On the convergence to saddle points of concave-convex functions, the gradient method and emergence of oscillations

Thomas Holding and Ioannis Lestas

Abstract— It is known that for a strictly concave-convex function, the gradient method introduced by Arrow and Hurwicz [1], has guaranteed global convergence to its saddle point. Nevertheless, there are classes of problems where the function considered is not strictly concave-convex, in which case convergence to a saddle point is not guaranteed.

In the paper we provide a characterization of the asymptotic behaviour of the gradient method, in the general case where this is applied to a general concave-convex function. We prove that for any initial conditions the gradient method is guaranteed to converge to a trajectory described by an explicit linear ODE. We further show that this result has a natural extension to subgradient methods, where the dynamics are constrained in a prescribed convex set. The results are used to provide simple characterizations of the limiting solutions for special classes of optimization problems, and modifications of the problem so as to avoid oscillations are also discussed.

I. INTRODUCTION

Finding the saddle point of a concave-convex function is a problem that is relevant in many applications in engineering and economics and has been addressed by various communities. It includes, for example, optimization problems that are reduced to finding the saddle point of a Lagrangian. The gradient method, first introduced by Arrow and Hurwicz [1] has been widely used in this context for network optimization problems as it leads to localised primal and dual update rules. It has thus been extensively used in areas such as resource allocation in communication and economic networks (e.g. [6], [7], [9], [3]). It is known that for a strictly concave-convex function, the gradient method is guaranteed to converge to its saddle point. Nevertheless, there are classes of problems where the function considered is not strictly concave-convex, and it has been observed that convergence to a saddle point is not guaranteed in this case, with the gradient dynamics leading instead to oscillatory solutions [1], [3], [4].

Our aim in this paper is to provide a characterization of the asymptotic behaviour of the gradient method, in this latter case where convergence is not guaranteed. In particular, we consider the general formulation where the gradient method is applied to any concave-convex function in C^2. One of our main results is to prove that for any initial conditions the gradient method is guaranteed to converge to a trajectory described by an explicit linear ODE. We further show that this result has a natural extension to subgradient

Thomas Holding is with the Cambridge Centre for Analysis, Centre for Mathematical Sciences, University of Cambridge, Wilberforce Road, Cambridge, CB3 0WA, United Kingdom; tjh60@cam.ac.uk

Ioannis Lestas is with the Department of Engineering, University of Cambridge, Trumpington Street, Cambridge, CB2 1PZ, United Kingdom and with the Cyprus University of Technology; icl20@cam.ac.uk

methods, where the dynamics are constrained in a prescribed convex set. These results are used to provide a simple characterization of the limiting solutions in special classes of optimization problems and ways of avoiding potential oscillations are also discussed. The proofs of the results in the paper also reveal some interesting geometric properties of concave-convex functions and their saddle points.

The rest of the paper is structured as follows. In section II we define various notions that will be needed to formulate the problem. The problem formulation is given in section III and the main results are stated in section IV. The proofs of the results are given in sections V and VI, with section V focusing on various geometric properties of concave-convex functions that are relevant to the problem. Some modification methods are then briefly discussed in section VIII and conclusions are finally drawn in in section IX.

II. PRELIMINARIES

A. Notation

Real numbers are denoted by $\mathbb{R}$. For vectors $x, y \in \mathbb{R}^n$ the inequality $x < y$ holds if the corresponding inequality holds for each pair of components, $d(x, y)$ is the Euclidean metric and $|x|$ denotes the Euclidean norm.

The space of $k-$times continuously differentiable functions is denoted by C^k. For a sufficiently differentiable function $f(x, y) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ we denote the vector of partial derivatives of f with respect to x as f_x, respectively f_y. The Hessian matrices with respect to x and y are denoted f_{xx} and f_{yy} with f_{xy} and f_{yx} denoting the matrices of mixed partial derivatives in the appropriate arrangement.

For a matrix $A \in \mathbb{R}^{n \times m}$ we denote its kernel and transpose by $\ker(A)$ and A^T respectively.

When we consider a concave-convex function $\varphi(x, y) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ (see Definition 1) we shall denote the pair $\mathbf{z} = (x, y) \in \mathbb{R}^{n+m}$ in bold, and write $\varphi(\mathbf{z}) = \varphi(x, y)$. The full Hessian matrix will then be denoted $\varphi_{\mathbf{zz}}$. Vectors in $\mathbb{R}^{n+m}$ and matrices acting on them will be denoted in bold font (e.g. $\mathbf{A}$). Saddle points (see Definition 2) of φ will be denoted $\bar{\mathbf{z}} = (\bar{x}, \bar{y}) \in \mathbb{R}^{n+m}$.

For subspaces $E \subseteq \mathbb{R}^n$ we denote the orthogonal complement as $E^\perp$, and for a set of vectors $E \subseteq \mathbb{R}^n$ we denote their span as $\mathrm{span}(E)$. The addition of a vector $v \in \mathbb{R}^n$ and a set $E \subseteq \mathbb{R}^n$ is defined as $v + E = \{v + u : u \in E\}$.

For a set $K \subset \mathbb{R}^n$, we denote the interior and relative interior of K as $\mathrm{int}\, K$ and $\mathrm{relint}\, K$ respectively.

For a closed convex set $K \subseteq \mathbb{R}^n$ and $\mathbf{z} \in \mathbb{R}^n$, we define the maximal orthogonal linear manifold to K through $\mathbf{z}$ as

$$M_K(\mathbf{z}) = \mathbf{z} + \mathrm{span}(\{\mathbf{u} - \mathbf{u}' : \mathbf{u}, \mathbf{u}' \in K\})^\perp \qquad \text{(II.1)}$$

and the normal cone to K through $\mathbf{z}$ as

$$N_K(\mathbf{z}) = \{\mathbf{w} \in \mathbb{R}^n : \mathbf{w}^T(\mathbf{z}'-\mathbf{z}) \leq 0 \text{ for all } \mathbf{z}' \in K\}. \quad \text{(II.2)}$$

If K is in addition non-empty, then we define the projection of $\mathbf{z}$ onto K as $\mathbf{P}_K(\mathbf{z}) = \operatorname{argmin}_{\mathbf{w} \in K} d(\mathbf{z}, \mathbf{w})$.

B. Gradient method

Definition 1 (Concave-Convex function)**.** We say that a function $g(x,y) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is (strictly) concave in x (respectively y) if for any fixed y (respectively x), $g(x,y)$ is (strictly) concave as a function of x, (respectively y). If g is concave in x and convex in y we call g concave-convex.

Throughout this paper we will assume that a concave-convex function φ satisfies

$$\begin{aligned} &\varphi(x,y) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \\ &\varphi \in C^2, \ \varphi \text{ is concave in } x \text{ and convex in } y \end{aligned} \quad \text{(II.3)}$$

Definition 2 (Saddle Point)**.** For a concave-convex function $\varphi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ we say that $(\bar{x}, \bar{y}) \in \mathbb{R}^{n+m}$ is a *saddle point* of φ if for all $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$ we have the inequality $\varphi(x,\bar{y}) \leq \varphi(\bar{x},\bar{y}) \leq \varphi(\bar{x},y)$.

If φ in addition satisfies (II.3) then $(\bar{x}, \bar{y})$ is a saddle point if and only if $\varphi_x(\bar{x},\bar{y}) = 0$ and $\varphi_y(\bar{x},\bar{y}) = 0$.

Definition 3 (Gradient method)**.** Given φ satisfying (II.3) we define the *gradient method* as the dynamical system

$$\dot{x} = \varphi_x \qquad \dot{y} = -\varphi_y \quad \text{(II.4)}$$

Definition 4 (Subgradient method)**.** Given φ satisfying (II.3) and a closed convex set $K \subseteq \mathbb{R}^{n+m}$ we define the *subgradient method* by

$$\dot{\mathbf{z}} = \mathbf{f}(\mathbf{z}) - \mathbf{P}_{N_K(\mathbf{z})}(\mathbf{f}(\mathbf{z})) \quad \text{where } \mathbf{f}(\mathbf{z}) = \begin{bmatrix} \varphi_x & -\varphi_y \end{bmatrix}^T \quad \text{(II.5)}$$

C. Applications to concave programming

Finding saddle points has a direct application to concave programming. Concave programming (see e.g. [10]) is concerned with the study of optimization problems of the form

$$\max_{x \in C, g(x) \geq 0} U(x) \quad \text{(II.6)}$$

where $U(x) : \mathbb{R}^n \to \mathbb{R}$, $g(x) : \mathbb{R}^n \to \mathbb{R}^m$ are concave functions, and $C \subseteq \mathbb{R}^n$ is convex. This is associated with the Lagrangian

$$\varphi(x,y) = U(x) + y^T g(x) \quad \text{(II.7)}$$

where $y \in \mathbb{R}^m_+$ are the Lagrange multipliers. Under Slater's condition (II.8), finding a saddle point $(\bar{x}, \bar{y})$ of (II.7) with $\bar{x} \in C$ and $\bar{y} \geq 0$ is equivalent to solving (II.6).

Theorem 5. *Let g be concave, and*

$$\exists x' \in \operatorname{relint} C \text{ with } g(x') > 0. \quad \text{(II.8)}$$

Then $\bar{x}$ is an optimum of (II.6) if and only if $\exists \bar{y}$ with $(\bar{x}, \bar{y})$ a saddle point of (II.7).

The min-max optimization problem associated finding a saddle point of (II.7) is the dual problem of (II.6). The

simple form of the gradient method (Definition 3) makes it an attractive method for finding saddle points. This use dates back to Arrow and Hurwicz[1] who proved the convergence to a saddle point under a strict concavity condition on φ.

More recently the localised structure of the system (II.4) when applied to network optimization problems has led to a renewed interest [3], [7], [9], [8], where the gradient method has been applied to various resource allocation problems.

D. Pathwise stability

A specific form of incremental stability, which we will refer to as pathwise stability, will be needed in the analysis that follows.

Definition 6 (Pathwise Stability)**.** Let $C^1 \ni \mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ and define a dynamical system $\dot{\mathbf{z}} = \mathbf{f}(\mathbf{z})$. Then we say the system is *pathwise stable* if for any two solutions $\mathbf{z}(t), \mathbf{z}'(t)$ the distance $d(\mathbf{z}(t), \mathbf{z}'(t))$, is non-increasing in time.

III. BASIC RESULTS ON THE GRADIENT METHOD/ PROBLEM FORMULATION

It was proved in [4] that the gradient method is pathwise stable, which is stated in the proposition below.

Proposition 7. *Assume (II.3), then the gradient method (II.4) is pathwise stable.*

Because saddle points are equilibrium points of the gradient method we obtain the well known result below.

Corollary 8. *Assume (II.3), then the distance of a solution of (II.4) to any saddle point is non-increasing in time.*

By an application of LaSalle's theorem we obtain:

Corollary 9. *Assume (II.3), then the gradient method (II.4) converges to a solution of (II.4) which has constant distance from any saddle point.*

Thus the limiting behaviour can be fully described by finding all solutions which lie a constant distance from *any* saddle point. The remainder of the paper will be concerned with the classification of these solutions.

In order to facilitate the presentation of the results, for a given concave-convex function φ we define the following sets:

- $\bar{\mathcal{S}}$ will denote the set of saddle points of φ.
- $\mathcal{S}$ will denote the set of solutions to (II.4) that are a constant distance from any saddle point of φ.
- Given $\bar{\mathbf{z}} \in \bar{\mathcal{S}}$, we denote the set of solutions to (II.4) that are a constant distance from $\bar{\mathbf{z}}$, (but not necessarily other saddle points), as $\mathcal{S}_{\bar{\mathbf{z}}}$.
- Given a closed convex set $K \subseteq \mathbb{R}^{n+m}$ we denote the set of solutions to the subgradient method (II.5) that lie a constant distance from any saddle point in K as $\mathcal{S}^K$.

It is later proved that $\mathcal{S}_{\bar{\mathbf{z}}} = \mathcal{S}$ but until then the distinction is important.

Note that if $\bar{\mathcal{S}} = \mathcal{S} \neq \emptyset$ then Corollary 9 gives the convergence of the gradient method to a saddle point.

IV. Main Results

This section summarises the main results of the paper.

Our main result is that solutions of the gradient method converge to solutions that satisfy an explicit linear ODE. Theorems (10-12) are stated for solutions in $\mathbb{R}^{n+m}$, but we also prove that the same results hold when the gradient method is restricted to a convex domain (Theorem 13). As explained in the previous section, in order to characterise the asymptotic behaviour of the gradient method, it is sufficient to characterise the set $\mathcal{S}$ defined in section III. For simplicity of notation we shall state the result for $\mathbf{0} \in \bar{\mathcal{S}}$; the general case may be obtained by a translation of coordinates.

Theorem 10. *Let* (II.3) *hold and* $\mathbf{0} \in \bar{\mathcal{S}}$ *then* $\mathcal{S} = \mathcal{S_0} = \mathcal{X}$ *where* $\mathcal{X}$ *is defined by* (VI.3). *In particular, solutions in* $\mathcal{S}$ *solve the linear ODE:*

$$\dot{\mathbf{z}}(t) = \mathbf{A}\mathbf{z}(t) \tag{IV.1}$$

where $\mathbf{A}$ *is a constant skew symmetric matrix given by*

$$\mathbf{A} = \begin{bmatrix} 0 & \varphi_{xy}(\mathbf{0}) \\ -\varphi_{yx}(\mathbf{0}) & 0 \end{bmatrix}. \tag{IV.2}$$

In general $\mathcal{X}$ is hard to compute. However, we can give characterisations of $\mathcal{X}$ by considering simpler forms of φ. In particular, the 'linear' case occurs when φ is a quadratic function, as then the gradient method (II.4) is a linear system of ODEs. In this case $\mathcal{S}$ has a simple explicit form in terms of the Hessian matrix of φ at $\mathbf{0} \in \bar{\mathcal{S}}$, and in general this provides an inclusion as described below.

Theorem 11. *Let* (II.3) *hold and* $\mathbf{0} \in \bar{\mathcal{S}}$. *Then define*

$$\mathcal{S}_{linear} = \mathrm{span}\{\mathbf{v} \in \ker(\mathbf{B}) : \mathbf{v} \text{ is an eigenvector of } \mathbf{A}\} \tag{IV.3}$$

where $\mathbf{A}$ *is defined by* (IV.2) *and* $\mathbf{B}$ *by*

$$\mathbf{B} = \begin{bmatrix} \varphi_{xx}(\mathbf{0}) & 0 \\ 0 & -\varphi_{yy}(\mathbf{0}) \end{bmatrix}. \tag{IV.4}$$

Then $\mathcal{S} \subseteq \mathcal{S}_{linear}$ *with equality if* φ *is a quadratic function.*

Here we draw an analogy with the recent study [2] on the discrete time gradient method in the quadratic case. There the gradient method is proved to be semi-convergent if and only if $\ker(\mathbf{B}) = \ker(\mathbf{A} + \mathbf{B})$, i.e. if $\mathcal{S}_{linear} \subseteq \bar{\mathcal{S}}$. Theorem 11 includes a continuous time version of this statement.

One of the main applications of the gradient method is to the dual formulation of concave optimization problems where some of the constraints are relaxed by Lagrange multipliers. When all the relaxed constraints are linear, φ has the form

$$\varphi(x, y) = U(x) + y^T(Dx + e) \tag{IV.5}$$

where D is a constant matrix and e a constant vector. One such case was studied by the authors previously in [4]. Under the assumption that U is analytic we obtain a simple exact characterisation of $\mathcal{S}$.

Theorem 12. *Let* φ *satisfying* (II.3) *be defined by* (IV.5) *with* U *analytic and* D, e *constant. Also assume that* $(\bar{x}, \bar{y}) = \bar{\mathbf{z}} \in$
$\bar{\mathcal{S}}$. *Then* $\mathcal{S}$ *is given by*

$$\mathcal{S} = \bar{\mathbf{z}} + \mathrm{span}\{(x, y) \in \mathcal{W} \times \mathbb{R}^m : (x, y) \text{ is}$$
$$\text{an eigenvector of } \begin{bmatrix} 0 & D \\ -D^T & 0 \end{bmatrix}\} \tag{IV.6}$$
$$\mathcal{W} = \{x \in \mathbb{R}^n : s \mapsto U(sx + \bar{x}) \text{ is linear for } s \in \mathbb{R}\}.$$

Furthermore $\mathcal{W}$ *is a linear manifold.*

All these results carry through easily to the case the gradient method is restricted to a closed convex domain.

Theorem 13. *Let* (II.3) *hold and* $K \subseteq \mathbb{R}^{n+m}$ *be a closed convex set with[1]* $\bar{\mathcal{S}} \cap \mathrm{int}\ K \neq \emptyset$. *Then the subgradient method* (II.5) *satisfies* $\mathcal{S}^K = \{\mathbf{z}(t) \in \mathcal{S} : \mathbf{z}(\mathbb{R}) \subseteq K\}$.

Lastly in section VIII we present a method of modifying a concave-convex function to ensure that $\mathcal{S} = \bar{\mathcal{S}}$ so that the gradient method is guaranteed to converge to a saddle point.

Sections V and VI provide sketches of the proofs of Theorems 10-12 above. The full proofs of these results have been omitted due to page constraints, and can be found in an extended version of this manuscript [5]. We also give below a brief outline of these derivations to improve the readability of the manuscript.

First in section V we use the pathwise stability of the gradient method (Proposition 7) and geometric arguments to establish convexity properties of $\mathcal{S}$. Lemma 14 and Lemma 15 tell us that $\bar{\mathcal{S}}$ is convex and can only be unbounded in degenerate cases. Lemma 16 gives an orthogonality condition between $\mathcal{S}$ and $\bar{\mathcal{S}}$ which roughly says that the larger $\bar{\mathcal{S}}$ is, the smaller $\mathcal{S}$ is. These allow us to prove the key result of the section, Lemma 18, which states that any convex combination of $\bar{\mathbf{z}} \in \bar{\mathcal{S}}$ and $\mathbf{z}(t) \in \mathcal{S}_{\bar{\mathbf{z}}}$ lies in $\mathcal{S}_{\bar{\mathbf{z}}}$.

In section VI we use the geometric results of section V to prove the main results (Theorems 10-12). We split the Hessian matrix of φ into symmetric and skew-symmetric parts, which allows us to express the gradients φ_x, φ_y in terms of line integrals from an (arbitrary) saddle point $\mathbf{0}$. This line integral formulation together with Lemma 18 allow us to prove Theorem 10, from which Theorem 11 is then deduced.

To prove Theorem 12 we construct a quantity $V(\mathbf{z})$ that is conserved by solutions in $\mathcal{S}$. In the case considered this has a natural interpretation in terms of the utility function $U(x)$ and the constraints $g(x)$.

V. Geometry of $\bar{\mathcal{S}}$ and $\mathcal{S}$

In this section we will use the gradient method to derive geometric properties of convex-concave functions. We will start with some simple results which are then used as a basis to derive Lemma 18 the main result of this section.

Lemma 14. *Assume* (II.3), *then* $\bar{\mathcal{S}}$ *is closed and convex.*

Proof. Closure follows from continuity of the derivatives of φ. For convexity let $\bar{\mathbf{a}}, \bar{\mathbf{b}} \in \bar{\mathcal{S}}$ and $\mathbf{c}$ lie on the line between them. Consider the two closed balls about $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ that meet

---

[1]It is possible to obtain results on the form of the set $\mathcal{S}^K$ without this assumption, but these are more technical, and due to page constraints, they will be considered in a subsequent work.

at the single point **c**, as in Figure 1. By Proposition 7, **c** is an equilibrium point as the motion is constrained to stay within both balls. It is hence a saddle point. ◻
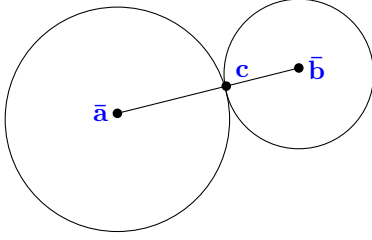

Fig. 1. $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ are two saddle points of φ which satisfies (II.3). By Proposition 7 any solution of (II.4) starting from **c** is constrained for all positive times to lie in each of the balls about $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$.

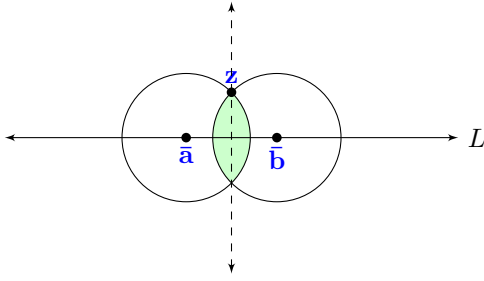Fig. 2. $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ are two saddle points of φ which satisfies (II.3). Solutions of (II.4) are constrained to lie in the shaded region for all positive time by Proposition 7.
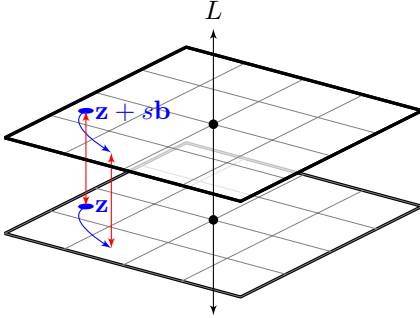

Fig. 3. L is a line of saddle points of φ satisfying (II.3). Solutions of (II.4) starting on hyperplanes normal to L are constrained to lie on these planes for all time. **z** lies on one normal hyperplane, and $\mathbf{z} + s\mathbf{b}$ lies on another. Considering the solutions of (II.4) starting from each we see that by Proposition 7 the distance between these two solutions must be constant and equal to $|s\mathbf{b}|$.

Lemma 15. *Assume* (II.3)*, and let the set of saddle points of φ contain the infinite line $L = \{\mathbf{a} + s\mathbf{b} : s \in \mathbb{R}\}$ for some $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{n+m}$. Then φ is translation invariant in the direction of L, i.e. $\varphi(\mathbf{z}) = \varphi(\mathbf{z} + s\mathbf{b})$ for any $s \in \mathbb{R}$.*

Proof. First we will prove that the motion of the gradient method is restricted to linear manifolds normal to L. Let **z** be a point and consider the motion of the gradient method starting from **z**. As illustrated in Figure 2 we pick two saddle points $\bar{\mathbf{a}}, \bar{\mathbf{b}}$ on L, then by Proposition 7 the motion starting from **z** is constrained to lie in the (shaded) intersection of the two closed balls about $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ which have **z** on their boundaries. The intersection of the regions generated by taking a sequence of pairs of saddle points off to infinity is contained in the linear manifold normal to L.

Next, as illustrated by Figure 3, for $s \in \mathbb{R}$ the motion starting from $\mathbf{z} + s\mathbf{b}$ is exactly the motion starting from **z** shifted by $s\mathbf{b}$. We omit the details.

This implies that φ is defined up to an additive constant on each linear manifold as the motion of the gradient method contains all the information about the derivatives of φ. As φ is constant on L, the continuity of φ completes the proof. ◻

We now use these techniques to prove orthogonality results about solutions in $\mathcal{S}$.

Lemma 16. *Let* (II.3) *hold and $\mathbf{z}(t) \in \mathcal{S}$, then $\mathbf{z}(t) \in M_{\bar{\mathcal{S}}}(\mathbf{z}(0))$ for all $t \in \mathbb{R}$.*

Proof. If $\bar{\mathcal{S}} = \{\bar{\mathbf{z}}\}$ or $\emptyset$ the claim is trivial. Otherwise we let $\bar{\mathbf{a}} \neq \bar{\mathbf{b}} \in \bar{\mathcal{S}}$ be arbitrary, and consider the spheres about $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ that touch $\mathbf{z}(t)$. By Proposition 7, $\mathbf{z}(t)$ is constrained to lie on the intersection of these two spheres which lies inside $M_L(\mathbf{z}(0))$ where L is the line segment between $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$. As $\bar{\mathbf{a}}$ and $\bar{\mathbf{b}}$ were arbitrary this proves the lemma. ◻

Lemma 17. *Let* (II.3) *hold, $\bar{\mathbf{z}} \in \bar{\mathcal{S}}$ and $\mathbf{z}(t) \in \mathcal{S}_{\bar{\mathbf{z}}}$ lie in $M_{\bar{\mathcal{S}}}(\mathbf{z}(0))$ for all t. Then $\mathbf{z}(t) \in \mathcal{S}$.*

Proof. If $\bar{\mathcal{S}} = \{\bar{\mathbf{z}}\}$ the claim is trivial. Let $\bar{\mathbf{a}} \in \bar{\mathcal{S}} \setminus \{\bar{\mathbf{z}}\}$ be arbitrary. Then by Lemma 14 the line segment L between $\bar{\mathbf{a}}$ and $\bar{\mathbf{z}}$ lies in $\bar{\mathcal{S}}$. Let **b** be the intersection of the extension of L to infinity in both directions and $M_{\bar{\mathcal{S}}}(\mathbf{z}(0))$. Then the definition of $M_{\bar{\mathcal{S}}}(\mathbf{z}(0))$ tells us that the extension of L meets $M_{\bar{\mathcal{S}}}(\mathbf{z}(0))$ at a right angle. We know that $d(\mathbf{z}, \bar{\mathbf{z}})$ and $d(\mathbf{b}, \bar{\mathbf{z}})$ are constant, which implies that $d(\mathbf{z}(t), \bar{\mathbf{a}})$ is also constant (as illustrated in Figure 4, we omit the details). ◻
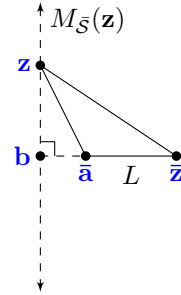

Fig. 4. $\bar{\mathbf{a}}$ and $\bar{\mathbf{z}}$ are saddle points of φ which satisfies (II.3), and L is the line segment between them. **z** is a point on a solution in $\mathcal{S}_{\bar{\mathbf{z}}}$ which lies on $M_{\bar{\mathcal{S}}}(\mathbf{z})$ which is orthogonal to L by definition. **b** is the point of intersection between $M_{\bar{\mathcal{S}}}(\mathbf{z})$ and the extension of L.

Using these orthogonality results we prove the key result of the section, a convexity result between $\mathcal{S}_{\bar{\mathbf{z}}}$ and $\bar{\mathbf{z}}$.

Lemma 18. *Let* (II.3) *hold, $\bar{\mathbf{z}} \in \bar{\mathcal{S}}$ and $\mathbf{z}(t) \in \mathcal{S}_{\bar{\mathbf{z}}}$. Then for any $s \in [0, 1]$, the convex combination $\mathbf{z}'(t) = (1-s)\bar{\mathbf{z}} + s\mathbf{z}(t)$ lies in $\mathcal{S}_{\bar{\mathbf{z}}}$. If in addition $\mathbf{z} \in \mathcal{S}$, then $\mathbf{z}'(t) \in \mathcal{S}$.*

Proof. Clearly $\mathbf{z}'$ is a constant distance from $\bar{\mathbf{z}}$. We must show that $\mathbf{z}'(t)$ is also a solution to (II.4). We argue in a similar way to Figure 3 but with spheres instead of planes. Let the solution to to (II.4) starting at $\mathbf{z}'(0)$ be denoted $\mathbf{z}''(t)$. We must show this is equal to $\mathbf{z}'(t)$. As $\mathbf{z}(t) \in \mathcal{S}$ it lies on a sphere about $\bar{\mathbf{z}}$, say of radius r, and by construction $\mathbf{z}'(0)$ lies on a smaller sphere about $\bar{\mathbf{z}}$ of radius rs. By Proposition 7,

$d(\mathbf{z}(t), \mathbf{z}''(t))$ and $d(\mathbf{z}''(t), \bar{\mathbf{z}})$ are non-increasing, so that $\mathbf{z}''(t)$ must be within rs of $\bar{\mathbf{z}}$ and within $r(1-s)$ of $\mathbf{z}(t)$. The only such point is $\mathbf{z}'(t) = (1-s)\bar{\mathbf{z}} + s\mathbf{z}(t)$ which proves the claim. For the additional statement, we consider another saddle point $\bar{\mathbf{a}} \in \bar{\mathcal{S}}$ and let L be the line segment connecting $\bar{\mathbf{a}}$ and $\bar{\mathbf{z}}$. By Lemma 16, $\mathbf{z}(t)$ lies in $M_{\bar{\mathcal{S}}}(\mathbf{z}(0))$, so by construction, $\mathbf{z}'(t) \in M_{\bar{\mathcal{S}}}(\mathbf{z}'(0))$, (as illustrated by Figure 5). Hence, by Lemma 17, $\mathbf{z}'(t) \in \mathcal{S}$. $\qquad\square$
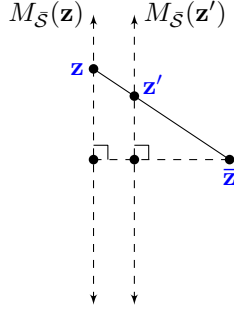

Fig. 5. $\bar{\mathbf{z}}$ is a saddle point of φ satisfying (II.3). $\mathbf{z}$ is a point on a solution in $\mathcal{S}$ and $\mathbf{z}'$ is a convex combination of $\mathbf{z}$ and $\bar{\mathbf{z}}$. $M_{\bar{\mathcal{S}}}(\mathbf{z})$ and $M_{\bar{\mathcal{S}}}(\mathbf{z}')$ are parallel to each other by definition.

Proposition 20 is a further convexity result that can be proved by means of similar methods, using Lemma 19 (their proof has been omitted).

Lemma 19. *Let* (II.3) *hold and let* $\mathbf{z}(t), \mathbf{z}'(t) \in \mathcal{S}$. *Then* $d(\mathbf{z}(t), \mathbf{z}'(t))$ *is constant.*

Proposition 20. *Let* (II.3) *hold, then* $\mathcal{S}$ *is convex.*

VI. CLASSIFICATION OF $\mathcal{S}$

We will now proceed with a full classification of $\mathcal{S}$ and prove Theorems 10-12. For notational convenience we will make the assumption (without loss of generality) that $\mathbf{0} \in \bar{\mathcal{S}}$. Then we compute $\varphi_x(\mathbf{z}), \varphi_y(\mathbf{z})$ from line integrals from $\mathbf{0}$ to $\mathbf{z}$. Indeed, letting $\hat{\mathbf{z}}$ be a unit vector parallel to $\mathbf{z}$, we have

$$\begin{bmatrix} \varphi_x(\mathbf{z}) \\ -\varphi_y(\mathbf{z}) \end{bmatrix} = \left(\int_0^{|\mathbf{z}|} \begin{bmatrix} \varphi_{xx}(s\hat{\mathbf{z}}) & \varphi_{xy}(s\hat{\mathbf{z}}) \\ -\varphi_{yx}(s\hat{\mathbf{z}}) & -\varphi_{yy}(s\hat{\mathbf{z}}) \end{bmatrix} ds \right) \hat{\mathbf{z}}. \quad \text{(VI.1)}$$

Motivated by this expression we define

$$\mathbf{A}(\mathbf{z}) = \begin{bmatrix} 0 & \varphi_{xy}(\mathbf{z}) \\ -\varphi_{yx}(\mathbf{z}) & 0 \end{bmatrix} \quad \mathbf{B}(\mathbf{z}) = \begin{bmatrix} \varphi_{xx}(\mathbf{z}) & 0 \\ 0 & -\varphi_{yy}(\mathbf{z}) \end{bmatrix}$$
$$\mathbf{B}'(\mathbf{z}) = \mathbf{B}(\mathbf{z}) + (\mathbf{A}(\mathbf{z}) - \mathbf{A}(\mathbf{0})). \quad \text{(VI.2)}$$

As $\mathbf{A}(\mathbf{z})$ is skew symmetric and $\mathbf{B}(\mathbf{z})$ is symmetric we have $\ker(\mathbf{B}'(\mathbf{z})) = \ker(\mathbf{B}(\mathbf{z})) \cap \ker(\mathbf{A}(\mathbf{z}) - \mathbf{A}(\mathbf{0}))$. Now we define,

$$\mathcal{X} = \{\mathbf{z}(t) : \dot{\mathbf{z}} = \mathbf{A}(\mathbf{0})\mathbf{z}, \forall t \in \mathbb{R}, 0 \leq r \leq |\mathbf{z}(0)|, \quad \text{(VI.3)}$$
$$\mathbf{z}(t) \in \ker(\mathbf{B}'(r\mathbf{z}(t)))\}.$$

We are now ready to prove the first main result.

Proof of Theorem 10. The proof of the inclusion $\mathcal{X} \subseteq \mathcal{S}$ is relatively straightforward and is omitted for brevity. We sketch the proof that $\mathcal{S}_0 \subseteq \mathcal{X}$, which is sufficient as $\mathcal{S} \subseteq \mathcal{S}_0$. Let $\mathbf{z}(t) \in \mathcal{S}_0$, then by (II.4), (VI.1) and (VI.2), we have

$$\dot{\mathbf{z}}(t) = \mathbf{A}(\mathbf{0})\mathbf{z}(t) + \int_0^{|\mathbf{z}(t)|} \mathbf{B}'(s\hat{\mathbf{z}}(t))\hat{\mathbf{z}}(t) \, ds. \quad \text{(VI.4)}$$

The convexity result Lemma 18 allows us to deduce a similar expression to (VI.4) for $r\mathbf{z}(t) \in \mathcal{S}_0$ for any $r \in [0, 1]$. By comparing these expressions we are able to show that the integrand in (VI.4) is independent of $s \in [0, 1]$, and is thus equal to its value at $s = 0$. Therefore $\dot{\mathbf{z}} = \mathbf{A}(\mathbf{0})\mathbf{z} + \mathbf{B}'(\mathbf{0})\mathbf{z}$, and as $\mathbf{A}(\mathbf{0})$ is symmetric, $\mathbf{B}'(\mathbf{0})$ is skew symmetric, and $|\mathbf{z}(t)|$ is constant, we can deduce that $\mathbf{B}'(\mathbf{0})\mathbf{z}$ vanishes. $\quad\square$

Corollary 21. *Let* (II.3) *hold and there be a saddle point* $\bar{\mathbf{z}}$ *which is locally asymptotically stable. Then* $\mathcal{S} = \bar{\mathcal{S}} = \{\bar{\mathbf{z}}\}$.

Proof. By local asymptotic stability of $\bar{\mathbf{z}}$, $\mathcal{S} \cap \mathcal{B} = \{\bar{\mathbf{z}}\}$ for some open ball $\mathcal{B}$ about $\bar{\mathbf{z}}$. Then by Proposition 20, $\mathcal{S}$ is convex, and we deduce that $\mathcal{S} = \{\bar{\mathbf{z}}\}$. $\quad\square$

To prove Theorem 11 and Theorem 12 we require the lemma below (this can be proved in the same way as a similar result in [4]).

Lemma 22. *Let* X *be a linear subspace of* $\mathbb{R}^n$ *and* $A \in \mathbb{R}^{n \times n}$ *a normal matrix. Let*

$$Y = \text{span}\{v \in X : v \text{ is an eigenvector of } A\}. \quad \text{(VI.5)}$$

Then Y *is the largest subset of* X *that is invariant under either of* A *or the group* e^{tA}.

Proof of Theorem 11. We start with $\mathcal{S}_{\text{linear}} \subseteq \mathcal{S}$ under the assumption that $\mathbf{B}'(\mathbf{z})$ is constant (i.e. φ is a quadratic function). We will use the characterisation of $\mathcal{S}$ given by Theorem 10. By Lemma 22, $\mathcal{S}_{\text{linear}}$ is invariant under $e^{t\mathbf{A}}$, so that $\mathbf{z}(0) \in \mathcal{S}_{\text{linear}} \implies \mathbf{z}(t) = e^{t\mathbf{A}}\mathbf{z}(0) \in \mathcal{S}_{\text{linear}}$. Hence if $\mathbf{z}(0) \in \mathcal{S}_{\text{linear}}$ then $\mathbf{z}(t) \in \ker(\mathbf{B}'(\mathbf{0}))$ for all time t, and as $\mathbf{B}'(\mathbf{z})$ is constant this is enough to show $\mathcal{S}_{\text{linear}} \subseteq \mathcal{S}$.

Next we prove $\mathcal{S} \subseteq \mathcal{S}_{\text{linear}}$. Let $\mathbf{z}(t) \in \mathcal{S}$, then by Theorem 10 taking $r = 0$ we have $\mathbf{z}(t) = e^{t\mathbf{A}} \in \ker(\mathbf{B}'(\mathbf{0}))$ for all $t \in \mathbb{R}$. Thus $\mathcal{S}$ lies inside the largest subset of $\ker(\mathbf{B}'(\mathbf{0}))$ that is invariant under the action of the group $e^{t\mathbf{A}}$, which by Lemma 22 is exactly $\mathcal{S}_{\text{linear}}$. $\quad\square$

In order to prove Theorem 12 we give a different interpretation of the condition in Theorem 10. The condition $\mathbf{z} \in \ker(\mathbf{B}(s\mathbf{z}))$ for all $s \in [0, 1]$ looks like a line integral condition. Indeed, if we define a function $V(\mathbf{z})$ by

$$V(\mathbf{z}) = \mathbf{z}^T \left(\int_0^1 \int_0^1 \mathbf{B}(ss'\mathbf{z})s \, ds' \, ds \right) \mathbf{z} \quad \text{(VI.6)}$$

then as $\mathbf{B}(\mathbf{z})$ is symmetric negative semi-definite we have that $V(\mathbf{z}) = 0$ if and only if $\mathbf{z} \in \ker(\mathbf{B}(s\mathbf{z}))$ for every $s \in [0, 1]$. This still leaves the condition $\mathbf{z} \in \ker(\mathbf{A}(s\mathbf{z}) - \mathbf{A}(\mathbf{0}))$ for all $s \in [0, 1]$, and the function V has no natural interpretation in general. However in the specific case where φ is the Lagrangian of a concave optimization problem where the relaxed constraints are linear, we do have an interpretation. After translating coordinates we deduce, that if φ is of the form (IV.5), then $V(\mathbf{z}) = U(x+\bar{x}) + \bar{y}^T g(x+\bar{x})$. Theorem 12 is now proved from Theorem 10 using Lemma 22 and the conserved quantity $V(\mathbf{z})$ (the proof has been omitted).

VII. THE CONVEX DOMAIN CASE

We now prove that $\mathcal{S}^K = \{\mathbf{z}(t) \in \mathcal{S} : \mathbf{z}(\mathbb{R}) \subseteq K\}$ under the assumption that there exists at least one saddle point in the interior of K.

Proof of Theorem 13. The $\supseteq$ inclusion is clear as if $\mathbf{z}(t)$ is a solution of the unconstrained gradient method (II.4) and remains in K then it is also a solution of (II.5) as the projection term vanishes. Now fix $\bar{\mathbf{z}} \in \bar{\mathcal{S}} \cap \text{int } K$, assume $\mathbf{z}(t) \in \mathcal{S}^K$ and define $W = \frac{1}{2}|\mathbf{z}(t) - \bar{\mathbf{z}}|^2$. Then by (II.5),

$$\dot{W} = (\mathbf{z} - \bar{\mathbf{z}})^T \mathbf{f}(\mathbf{z}) - (\mathbf{z} - \bar{\mathbf{z}})^T \mathbf{P}_{N_K(\mathbf{z})}(\mathbf{f}(\mathbf{z})) = 0 \quad \text{(VII.1)}$$

Both terms are non-positive, the first by Proposition 7 on the gradient method on the full space, and the second by the definition of the normal cone and the convexity of K. The second term vanishing implies[2] that the convex projection is zero, hence $\mathbf{z}(t)$ is a solution to the unconstrained gradient method (II.4). The first term being zero allows us to deduce that $\mathbf{z}(t) \in \mathcal{S}_{\bar{\mathbf{z}}}$ which is equal to $\mathcal{S}$ by Theorem 10. $\qquad\square$

VIII. Modification method for convergence

In this section we will consider methods for modifying φ so that the gradient method converges to a saddle point. We will make statements about $\mathcal{S}$, which carry over directly to the subgradient method on a convex domain K using the relationship proved in Theorem 13. However to ease presentation we will work on the full space.

Various modification methods have been proposed for Lagrangians (e.g. [1], [3]). We will consider a method that was studied in [4]. An important feature of this method, is that when applied to a Lagrangian originating from a network optimization problem, (and with suitable choice of the matrix M), it preserves the localised[3] structure of update rules and requires no additional information transfer. Here we present a natural generalisation of the method to any concave-convex function φ. We define the modified concave-convex function φ' as,

$$\varphi'(x', x, y) = \varphi(x, y) + \psi(Mx - x')$$

$$\psi : \mathbb{R}^l \to \mathbb{R}, M \in \mathbb{R}^{l \times n} \text{ is a constant matrix} \quad \text{(VIII.1)}$$

$$\psi \in C^2 \text{ is strictly concave with max } \psi(0) = 0$$

where x' is a set of l auxiliary variables. It is easy to see that under this condition φ' is concave-convex in $((x', x), y)$. We will denote the set of saddles points of φ' as $\bar{\mathcal{S}}'$ and the set of solutions to (II.4) on φ' which are at a constant distance from elements of $\bar{\mathcal{S}}'$ as $\mathcal{S}'$.

There is an equivalence between saddle points of φ and φ'. If $(x, y) \in \bar{\mathcal{S}}$ then a simple computation shows that $(Mx', x, y) \in \bar{\mathcal{S}}'$. Conversely if $(x', x, y) \in \bar{\mathcal{S}}'$ then we must have $Mx = x'$ and $(x, y) \in \bar{\mathcal{S}}$. In this way searching for a saddle point of φ may be done by searching for a saddle point (x', x, y) of φ' and discarding the extra x' variables.

The condition we require on the matrix M is given in the statement of the result below, and the reason for this assumption is evident in the proof. We do remark however,

that taking $l = n$ and M as the $n \times n$ identity matrix will always satisfy the given condition and preserve the locality of the gradient method in network optimization problems [4].

Proposition 23. *Let* (II.3) *hold,* φ' *satisfy* (VIII.1) *and* $M \in \mathbb{R}^{l \times n}$ *be such that* $\ker(M) \cap \ker(\varphi_{xx}(\bar{\mathbf{z}})) = \{0\}$ *for some* $\bar{\mathbf{z}} \in \bar{\mathcal{S}}$. *Then* $\mathcal{S}' = \bar{\mathcal{S}}'$.

Proof. Without loss of generality we take $\bar{\mathbf{z}} = \mathbf{0}$. We will use the classification of $\mathcal{S}$ given by Theorem 10. Let $\mathbf{z}(t) = (x'(t), x(t), y(t)) \in \mathcal{S}$, then by the form of $\mathbf{A}(\mathbf{0})$ we deduce that $x'(t)$ is constant and by $\mathbf{B}'(s\mathbf{z})\mathbf{z} = 0$ for $s \in [0, 1]$, that

$$0 = \mathbf{z}^T \mathbf{B}'(s\mathbf{z})\mathbf{z} = u^T \psi_{uu} u + x^T \varphi_{xx} x - y^T \varphi_{yy} y \quad \text{(VIII.2)}$$

where ψ_{uu} is the Hessian matrix of ψ evaluated at $u = Mx - x'$. As each term is non-positive and ψ is strictly concave we deduce that $Mx - x' = 0$ and $x \in \ker(\varphi_{xx}(\mathbf{0}))$. Thus $Mx(t)$ is constant. By the condition that $\ker(M) \cap \ker(\varphi_{xx}) = \{0\}$ we deduce that $x(t)$ is constant. Then the form of $\mathbf{A}(\mathbf{0})$ allows us to deduce that $y(t)$ is also constant. $\qquad\square$

IX. Conclusion

We have considered in the paper the problem of convergence to a saddle point of a general concave-convex function that is not necessarily strictly concave-convex. It has been shown that for all initial conditions the gradient method is guaranteed to converge to a trajectory that satisfies a linear ODE. Extensions have also been given to subgradient methods that constrain the dynamics to a convex domain. Simple characterizations of the limiting solutions have been given in special cases and modified schemes with guaranteed convergence have also been discussed. Our aim is to further exploit the geometric viewpoint in the paper to investigate improved schemes in this context with higher order dynamics.

---

[2]Note that, as $\bar{\mathbf{z}}$ is *strictly* inside K, it cannot vanish due to orthogonality.

[3]It should be noted that other modification methods can also lead to distributed dynamics, with potentially improved convergence rate, but these often require additional information transfer between nodes (see [4] for a more detailed discussion). The results in the paper are also applicable in these cases to prove convergence on a general convex domain, but the analysis has been omitted due to page constraints.

References

[1] Kenneth J. Arrow, Leonid Hurwicz, and Hirofumi Uzawa. *Studies in linear and non-linear programming.* Standford University Press, 1958.
[2] Zhong-Zhi Bai. On semi-convergence of hermitian and skew-hermitian splitting methods for singular linear systems. *Computing*, 89(3-4):171–197, 2010.
[3] Diego Feijer and Fernando Paganini. Stability of primal-dual gradient dynamics and applications to network optimization. *Automatica J. IFAC*, 46(12):1974–1981, 2010.
[4] T. Holding and I. Lestas. On the emergence of oscillations in distributed resource allocation. In *52nd IEEE Conference on Decision and Control*, December 2013.
[5] T. Holding and I. Lestas. On the convergence to saddle points of concave-convex functions, the gradient method and emergence of oscillations. Technical report, Cambridge University, September 2014. CUED/B-ELECT/TR.92.
[6] Leonid Hurwicz. The design of mechanisms for resource allocation. *The American Economic Review*, 63(2):1–30, May 1973.
[7] F. Kelly, A. Maulloo, and D. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49(3):237–252, March 1998.
[8] I. Lestas and G. Vinnicombe. Combined control of routing and flow: a multipath routing approach. In *43rd IEEE Conference on Decision and Control*, December 2004.
[9] R. Srikant. *The mathematics of Internet congestion control.* Birkhauser, 2004.
[10] Boyd Stephen and Vandenberghe Lieven. *Convex Optimization.* Cambridge University Press, New York, NY, USA, 2004.