is divided into three parts. In the first part, we show that given an iterate v_h^k in the admissible set $\mathbb{U}_\epsilon$, one can find a path from v_h^k which is contained in $\mathbb{U}_\epsilon$, provided v_h^k is not a zero of G. We show by a contradiction argument that G decreases in norm along such a path at the rate given in the algorithm. In the second part, we show that there exists a subsequence of the iterates converging to a zero of G. Finally in the third part, we give the convergence rate.

Part 1: The damped Newton's method is well defined. Assume that $v_h^k \in \mathbb{U}$. We first note that it follows from the definitions that for all $\tau \in [0,1]$ we have

$$G(v_h^k) + \big(G'(v_h^k) - \nu I\big)(p^k(\tau) - v_h^k) = (1-\tau)G(v_h^k). \tag{3.1}$$

Assume that $G(v_h^k) \neq 0$. We claim that there exists $\tau_k' \in (0,1]$ such that

$$||G(p^k(\tau))|| \leq (1 - \delta\tau)||G(v_h^k)||, \ \forall \tau \in [0, \tau_k'].$$

If such a τ_k' does not exist, there would exist a sequence τ_l converging to 0 such that

$$||G(p^k(\tau_l))|| > (1 - \delta\tau_l)||G(v_h^k)||, \ \forall k. \tag{3.2}$$

Since G is C^1 by assumption, we have

$$G(p^k(\tau_l)) = G(v_h^k) + G'(v_h^k)(p^k(\tau_l) - v_h^k) + ||p^k(\tau_l) - v_h^k||\psi(p^k(\tau_l) - v_h^k),$$

for a function $\psi : \mathbb{R}^M \to \mathbb{R}^M$ such that $||\psi(p^k(\tau_l) - v_h^k)|| \to 0$ as $||p^k(\tau_l) - v_h^k|| \to 0$ or equivalently $\tau_l \to 0$. Thus, by (3.1), we have

$$G(p^k(\tau_l)) = (1-\tau_l)G(v_h^k) + \nu||p^k(\tau_l) - v_h^k|| + ||p^k(\tau_l) - v_h^k||\psi(p^k(\tau_l) - v_h^k), \tag{3.3}$$

and thus by (3.2)

$$\begin{aligned}(1-\delta\tau_l)||G(v_h^k)|| < ||G(p^k(\tau_l))|| \leq (1-\tau_l)||G(v_h^k)|| + \nu||p^k(\tau_l) - v_h^k|| \\ + ||p^k(\tau_l) - v_h^k|| \, ||\psi(p^k(\tau_l) - v_h^k)||.\end{aligned}$$

This implies that $\tau_l(1-\delta)||G(v_h^k)|| < \nu||p^k(\tau_l)-v_h^k|| + ||p^k(\tau_l)-v_h^k||\,||\psi(p^k(\tau_l)-v_h^k)||$ which gives using the definition of $p^k(\tau_l)$

$$(1-\delta)||G(v_h^k)|| < \nu||p^k(\tau_l)-v_h^k|| + ||G'(v_h^k)^{-1}G(v_h^k)||\,||\psi(p^k(\tau_l)-v_h^k)||.$$

Taking the limit as $||p^k(\tau_l)-v_h^k|| \to 0$, i.e. $\tau_l \to 0$, we obtain $(1-\delta)||G(v_h^k)|| \leq 0$, a contradiction as $G(v_h^k) \neq 0$ by assumption.

We now show that there exists $\overline{\tau}_k \leq \min\{\,\tau_k', \tau_k\,\}$, with $\overline{\tau}_k > 0$ such that for all $\tau \in [0, \overline{\tau}_k], p^k(\tau) \in \mathbb{U}_\epsilon$ when $v_h^k \in \mathbb{U}_\epsilon$.

If $p^k(\tau) \notin \mathbb{U}_\epsilon$ for some $\tau \leq \min\{\,\tau_k', \tau_k\,\}$ it must be that at some time $\overline{\tau}_k \leq \tau, p^k(\overline{\tau}_k) \in \partial\mathbb{U}_\epsilon$. Thus for some $i_k \in \{\,1,\ldots,M\,\}$ we have $G_{i_k}(p(\overline{\tau}_k)) = -\epsilon$. Let us assume that $\overline{\tau}_k$ is chosen so that $G_i(p^k(\tau)) > -\epsilon$ for all $\tau \in [0, \overline{\tau}_k)$ and all $i \in \{\,1,\ldots,M\,\}$.

We have

$$||G(p^k(\overline{\tau}_k)) - G(v_h^k)|| \geq |G_{i_k}(p^k(\overline{\tau}_k)) - G_{i_k}(v_h^k)| = |-\epsilon - G_{i_k}(v_h^k)| > 0, \tag{3.4}$$

since $G_{i_k}(v_h^k) > -\epsilon$ as $v_h^k = p^k(0) \in \mathbb{U}_\epsilon$ by assumption. We conclude that $\overline{\tau}_k > 0$. Otherwise $p^k(\overline{\tau}_k)) = p^k(0) = v_h^k$ which contradicts (3.4) as the first term on the left in (3.4) is equal to 0. By construction of $\overline{\tau}_k$, $p^k(\tau) \in \mathbb{U}_\epsilon$ for all $\tau \in [0, \overline{\tau}_k)$.

Part 2: The sequence v_h^k has a subsequence converging to a zero u_h of G. Let

$$\mathcal{K} = \{\, v \in \mathbb{U}_\epsilon, ||G(v)|| \leq ||G(v_h^0)||\,\}.$$

Since G is proper, $\mathcal{K}$ is compact. By construction $||G(v_h^{k+1})|| < ||G(v_h^k)||$ and $v_h^k \in \mathbb{U}_\epsilon$ for all k. We therefore have $v_h^k \in \mathcal{K}$ for all k. Thus up to a subsequence, the sequence v_h^k converges to an element $u_h \in \overline{\mathbb{U}_\epsilon}$ and since the sequence $||G(v_h^k)||$ is strictly decreasing, there exists $\eta \geq 0$ such that $||G(v_h^k)|| \to \eta$.

If $\eta = 0$, then $G(u_h) = 0$ and the subsequence converges to a zero of G. Now we show that it is not possible to have $\eta \neq 0$.

Assume that $\eta > 0$. If the sequence i_k of Step 3 of the algorithm were bounded, there would exist a constant $\xi > 0$ such that $\rho^{i_k} \geq \xi$ for all k. This would imply that $1-\delta\rho^{i_k} \leq 1-\delta\xi$ and thus $||G(v_h^{k+1})|| \leq (1-\delta\rho^{i_k})||G(v_h^k)|| \leq (1-\delta\xi)||G(v_h^k)||$. It follows that $\eta \leq (1-\delta\xi)\eta$. Thus $\eta = 0$ which is a contradiction. We therefore have $\lim_{k\to\infty} i_k = \infty$ and consequently $\lim_{k\to\infty} \rho^{i_k-1} = 0$. Put

$$\hat{\tau}_k = \rho^{i_k-1}.$$

By definition of i_k and (3.3), we have

$$\begin{aligned}(1-\delta\hat{\tau}_k)||G(v_h^k)|| < ||G(p(\hat{\tau}_k))|| &\leq (1-\hat{\tau}_k)||G(v_h^k)|| + \nu||p^k(\hat{\tau}_k)-v_h^k|| \\ &\quad + ||p^k(\hat{\tau}_k)-v_h^k||\,||\psi(p^k(\hat{\tau}_k)-v_h^k)||,\end{aligned}$$

and thus

$$(1-\delta)||G(v_h^k)|| < \nu||p^k(\hat{\tau}_k)-v_h^k|| + ||p^k(\hat{\tau}_k)-v_h^k||\,||\psi(p^k(\hat{\tau}_k)-v_h^k)||.$$

Recall that the sequence $||G(v_h^k)||$ is bounded. Moreover $u_h \in \mathbb{U}_\epsilon$ and thus $\det\big(G'(u_h)-\nu I\big) \neq 0$. Since $G \in C^1(\mathbb{U}_\epsilon, \mathbb{R}^M)$ it follows that for k sufficiently large $||\big(G'(v_h^k) -$

$\nu I\big)^{-1}|| \leq C$ for a constant C, c.f. [8, \$ 2.3.3]. Therefore $||p^k(\hat{\tau}_k) - v_h^k|| \leq C\hat{\tau}_k$ for a constant C. We conclude that

$$(1-\delta)||G(v_h^k)|| < C\nu\hat{\tau}_k + C\hat{\tau}_k||\psi(p^k(\hat{\tau}_k) - v_h^k)||.$$

Since $\hat{\tau}_k \to 0$, we obtain $\eta = 0$, a contradiction.

Part 3: We first prove that for the subsequence k_l obtained in Part 2, we have the rate $||v_h^{k_l+1} - u_h|| \leq \nu||v_h^{k_l} - u_h|| + C||v_h^{k_l} - u_h||\, ||\psi(v_h^{k_l} - u_h)||$, when $i_k = 0$, for $k_l \geq k_0$ and k_0 sufficiently large, and where $||\psi(v_h^{k_l} - u_h)|| \to 0$ as $v_h^{k_l} - u_h \to 0$.

For convenience, we denote the subsequence $v_h^{k_l}$ by v_h^k below. We have with $\tilde{\tau}_k = \rho^{i_k}$, $v_h^{k+1} = v_h^k - \tilde{\tau}_k\big(G'(v_h^k) - \nu I\big)^{-1}G(v_h^k)$ and

$$\begin{aligned}\big(G'(v_h^k) - \nu I\big)(v_h^{k+1} - u_h) &= \big(G'(v_h^k) - \nu I\big)\big(v_h^k - u_h - \tilde{\tau}_k\big(G'(v_h^k) - \nu I\big)^{-1}G(v_h^k)\big) \\ &= \big(G'(v_h^k) - \nu I\big)(v_h^k - u_h) - \tilde{\tau}_k G(v_h^k).\end{aligned}$$

Moreover $0 = G(u_h) = G(v_h^k) - G'(v_h^k)(v_h^k - u_h) + ||v_h^k - u_h||\psi(v_h^k - u_h)$. Thus $\big(G'(v_h^k) - \nu I\big)(v_h^{k+1} - u_h) = (1-\tilde{\tau}_k)G(v_h^k) - \nu(v_h^k - u_h) + ||v_h^k - u_h||\psi(v_h^k - u_h)$. If $i_k = 0$, we have $\tilde{\tau}_k = 1$ and in that case $\big(G'(v_h^k) - \nu I\big)(v_h^{k+1} - u_h) = -\nu(v_h^k - u_h) + ||v_h^k - u_h||\psi(v_h^k - u_h)$. Since $\big(G'(v_h^k) - \nu I\big)^{-1}$ is uniformly bounded for k sufficiently large, the claim follows when $\tilde{\tau}_k = 1$.

The proof of the last statement is standard. □

4. Convergence of the damped Newton's method for the discretization

Let M denote the cardinality of Ω_h and denote the points of Ω_h by $x^i, i = 1, \ldots, M$. The set of mesh functions on Ω_h is identified with $\mathbb{R}^M$. We note that mesh functions on Ω_h are naturally extended to $\mathbb{Z}_h^d$ using (1.7). Mesh functions which are extended to $\mathbb{Z}_h^d$ using (1.7) and for which $\Delta_{he} v_h(x) > 0$ for all $x \in \Omega_h$ and $e \in V(x)$, form the subset of $\mathcal{C}_h^{Y,V}$ of strictly V-discrete convex mesh functions. The latter subset is identified with an open subset $\mathbb{U}$ of $\mathbb{R}^M$ by mapping v_h to $(v_h(x_i))_{i=1,\ldots,M}$.

With an abuse of notation, we consider a map $G : \mathbb{U} \to \mathbb{R}^M$ defined by

$$G_i(v_h) = \omega_V(R, v_h, \{\, x^i \,\}) - \int_{C_{x^i}} f(t)dt - w\, v_h(x^1) \text{ for } i = 1, \ldots, M, \tag{4.1}$$

and we recall that $w < 0$ is a constant. We now make the assumption that $f > 0$ in Ω. Define as in the previous section for $\epsilon > 0$, $\mathbb{U}_\epsilon = \{\, v_h \in \mathbb{U}, G_i(v_h) > -\epsilon, i = 1, \ldots, M \,\}$.

Lemma 4.1. *The set $\mathbb{U}_\epsilon$ is non empty.*

Proof. Let u_h^0 denote the solution of (1.8) for $V = V_{max}$ and for which $u_h(x^1) = 0$, c.f. Lemma 2.1. We have for $i = 1, \ldots, M$,

$$G_i(u_h^0) = \omega_V(R, u_h^0, \{\, x^i \,\}) - \int_{C_{x^i}} f(t)dt \geq \omega_{V_{max}}(R, u_h^0, \{\, x^i \,\}) - \int_{C_{x^i}} f(t)dt = 0 > -\epsilon.$$

This completes the proof. □

The goal of this section is to verify the assumptions of Theorem 3.1 for the equation $G(v_h) = 0$ with G given by (4.1). We will prove the following theorem

Theorem 4.2. *Assume that $w < 0$. Given an initial guess $v_h^0 \in \mathbb{U}_\epsilon$, the iterate v_h^k of the damped Newton's method is well defined, and there is a subsequence $v_h^{k_l}$ which converges to a solution u_h of* (1.8).

The proof of Theorem 4.2 proceeds in several steps. The path condition in Theorem 3.1 is proved in Theorem 4.3 under the assumption that the mapping G is differentiable with $G'(v) - \nu I$ invertible, for $\nu \geq 0$ and I the identity matrix. The C^1 continuity of G is established in Theorem 4.6 and the invertibility assumption in Theorem 4.8 under the assumption $\nu > -w$. Finally, in Theorem 4.9, we prove that the mapping G is proper for $w < 0$. We are then in a position to give the proof of Theorem 4.2.

Theorem 4.3. *Assume that G is differentiable at $v \in \mathbb{U}$ and $G'(v) - \nu I$ is invertible for $\nu \geq 0$ and I the identity matrix. There exists $\check{\tau}$ in $(0, 1]$ such that for all $0 < \tau \leq \check{\tau}, p(\tau) = v - \tau\big(G'(v) - \nu I\big)^{-1} G(v) \in \mathbb{U}$.*

Proof. Since $v \in \mathbb{U}$, $\Delta_{he} v(x^i) > 0$ for all $i = 1, \ldots, M$ and $e \in V(x)$. Let $c_0 = \min\{\, \Delta_{he} v_h(x^i), i = 1, \ldots, M, e \in V(x) \,\}$ and let $\zeta = c_0/8$.

We claim that if $|w - v|_\infty := \max\{\, |w(x) - v(x)|, x \in \Omega_h \,\} \leq \zeta$, then $\Delta_{he} w(x^i) > 0, i = 1, \ldots, M, e \in V(x)$.

Indeed, for $x \in \Omega_h$, when $x \pm e \in \Omega_h$

$$
\begin{aligned}
\Delta_{he} w(x) &= w(x + he) - 2w(x) + w(x - he) \\
&= \Delta_{he} v(x) + \big((w - v)(x + he) - 2(w - v)(x) + (w - v)(x - he)\big) \\
&\geq \Delta_{he} v(x) - 4\zeta \geq c_0 - 4\zeta = \frac{c_0}{2} > 0,
\end{aligned}
$$

where we used $|w - v|_\infty \leq \zeta$.

If $x + e \notin \Omega_h$, using (1.7), we have $w(x + e) = w(y) + (x + e - y) \cdot q$ for some $q \in Y$ and $y \in \partial\Omega_h$. We have $v(x + e) \leq v(y) + (x + e - y) \cdot q$ and thus

$$
w(x + e) - v(x + e) \geq w(y) - v(y) \geq -\zeta.
$$

Similarly, when $x - e \notin \Omega_h$, we have $w(x - e) - v(x - e) \geq -\zeta$. Thus, arguing as above, for all $e \in V(x)$, $\Delta_{he} w(x) > c_0/2 > 0$, which proves the claim.

Since $|w - v|_\infty \leq C||w - v||$ and

$$
||p(\tau) - v|| \leq \tau||\big(G'(v) - \nu I\big)^{-1}||\, ||G(v)||,
$$

we have

$$
|p(\tau) - v|_\infty \leq C||p(\tau) - v|| \leq C\tau||\big(G'(v) - \nu I\big)^{-1}||\, ||G(v)|| \leq \zeta,
$$

for τ sufficiently small. Here, since v is fixed, $||\big(G'(v) - \nu I\big)^{-1}||\, ||G(v)||$ is a constant independent of τ. It follows that $p(\tau) \in \mathbb{U}$ for τ sufficiently small. The proof is complete. □

Let $\#W$ denote the cardinality of the set W. Given $\lambda \in \mathbb{R}^{\#W}$, we write $\lambda = (\lambda_a)_{a\in W}$ by an abuse of notation, instead of the more familiar notation $\lambda = (\lambda_j)_{j=1,\dots,\#W}$.

We fix i in $\{ 1, \dots, M \}$ and for $\lambda \in \mathbb{R}^{\#V(x^i)}$, let

$$Q(\lambda) = \{ p \in \mathbb{R}^d, p \cdot e \leq \lambda_e, \ \forall e \in V(x^i) \},$$

where for simplicity we do not mention the dependence of $Q(\lambda)$ on the index i.

Consider the mapping S defined by $S(\lambda) := \int_{Q(\lambda)} R(p)dp$. For a given index j we are interested in the variations of $G_i(v)$ with respect to $v(x^j)$, i.e. the derivative at $v(x^j)$ of the application which is the composite of S and the mapping

$$r \mapsto \lambda = \left(\frac{w(x^i + ha) - w(x^i)}{h} \right)_{a \in V(x^i)},$$

where w is the mesh function defined by

$$w(x) = v(x), x \neq x^j, \quad w(x^j) = r.$$

Let $z_i, i = 1, \dots, N \in \mathbb{R}^d$. Let $(r_1, \dots, r_N)$ denote the canonical basis of $\mathbb{R}^N$. Given $\lambda \in \mathbb{R}^N$, define

$$\tilde{Q}(\lambda) := \{ x \in \mathbb{R}^d, x \cdot z_i \leq \lambda_i, i = 1, \dots, N \}.$$

Assume that the vectors z_i are chosen such that the polytope $\hat{Q}(\lambda)$ is bounded. We will need the following lemma [6, Lemma 16].

Lemma 4.4. *Let $\rho : \mathbb{R}^d \to \mathbb{R}$ be a continuous function. Define*

$$\tilde{S}(\lambda) := \int_{\tilde{Q}(\lambda)} \rho(p)dp$$

(1) *If $z_i \neq 0$ for all i, $\tilde{S}$ is continuous on $\mathbb{R}^d$.*
(2) *$\forall R \geq 0$, $\exists Q_R \subset \mathbb{R}^d$ compact such that $\forall \lambda' \in \mathbb{R}^N$, $\max_i |\lambda' - \lambda| \leq R$ implies $\tilde{Q}(\lambda') \subseteq Q_R$.*
(3) *There exists a function $\eta_R : \mathbb{R}^+ \to \mathbb{R}^+$ satisfying $\lim_{s\to 0} \eta_R(s) = 0$ such that for all $x, y \in Q_R$, $|\rho(x) - \rho(y)| \leq \eta_R(||x - y||)$.*
(4) *Let $i_0 \in \{ 1, \dots, N \}, \lambda \in R^N$ and $t \geq 0$. We have*

$$\frac{1}{t}(\tilde{S}(\lambda + tr_{i_0}) - \tilde{S}(\lambda)) = \frac{1}{t} \int_0^t g_{i_0}(\lambda + sr_{i_0})ds,$$

where $g_{i_0}(\overline{\lambda}) := 1/||z_{i_0}|| \int_{\tilde{Q}(\overline{\lambda}) \cap \{ x \in \mathbb{R}^d, x \cdot z_{i_0} = \overline{\lambda}_{i_0} \}} \rho(p)dp$.
(5) *Let Π_{i_0} denote the orthogonal projection onto the hyperplane orthogonal to z_{i_0}. Put $P_{i_0}(\lambda) = \Pi\big(\tilde{Q}(\lambda) \cap \{ x \in \mathbb{R}^d, x \cdot z_{i_0} = \lambda_{i_0} \}\big)$. We have*

$$P_{i_0}(\lambda) = \{ y \in \mathbb{R}^d, y \cdot z_{i_0} = 0 \ \text{and} \ y \cdot \Pi_{i_0}(z_i) \leq \lambda_i - \lambda_{i_0} z_i \cdot z_{i_0}/||z_{i_0}||^2, i \neq i_0 \}.$$

(6) *Define $\rho_\lambda(y) = \rho(y + \lambda_{i_0} z_{i_0}/||z_{i_0}||^2)$. We have $||z_{i_0}|| \, |g_{i_0}(\lambda) - g_{i_0}(\lambda')| \leq |A_1| + |A_2|$, where*

$$A_1 := \int_{P_{i_0}(\lambda)} (\rho_\lambda(y) - \rho_{\lambda'}(y))dy$$

$$A_2 := \int_{P_{i_0}(\lambda)} \rho_{\lambda'}(y))dy - \int_{P_{i_0}(\lambda')} \rho_{\lambda'}(y))dy.$$

(7) *As* $\lambda' \to \lambda$, $|A_1| \to 0$ *by (3) above and if the vectors* $\Pi_{i_0}(z_i)$ *are non zero,* $|A_2| \to 0$ *using (1) above, giving the* C^1 *continuity of* $\tilde{S}$.

Lemma 4.5. *Let* W *be a set of non zero vectors which contains the canonical basis of* $\mathbb{R}^d$ *with the property that for* e *and* f *in* V, $e = rf$ *for a scalar* r *if and only if* $r = -1$. *Assume furthermore that* W *is symmetric with respect to the origin. Let* R *be continuous on* Ω^* *and assume that for* $\lambda \in \mathbb{R}^{\#W}, Q(\lambda) \subset \Omega^*$. *Then the mapping*

$$\lambda \mapsto S(\lambda) := \int_{Q(\lambda)} R(p)dp,$$

is C^1 *continuous on* $\{ \lambda \in \mathbb{R}^{\#W}, \lambda_{-a} + \lambda_a > 0, \forall a \in W \}$ *with*

$$\frac{\partial S}{\partial \lambda_a} = \frac{1}{||a||} \int_{Q(\lambda) \cap \{ p \in \mathbb{R}^d, p \cdot a = \lambda_a \}} R(p)dp.$$

Proof. Since W contains the canonical basis of $\mathbb{R}^d$, the polytope $Q(\lambda)$ is bounded. Let $e, a \in W$ such that $e \neq a$. Let Π_a denote the orthogonal projection onto the hyperplane orthogonal to a. By Lemma 4.4 we have

$$\Pi_a\big(Q(\lambda) \cap \{ x \in \mathbb{R}^d, x \cdot a = \lambda_a \}\big) = \\ \{ y \in \mathbb{R}^d, y \cdot a = 0 \text{ and } y \cdot \Pi_a(e) \leq \lambda_e - \lambda_a \, e \cdot a/||a||^2, e \neq a \}.$$

If $e = -a$ $\lambda_{-a} - \lambda_a \, (-a) \cdot a/||a||^2 = \lambda_{-a} + \lambda_a > 0$ by assumption. We then have trivially $y \cdot \Pi_a(-a) = 0 < \lambda_{-a} + \lambda_a$. Thus

$$\Pi_a\big(Q(\lambda) \cap \{ x \in \mathbb{R}^d, x \cdot a = \lambda_a \}\big) = \\ \{ y \in \mathbb{R}^d, y \cdot a = 0 \text{ and } y \cdot \Pi_a(e) \leq \lambda_e - \lambda_a \, e \cdot a/||a||^2, e \notin \{ a, -a \} \}.$$

If $e \notin \{ a, -a \}$, e and a are independent which implies $\Pi_a(e) \neq 0$. The C^1 continuity of S follows as for Lemma 4.4 □

Recall that for $x \in \Omega_h$, $V(x)$ satisfies the assumptions on W in Lemma 4.5. Define for $v \in \mathbb{U}$

$$\lambda_{ha}(v)(x) = \frac{v(x+ha) - v(x)}{h}.$$

Recall that $\partial_V v(x) = \{ p \in \mathbb{R}^d, p \cdot (he) \leq \lambda_{he}(v)(x), \forall e \in V(x) \}$. Thus, if we put $\lambda \equiv (\lambda_{he}(v)(x))_{e \in V(x)}$ we have

$$\partial_V v(x) = Q(\lambda).$$

We omit the dependence of λ on x as it will be clear from the context.

Theorem 4.6. *The mapping* G *is* C^1 *continuous on* $\mathbb{U}$ *for* R *continuous on* Ω^*.

Proof. By Lemma 2.3, for $v \in \mathbb{U}$, $\partial_V v(x) \subset \Omega^*$ for all $x \in \Omega_h$. Since elements of $\mathbb{U}$ are strictly discrete convex, we have for $v \in \mathbb{U}$, $\lambda_{ha}(v)(x) + \lambda_{-ha}(v)(x) > 0$ for all $x \in \Omega_h$. We note that the mapping $v_h \mapsto v_h(x^1)$ is C^1 continuous. On the other hand, the mapping $v_h \mapsto \omega_V(R, v_h, \{ x^i \})$ is the composite of S and the mapping $v_h \mapsto \left(\frac{v_h(x^i+ha) - v_h(x^i)}{h} \right)_{a \in V(x^i)}$. By Lemma 4.5 the functional G is C^1 continuous on $\mathbb{U}$. □

For $x \in \Omega_h$ and $e \in V(x)$ such that $x + he \notin \Omega_h$, we define

$$\Gamma(x+he) = \operatorname*{argmin}_{y \in \partial\Omega_h} v_h(y) + k_Y(x-y).$$

A priori, $\Gamma(x+he)$ is multi-valued. We assume that for the implementation a unique choice is made for pairs (x,e) such that $x + he \notin \Omega_h$. If $x + he \in \Omega_h$, we put $\Gamma(x+he) = x + he$.

Given $i \in \{1, \dots, M\}$, $\partial_V v(x^i)$ and hence $\omega_V(R, v, \{x^i\})$ depends on $x^i, x^i + he, e \in V(x^i)$ when $x^i + he \in \Omega_h$ and on $\Gamma(x^i + he)$ when $x^i + he \notin \Omega_h$. We have for $i, j \in \{1, \dots, M\}$

(4.2) if $x^j \notin \{x^i + he, e \in V(x^i)\}, x^j \notin \partial\Omega_h$ and $j \neq i$,

$$\partial\omega_V(R, v, \{x^i\})/\partial v(x^j) = 0.$$

(4.3) If $x^j \notin \{x^i + he, e \in V(x^i)\}, x^j \in \partial\Omega_h$ and $j \neq i$,

$$\frac{\partial\omega_V(R, v, \{x^i\})}{\partial v(x^j)} = \sum_{\substack{e' \in V(x^i) \\ x_i + he' \notin \Omega_h, \Gamma(x^i + he') = x^j}} \frac{1}{h||e'||} \int_{Q(\lambda) \cap \{p \in \mathbb{R}^d, p \cdot (he') = \lambda_{he'}\}} R(p)dp.$$

For $e \in V(x^i)$ and $x^i + he \in \Omega_h \setminus \partial\Omega_h$

$$\frac{\partial\omega_V(R, v, \{x^i\})}{\partial v(x^i + he)} = \frac{1}{h||e||} \int_{Q(\lambda) \cap \{p \in \mathbb{R}^d, p \cdot (he) = \lambda_{he}\}} R(p)dp, \tag{4.4}$$

while for $e \in V(x^i)$ and $x^i + he \in \partial\Omega_h$

$$\frac{\partial\omega_V(R, v, \{x^i\})}{\partial v(x^i + he)} = \sum_{\substack{e' \in V(x^i) \\ \Gamma(x^i + he') = x^i + he}} \frac{1}{h||e'||} \int_{Q(\lambda) \cap \{p \in \mathbb{R}^d, p \cdot (he') = \lambda_{he'}\}} R(p)dp, \tag{4.5}$$

where we also used $\Gamma(x^i + he) = x^i + he$. That is, the sum in (4.5) also includes the term on the right hand side of (4.4).

When $x^i \notin \partial\Omega_h$

$$\frac{\partial\omega_V(R, v, \{x^i\})}{\partial v(x^i)} = -\sum_{e \in V(x^i)} \frac{1}{h||e||} \int_{Q(\lambda) \cap \{p \in \mathbb{R}^d, p \cdot (he) = \lambda_{he}\}} R(p)dp. \tag{4.6}$$

Finally when $x^i \in \partial\Omega_h$

$$\begin{aligned} \frac{\partial\omega_V(R, v, \{x^i\})}{\partial v(x^i)} &= -\sum_{e \in V(x^i)} \frac{1}{h||e||} \int_{Q(\lambda) \cap \{p \in \mathbb{R}^d, p \cdot (he) = \lambda_{he}\}} R(p)dp \\ &\quad + \sum_{\substack{e' \in V(x^i) \\ x_i + he' \notin \Omega_h, \Gamma(x^i + he') = x^i}} \frac{1}{h||e'||} \int_{Q(\lambda) \cap \{p \in \mathbb{R}^d, p \cdot (he') = \lambda_{he'}\}} R(p)dp \\ &= -\sum_{\substack{e \in V(x^i) \\ \Gamma(x^i + he) \neq x^i}} \frac{1}{h||e||} \int_{Q(\lambda) \cap \{p \in \mathbb{R}^d, p \cdot (he) = \lambda_{he}\}} R(p)dp. \end{aligned} \tag{4.7}$$

Put

$$L_e = \frac{1}{h||e||} \int_{Q(\lambda)\cap\{\, p\in\mathbb{R}^d, p\cdot(he)=\lambda_{he}\,\}} R(p)dp.$$

Let $A = (a_{ij})_{i,j=1,\ldots,M}$ be a $M \times M$ matrix. A row i of A is diagonally dominant if $|a_{ii}| \geq \sum_{j\neq i} |a_{ij}|$ and strictly diagonally dominant if the strict inequality holds. A matrix A is said to be weakly diagonally dominant if all its rows are diagonally dominant. A matrix A is strictly diagonally dominant if all its rows are strictly diagonally dominant.

We will prove that for v in $\mathbb{U}$, $G'(v) - \nu I$ is invertible as a strictly diagonally dominant under a condition on ν and the constant w in (1.6)

Theorem 4.7. *At each v in $\mathbb{U}$, the matrix B with entries $\partial\omega_V(R, v, \{\, x^i \,\})/\partial v(x^j)$ $i, j = 1, \ldots, M$, is a weakly diagonally dominant matrix with $B_{ii} \leq 0$, $B_{ij} \geq 0, j \neq i$ and $|B_{ii}| = \sum_{j\neq i} |B_{ij}|$, $i, j = 1, \ldots, M$.*

Proof. Note that the entries of B are computed from (4.2)–(4.7). This also shows that $B_{ii} \leq 0$ and $B_{ij} \geq 0, j \neq i$, $i, j = 1, \ldots, M$. Recall that for a given index $i = 1, \ldots, M$, and $e \in V(x^i)$, either $x_i + he \in \Omega_h$ or $x_i + he \notin \Omega_h$ and when $x_i + he \notin \Omega_h, \Gamma(x_i + he) = x_j$ for some $x_j \in \partial\Omega_h$.

Using (4.6), we have for $x^i \notin \partial\Omega_h, i \neq 1$

$$\begin{aligned} \left|\frac{\partial\omega_V(R, v, \{\, x^i \,\})}{\partial v(x^i)}\right| &= \left| \sum_{e\in V(x^i)} \frac{1}{h||e||} \int_{Q(\lambda)\cap\{\, p\in\mathbb{R}^d, p\cdot(he)=\lambda_{he}\,\}} R(p)dp \right| \\ &= \sum_{e\in V(x^i)} \frac{1}{h||e||} \int_{Q(\lambda)\cap\{\, p\in\mathbb{R}^d, p\cdot(he)=\lambda_{he}\,\}} R(p)dp. \end{aligned}$$

Possible non zero elements of row i correspond to those directions e' for which $x^i + he' \in \Omega_h$ and the directions e'' for which $x^i + he'' \notin \Omega_h$ but for which $\Gamma(x^i + he'') = x^j, x^j \in \partial\Omega_h, j \neq i$. Note that when $\Gamma(x^i + he') = x^j$, the expression (4.5) of $\partial G_i(v)/\partial v(x^j)$ takes into account all directions e' for which $\Gamma(x^i + he') = x_j$. Recall also that by assumption $\Gamma(x^i + he_1) = \Gamma(x^i + he_2)$ only when $e_1 = e_2$.

We claim that, if $x^i \notin \partial\Omega_h$

$$\left|\frac{\partial\omega_V(R, v, \{\, x^i \,\})}{\partial v(x^i)}\right| = \sum_{j\neq i} \left|\frac{\partial\omega_V(R, v, \{\, x^i \,\})}{\partial v(x^j)}\right|. \tag{4.8}$$

Let $e \in V(x^i)$. If $x^i + he \in \Omega_h \setminus \partial\Omega_h$, we have $L_e = \partial\omega_V(R, v, \{\, x^i \,\})/\partial v(x^i + he)$ by (4.4). If $x^i + he \in \partial\Omega_h$, $\partial\omega_V(R, v, \{\, x^i \,\})/\partial v(x^i + he)$ is the sum of L_e and positive terms by (4.5). If $x^i + he \notin \Omega_h$, $\partial\omega_V(R, v, \{\, x^i \,\})/\partial v(x^j)$ where $x^j = \Gamma(x + he)$ is also the sum of L_e and positive terms by (4.5). We conclude that $\sum_{e\in V(x^i)} L_e \leq \sum_{j\neq i} \partial\omega_V(R, v, \{\, x^i \,\})/\partial v(x^j)$.

On the other hand, by (4.2)–(4.5), either $\partial\omega_V(R, v, \{\, x^i \,\})/\partial v(x^j)$ is a sum of terms $L_e, e \in V(x^i), \Gamma(x^i + he) = x^j$ or $\partial\omega_V(R, v, \{\, x^i \,\})/\partial v(x^j) = 0$. Since Γ is one-to-one, we obtain $\sum_{j\neq i} \partial\omega_V(R, v, \{\, x^i \,\})/\partial v(x^j) \leq \sum_{e\in V(x^i)} L_e$. This proves (4.8).

If $x^i \in \partial\Omega_h$, we have using (4.7)

$$\left|\frac{\partial \omega_V(R,v,\{\, x^i \,\})}{\partial v(x^i)}\right| = \sum_{\substack{e\in V(x^i)\\ \Gamma(x^i+he)\neq x^i}} \frac{1}{h||e||}\int_{Q(\lambda)\cap\{\, p\in\mathbb{R}^d, p\cdot(he)=\lambda_{he}\,\}} R(p)dp.$$

If $\Gamma(x^i+he)=x^j$ and $j\neq i$, then $\Gamma(x^i+he)\neq x^i$. The same argument as above then shows that (4.8) holds when $x^i\in\partial\Omega_h$. This shows that for all $i,j=1,\dots,M$, $|B_{ii}|=\sum_{j\neq i}|B_{ij}|$. □

Theorem 4.8. *At each v in $\mathbb{U}$ the matrix $G'(v)-\nu I$ is invertible for $\nu>-w$ where $w\leq 0$ is the constant in* (1.6).

Proof. Let D be the $M\times M$ matrix with $D_{ij}=0, j\neq 1$ and $D_{i1}=-w$, $i,j=1,\dots,M$. Put $A=G'(v)-\nu I$. We have $A=B+D-\nu I$ where B is the matrix from Theorem 4.7. Recall that $B_{ii}\leq 0$, $B_{ij}\geq 0, j\neq i$ and $|B_{ii}|=\sum_{j\neq i}|B_{ij}|$, $i,j=1,\dots,M$.

Since $w\leq 0$ and $\nu>-w\geq 0$ we get $|A_{11}|=|B_{11}-w-\nu|=-B_{11}+w+\nu=|B_{11}|+w+\nu=\sum_{j\neq 1}|B_{1j}|+w+\nu=\sum_{j\neq 1}|A_{1j}|+w+\nu>\sum_{j\neq 1}|A_{1j}|$.

Let $1<i\leq M$. We have $A_{ii}=B_{ii}-\nu$. Thus $|A_{ii}|=-B_{ii}+\nu$. If $B_{i1}=0$, then $A_{i1}=-w$. We have

$$\sum_{j\neq i}|A_{ij}|=|A_{i1}|+\sum_{j\notin\{\, i,1\}}|A_{ij}|=|w|+\sum_{j\neq i}|B_{ij}|=-w+|B_{ii}|=-w-B_{ii}=|A_{ii}|-\nu-w.$$

It follows that $|A_{ii}|-\sum_{j\neq i}|A_{ij}|=w+\nu>0$ since $\nu>-w$.

If $1<i\leq M$ and $B_{i1}\neq 0$, then $A_{i1}=-w+B_{i1}$. Since $w\leq 0$ and $B_{i1}\geq 0$, we have $|A_{i1}|=-w+B_{i1}$. Also, for $j\neq i$ and $j\neq 1$, $A_{ij}=B_{ij}$ and $|A_{ii}|=|B_{ii}|+\nu$. We have

$$|A_{ii}|-\sum_{j\neq i}|A_{ij}|=|B_{ii}|+\nu-\sum_{\substack{j\neq i\\ j\neq 1}}|A_{ij}|-|A_{i1}|=|B_{ii}|+\nu-\sum_{\substack{j\neq i\\ j\neq 1}}|B_{ij}|-|A_{i1}|$$

$$=\sum_{j\neq i}|B_{ij}|+\nu-\sum_{\substack{j\neq i\\ j\neq 1}}|B_{ij}|-|A_{i1}|=\nu+|B_{i1}|-|A_{i1}|=\nu+|B_{i1}|+w-B_{i1}=\nu+w>0.$$

We conclude that $G'(v)-\nu I$ is a strictly diagonally dominant matrix and is hence invertible. This completes the proof. □

Theorem 4.9. *The mapping $G:\mathbb{U}_\epsilon\to\mathbb{R}^M$ is proper for $w<0$.*

Proof. Let K be a compact subset of $\mathbb{R}^M$. Since G is continuous by Lemma 4.6, $G^{-1}(K)$ is closed. As K is bounded, there exists a constant C such that for $v\in G^{-1}(K)$ we have $||G(v)||\leq C$. Since for $i=1,\dots,M$, $G_i(v)=\omega_V(R,v,\{\, x^i\,\})$ $-\int_{C_{x^i}}f(t)dt-w\,v(x^1)$ we obtain

$$|w|\,|v(x^1)|\leq ||G(v)||+\omega_V(R,v,\{\, x^i\,\})+\int_{C_{x^i}}f(t)dt\leq ||G(v)||+\int_{\Omega^*}R(p)dp+\int_{\Omega}f(t)dt.$$

It follows that for $v\in G^{-1}(K)$, we have $|v(x^1)|\leq C/|w|$ for a constant C. Using Lemma 2.2 we obtain for $x\in\Omega_h$ and $v\in\mathbb{U}$, $|v(x)|\leq|v(x^1)|+|v(x)-v(x^1)|\leq C$. This shows that $G^{-1}(K)$ is a bounded subset of $\mathbb{U}_\epsilon$. We concluded that $G^{-1}(K)$ is compact. □

We can now give the proof of Theorem 4.2.

Proof of Theorem 4.2. The result follows immediately from Theorem 3.1, Theorem 4.3 and Theorems 4.6-4.9. □

Theorem 4.10. *There exists a solution to* (1.9).

Proof. Taking limits of solutions to (1.8) as $w \to 0$, we obtain a solution to (1.9). □

Remark 4.11. *To combine* (1.5) *with the unicity condition* $u(x^1) = 0$ *in a single equation as for* (1.6)*, we need* $w \neq 0$*. That then required* $\nu > -w > 0$*.*

Remark 4.12. *It is possible to prove convergence rates for the damped Newton's method better than the ones in Theorem 3.1, with further regularity assumptions on the density* R *as in* [5]*. We wish to do that in the more general setting of generated Jacobians.*

Remark 4.13. *Theorem 3.1 does not guarantee a fast decrease of the error. For a fixed* δ*, it may be possible to have* i_k *large. In fact,* i_k *may depend on the iterate* v_h^k*. Since* $G'(x) - \nu I$ *is invertible for all* $x \in \Omega_h$ *and* $\nu > -w$*, it should be possible as in the proof of* [6, Proposition 24, step 2] *to prove that* i_k *depends continuously on* v_h^k *and use a compactness argument to get a uniform bound on* i_k*.*

Acknowledgments

The author was partially supported by NSF grant DMS-1720276. The author would like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme "Geometry, compatibility and structure preservation in computational differential equations" where a significant part of this work was undertaken. This work was partially supported by EPSRC grant no EP/K032208/1.

References

[1] G. Awanou. The second boundary value problem for a discrete Monge-Ampère equation. *Journal of Scientific Computing*, 97(1):22, 2023.
[2] I. J. Bakelman. *Convex analysis and nonlinear geometric elliptic equations.* Springer-Verlag, Berlin, 1994. With an obituary for the author by William Rundell, Edited by Steven D. Taliaferro.
[3] S. Bütikofer. *Generalized Newton-type methods for nonsmooth equations in optimization and complementarity problems.* PhD thesis, ETH Zurich, 2008.
[4] J. H. G. Fu. Monge-Ampère functions. I, II. *Indiana Univ. Math. J.*, 38(3):745–771, 773–789, 1989.
[5] J. Kitagawa, Q. Mérigot, and B. Thibert. Convergence of a Newton algorithm for semi-discrete optimal transport. *J. Eur. Math. Soc. (JEMS)*, 21(9):2603–2651, 2019.
[6] Q. Mérigot, J. Meyron, and B. Thibert. An algorithm for optimal transport between a simplex soup and a point cloud. *SIAM J. Imaging Sci.*, 11(2):1363–1389, 2018.
[7] J.-M. Mirebeau. Discretization of the 3D Monge-Ampère operator, between wide stencils and power diagrams. *ESAIM Math. Model. Numer. Anal.*, 49(5):1511–1523, 2015.
[8] J. M. Ortega and W. C. Rheinboldt. *Iterative solution of nonlinear equations in several variables*, volume 30 of *Classics in Applied Mathematics.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. Reprint of the 1970 original.

Department of Mathematics, Statistics, and Computer Science, M/C 249. University of Illinois at Chicago, Chicago, IL 60607-7045, USA

Email address: awanou@uic.edu

URL: http://www.math.uic.edu/~awanou