52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

On the emergence of oscillations in distributed resource allocation

Thomas Holding and Ioannis Lestas

***Abstract*— We consider the problem of resource allocation in a decentralized market where users and suppliers trade for a single commodity. Due to the lack of strict concavity, convergence to the optimal solution by means of classical gradient type dynamics for the prices and demands, is not guaranteed. In the paper we explicitly characterize in this case the limiting behaviours of trajectories. Methods of modifying the dynamics are also given, such that convergence to an optimal solution is guaranteed.**

I. Introduction

Problems of distributed resource allocation have been extensively studied by several scientific communities as a result of their significance in important applications such as allocation of resources in a communication networks (e.g. Internet congestion control, multipath routing), or market mechanisms in economic networks (e.g. [7], [10], [6], [13]). A classical approach in this context is to consider the problem of maximizing an aggregate user utility by means of appropriate decentralized update schemes for the primal and dual variables, so as to reach a saddle point of the corresponding Lagrangian. Nevertheless in many cases the underlying structure of the problem leads to a Lagrangian which is not strictly concave-convex, in which case convergence to the desired optimal solution, by means of classical gradient type dynamics, can be problematic, (e.g. [1], [3]).

We focus in this paper on such a case, by considering the problem of distributed resource allocation in an economic network where consumers and suppliers trade for a single commodity (such a setting could be relevant, for example, in electricity markets, but it can also be shown to be closely related to formulations that have been used for multipath routing [7]). We discuss that classical update schemes for the supply and demand might not converge to the optimal solution, and one of our main results is to fully characterize in this case the asymptotic behaviour of trajectories. In particular, it is shown that all trajectories are guaranteed to converge to either a saddle point or to oscillatory solutions that are explicitly characterized in terms of the topology of the network. Furthermore, we discuss how appropriate modifications, which lead to higher order dynamics in the local update schemes, provide guarantees for convergence to the desired equilibrium point.

The paper is structured as follows. In section II various preliminary results and definitions are provided. These include the notions of stability that will be used, and also results on the Arrow gradient method for convergence to a saddle point of a concave-convex function. In section III we describe the problem formulation associated with the economic network that will be studied. In section IV we state the main results of the paper and provide a sketch of their proof. These classify the asymptotic behaviour of the system described in section III and give an explicit form of trajectories that lie in compact invariant sets. Section V discusses methods of modifying the update schemes such that convergence to an equilibrium point that maximizes user utility can be guaranteed.

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; icl20@cam.ac.uk

II. Preliminaries

A. Notation

Real numbers are denoted by $\mathbb{R}$, and the non-negative reals by $\mathbb{R}_+$. For vectors $x, y \in \mathbb{R}^n$ the inequality $x < y$ means $x_i < y_i$ for all $i = 1, \dots n$, and $|x|$ denotes the Euclidean norm of x.

The space C^k denote k times continuously differentiable functions. For a sufficiently differentiable function $f(x, y) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ we denote the partial derivatives with respect to each component of the vectors x or y by f_{x_i} or f_{y_i} respectively. f_x will denote the gradient of f with respect to x, i.e. the vector $(f_{x_1}, f_{x_2}, \dots f_{x_n})$, respectively f_y. Its Hessian matrix with respect to x and y are denoted f_{xx} and f_{yy} respectively, while f_{xy} and f_{yx} denote the matrices of partial derivatives $f_{x_i y_j}$ in the appropriate arrangement. For a function $x(t)$ with time dependence, $\dot{x}, \ddot{x}$ will denote the first and second derivatives with respect to t.

For a matrix $A \in \mathbb{R}^{n \times m}$, its entries are denoted A_{ij} and its kernel and transpose are denoted $\ker(A)$ and A^T respectively. The matrix $\mathrm{diag}(J_1, \dots, J_n)$ with $J_1, \dots, J_n$ square matrices, of possibly different sizes, will denote the block diagonal matrix with diagonal entries $J_1, \dots, J_n$.

The distance from a vector $x \in \mathbb{R}^n$ to a subset $E \subset \mathbb{R}^n$ will be denoted and defined by $\mathrm{dist}(x, E) = \mathrm{dist}(E, x) = \inf\{|x - y| : y \in E\}$.

We will use the notation $\bar{z} = (\bar{x}, \bar{y})$ to denote a saddle point of a function $f(x, y)$ (see definition II.8).

B. Stability

We begin with the definitions of stability and some statements of well known results (e.g.[8]) that will be used later. For the purposes of this section we will look at the autonomous differential equation

$$\dot{x} = f(x) \tag{II.1}$$

where $x \in \mathbb{R}^N$, and $f : D \to \mathbb{R}^N$ is nice enough for the solution to exist and be unique for any initial condition x^0. We will assume D is a simply connected and open subset of $\mathbb{R}^N$. We will assume without loss of generality that the origin is the equilibrium point. Let $x_{x^0}(t)$, be the solution starting at x^0 at $t = 0$.

Definition II.1 (Stability)**.** *The equilibrium point $x = 0$ of* (II.1) *is*

- *Stable iff for any $\epsilon > 0$ there is a $\delta > 0$ such that if $\|x^0\| \leq \delta$ then $\|x_{x^0}(t)\| \leq \epsilon$ for $t \geq 0$.*
- *Asymptotically stable iff stable and for any $\epsilon > 0$ there is a $\delta > 0$ such that if $\|x^0\| \leq \delta$ then $\|x_{x^0}(t)\| \to 0$ as $t \to \infty$.*
- *Globally asymptotically stable iff it is stable and for any initial condition x^0, $\|x_{x^0}(t)\| \to 0$ as $t \to \infty$.*
- *(Globally) exponentially stable iff it is (globally) asymptotically stable and $\|x_{x^0}(t)\|$ decays exponentially to 0 as $t \to \infty$.*

We can extend the definitions of stability to invariant sets.

Definition II.2 (Stability of invariant sets)**.** *An invariant set E has the same definition of stability as an equilibrium point, except we replace the norm conditions $\|x_{x^0}(t)\|$ with distance conditions $\operatorname{dist}(x_{x^0}(t), E)$.*

The notion of stability of a compact invariant set can be used to define stability of solutions. We will also be using a stronger notion in this context. If when we make a small perturbation to the initial condition then the new solution remains close to the old solution through time we say that the solution is pathwise stable. More precisely:

Definition II.3 (Pathwise stability)**.** *We say that a solution $x(t)$ of* (II.1) *is pathwise stable, iff for any $\epsilon > 0$ there is a $\delta > 0$ such that for any other solution $y(t)$ of* (II.1)*, with $\|y(0) - x(0)\| \leq \delta$ we have $\|y(t) - x(t)\| \leq \epsilon$ for all $t \geq 0$. The other forms of stability can be defined for solutions in the same manner. If a family of solutions of* (II.1) *is each pathwise stable with the same δ for each ϵ then we say the family is uniformly pathwise stable. Finally, we say that a solution $x(t)$ of* (II.1) *is globally pathwise stable if there is a constant C such that for any other solution $y(t)$, for all time $t \geq 0$, $\|y(t) - x(t)\| \leq C \|y(0) - x(0)\|$. We extend this to uniform globally pathwise stability in the same way as before.*

Remark II.4. *For a solution $x(t)$ of* (II.1)*, pathwise stability is a stronger condition than the stability of the set $\{x(t) : t \in \mathbb{R}\}$ in the sense of closed invariant sets.*

Lemma II.5. *If a family $\mathcal{F}$ of solutions to* (II.1) *is uniformly pathwise stable, then it is stable, by which we mean that the invariant set*

$$\bigcup_{x(t) \text{ is a solution in } \mathcal{F}} \{x(t) : t \in \mathbb{R}\}$$

is stable. The other forms of stability also carry over from uniform pathwise to the stability of the family.

C. Arrow-Hurwicz gradient method

In this section we summarise the results of Arrow and Hurwicz[1]. The reader is encouraged to consult [3] for a more modern view with various applications.

Definition II.6 (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.*

Definition II.7 (α-strongly concave)**.** *For $\alpha > 0$, a continuously twice differentiable function $f(x) : \mathbb{R}^d \to \mathbb{R}$ is α-strongly concave in a subspace $\mathcal{H} \subset \mathbb{R}^d$ iff its Hessian matrix f_{xx} is negative definite and for any $v \in \mathbb{R}^d$ and for any $\tilde{x} \in \mathcal{H}$, $v^T f_{xx}(\tilde{x}) v \leq -\alpha |v|^2$. If $\mathcal{H} = \mathbb{R}^d$ we simply say that f is α-strongly concave. When we say f is strongly concave we mean that f is α-strongly concave for some $\alpha > 0$. We define strongly convex in the same way, but with f_{xx} positive definite and $v^T f_{xx} v \geq \alpha |v|^2$.*

Concave programming is concerned with an optimisation problem of maximising a *concave* function $f(x) : \mathbb{R}^n \to \mathbb{R}$ subject to some restrictions $Ax = b$, $g(x) \geq 0$, where $g : \mathbb{R}^n \to \mathbb{R}^m$ is a concave function. Some well known preliminary results will now be stated without proof (see for example [1] or the more recent [11]). We will be considering in the paper *primal* problems of the form

$$\max_{x \geq 0,\, g(x) \geq 0} f(x) \tag{II.2}$$

where $f(x) : \mathbb{R}^n \to \mathbb{R}$, $g(x) : \mathbb{R}^n \to \mathbb{R}^m$ are concave functions.
The Lagrangian for (II.2) is

$$\varphi(x, y) = f(x) + y^T g(x) \tag{II.3}$$

where $y \in \mathbb{R}^m_+$ are the Lagrange multipliers.

Definition II.8 (Saddle point)**.** *A saddle point of $\varphi(x, y) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is a non-negative pair $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^m$ such that $\forall x, y \geq 0$,*

$$\varphi(x, \bar{y}) \leq \varphi(\bar{x}, \bar{y}) \leq \varphi(\bar{x}, y)$$

Under Slater's condition (II.4), solving (II.2) is equivalent to finding a saddle of (II.3).

Theorem II.9. *Let g be concave on $x \geq 0$, and*

$$\exists x' > 0 \text{ with } g(x') > 0 \tag{II.4}$$

then $\bar{x}$ is an optimum of (II.2) *iff $\exists \bar{y}$ with $(\bar{x}, \bar{y})$ a saddle point of* (II.3).

The min max optimization problem associated with finding a saddle of (II.3) is the dual problem of (II.2).
Suppose that φ is a Lagrangian such that:

$$\begin{aligned} &\varphi(x, y) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \\ &\varphi \in C^2,\ \varphi \text{ is concave in } x \text{ and convex in } y \end{aligned} \tag{II.5}$$

We wish to design a dynamical system that will enable us to converge to a saddle point. An obvious choice would be to send x in the direction of increasing φ, and y in the direction of decreasing φ. So we choose

$$\dot{x_i} = \begin{cases} 0 & \text{if } x_i = 0 \text{ and } \varphi_{x_i} < 0 \\ \varphi_{x_i} & \text{otherwise} \end{cases}$$
$$\dot{y_i} = \begin{cases} 0 & \text{if } y_i = 0 \text{ and } \varphi_{y_i} > 0 \\ -\varphi_{y_i} & \text{otherwise} \end{cases} \tag{II.6}$$

where the cases above are used to keep $x, y \geq 0$. This is called the Arrow-Hurwicz gradient method. The piecewise definition is not locally Lipschitz, so even existence is not assured. However if we assume some regularity then existence and uniqueness do follow.

Definition II.10. *A regular solution to* (II.6) *applied to* (II.5) *is a solution $z(t) = (x(t), y(t))$ where any sequence of zeros of z_i, $(t_k)_1^\infty$, with $t_k \to 0$ can only (eventually) lie in an interval $[0, t_0]$ in which $z_i(t) = 0$.*

This means that the solution cannot bounce off the edge near $t = 0$ an infinite number of times.

Definition II.11. *We say that a coordinate i of x in a solution pair (x, y) of* (II.6) *is,*

- *Active if $x_i > 0$ or $\varphi_{x_i} > 0$.*
- *Semi-active if $x_i = 0$ and $\varphi_{x_i} = 0$.*
- *Inactive if $x_i = 0$ and $\varphi_{x_i} < 0$.*

We use the same definitions for the components of y, but with $-\varphi_{y_i}$.

Note that x_i (or y_i) cannot be negative, so a coordinate is either active, inactive or semi-active.

Theorem II.12 (Arrow, Hurwicz[1])**.** *Let $\varphi(x, y) \in C^2$, be strictly concave in $x \in \mathbb{R}^n$, convex in $y \in \mathbb{R}^m$ and let $(x^0, y^0) \geq 0$. Then* (II.6) *has exactly one regular solution with initial condition (x^0, y^0).*

Definition II.13. *We say a solution $(x(t), y(t))$ of the gradient method* (II.6) *applied to* (II.5) *is proper, iff for any coordinate x_i, we have that $\forall t, x_i(t) = 0$, or that $x_i(t) = 0$ at only isolated times. We also require the same for the y_j coordinates. We say that a solution is totally proper if all coordinates are zero only at isolated times.*

A proper solution can be considered as a solution to the set of ODEs where we drop coordinates that are always 0, and remove the conditional definition that keeps the other coordinates non-negative. In practical terms, we can just ignore the boundaries for these solutions.

Theorem II.14 (Arrow, Hurwicz[1])**.** *Let* (II.5) *hold, and in addition $\varphi(x, y)$ be strictly concave in x. Let φ have a saddle point $\bar{z} = (\bar{x}, \bar{y}) \geq 0$. Then $\bar{x}$ is unique, and when the gradient method* (II.6) *is applied to φ, $x(t) \to \bar{x}$ for any initial condition $z^0 = (x^0, y^0) \geq 0$.*

If φ is only concave-convex, then the gradient method may have solutions that do not converge to a saddle. However we can make some statements about such solutions.

Theorem II.15 (Arrow, Hurwicz[1])**.** *Let* (II.5) *hold, $\bar{z}$ be a saddle point and $z(t)$ be a solution of* (II.6) *lying in a compact invariant set. Then $|z(t) - \bar{z}|$ is constant.*

III. Economic network

The main results in the paper are associated with the latter case in section II-C where the Lagrangian is only concave-convex, but not strictly concave. In particular, we consider the problem of distributed resource allocation in an economic network where consumers and suppliers trade for a single commodity. As it will be discussed within the paper, the underlying structure of the problem leads to a lack of strict concavity in the Lagrangian, hence classical pricing and user dynamics can fail to converge to the desired solutions. Our main result is to fully characterize in this case the asymptotic behaviour and the global convergence properties.

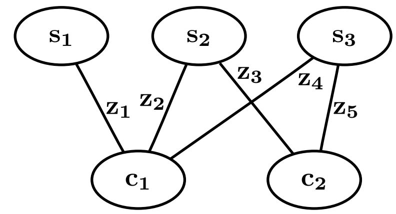

Fig. 1. An example network.

We define a network $\mathcal{G} = (S, C, Z)$ of suppliers $S = (s_i)_1^m$, consumers $C = (c_i)_1^n$ and the links between them $Z = (z_l)_1^L$. The links determine the suppliers each consumer can trade with. We will always assume that the network is connected.

Notation. *We write $\sim$ to indicate connection between consumers, suppliers and links. So that $c_i \sim z_l$ iff c_i is connected to z_l, $s_j \sim z_l$ iff s_j is connected to z_l, and $s_j \sim c_i$ iff the supplier s_j is connected to the consumer c_i via some link z_l.*

If c_i is specified then the sum $\sum_{z_l \sim c_i}$ will be over all links z_l that are connected to c_i, and similarly in other cases.

We will overload z_l to be both the lth link, and the amount of goods being sent down it. It forms a vector $z \in \mathbb{R}^L$.

Each consumer c_i has a strictly concave utility function $U_i : \mathbb{R}_+ \to \mathbb{R}$, and each supplier s_j has a maximum output $Y_j \geq 0$. We also denote $Y = (Y_1, Y_2, \ldots, Y_n)$. The primal problem of maximising total utility is then

Primal.

$$\max_{z \geq 0, x \leq Y} \sum_{c_i} U_i(x_i) \tag{III.1}$$

where $x \in \mathbb{R}^n$ with $x_i = \sum_{c_i \sim z_l} z_l$.

It is easy to see that this problem is concave and it is also assumed that Slater's condition (II.4) is satisfied, so we formulate the dual problem using the Lagrangian

$$\varphi(z, p) = \sum_{c_i} U_i(x_i) + \sum_{s_j} p_j \left(Y_j - \sum_{z_l \sim s_j} z_l \right) \tag{III.2}$$

where $p \in {\mathbb{R}_+}^m$ are Lagrangian multipliers (or marginal prices in economic terms).

Dual.

$$\textit{Find a saddle of } \varphi(z,p) \textit{ in } z \geq 0, p \geq 0 \tag{III.3}$$

with $\varphi(z,p)$ *as defined in* (III.2).

A standard approach with an economic interpretation is to use the gradient method to solve the problem, which leads to decentralized update rules for the prices and demands.

In particular, the gradient method (II.6) applied to (III.2) yields the system:

$$\begin{aligned} \dot{z_l} &= M_{z_l}\left(U_i'(x_i) - p_j\right) \quad \text{where } s_j \sim z_l \sim c_i \\ \dot{p_j} &= M_{p_j}\left(\sum_{z_l \sim s_j} z_l - Y_j\right) \end{aligned} \tag{III.4}$$

where U_i' is the derivative of U_i. The M_{z_l} and M_{p_j} are operators that act to keep the supply along the links and the prices non-negative and are defined by

$$M_{p_j} q = \begin{cases} 0 & \text{if } q < 0 \text{ and } p_j = 0 \\ q & \text{otherwise} \end{cases}$$

We will look at a slight generalisation, where the equations have an extra set of positive constants k_l, k_j' added. We will also rewrite the equation in vector form, defining two matrices that determine the connections between suppliers, consumers and the links.

The matrices $H \in \mathbb{R}^{n \times L}$ and $A \in \mathbb{R}^{m \times L}$ are defined by

$$\begin{aligned} H_{il} &= 1 \text{ if } c_i \sim z_l, \text{ and } 0 \text{ otherwise} \\ A_{jl} &= 1 \text{ if } s_j \sim z_l, \text{ and } 0 \text{ otherwise} \end{aligned}$$

To illustrate this we give these in the case of the example in figure 1.

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \qquad A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

The final system is, (using the notation in (III.4)),

$$\begin{aligned} \dot{z} &= M_z K'(H^T U'(Hz) - A^T p) \\ \dot{p} &= M_p K(Az - Y) \end{aligned} \tag{III.5}$$

where M_p, M_z act diagonally by M_{p_j} and M_{z_l} respectively, $K = \text{diag}(k_1, \ldots, k_m)$, $K' = \text{diag}(k_1', \ldots, k_l')$, and U' is defined by

$$(U'(y))_i = U_i'(y_i) \quad \text{for } i = 1, \ldots, n.$$

Remark III.1. *As defined, M_z and M_p are operators not matrices, but for simplicity we will from now on view them as 0-1-matrices that are defined to depend on the expressions they are applied to in* (III.5).

Remark III.2. *Using* (III.2), (III.5) *could also be formulated as*

$$\begin{aligned} \dot{z} &= M_z K' \varphi_z \\ \dot{p} &= M_p K(-\varphi_p) \end{aligned}$$

Solutions of III.5 are not unique backwards in time. In particular for a solution with a coordinate inactive for all time, we can have other solutions that start with this coordinate positive, join the solution with the inactive coordinate and be equal to it from then on. This makes statements about invariance technical. For space reasons we will restrict the discussion in this paper to when (III.2) has a saddle with all coordinates positive. Note that when all saddles have a coordinate (e.g. z_i) zero, solutions will converge to the plane $\{z_i = 0\}$, hit it and remain on this plane thereafter.

Finally, we will put a regularity requirement on the utility functions:

$$\begin{gathered} U_i : \mathbb{R}_+ \to \mathbb{R}, \qquad U_i \in C^2, \\ U_i \text{ is strictly increasing and strictly concave.} \\ \varphi \text{ in (III.2) has at least one saddle } (\bar{z}, \bar{p}) > 0 \end{gathered} \tag{III.6}$$

IV. Asymptotic behaviour

The goal of this section is to exactly determine the asymptotic behaviour of the system (III.5). We will classify the exact limiting solutions and the convergence to them. This main result is stated as the following theorem.

Theorem IV.1. *Let* (III.6) *hold. Then the solutions of* (III.5) *satisfy the following:*

(i) For all initial conditions, a solution of (III.5) *converges to a proper solution of the form* (IV.3)-(IV.4)*, or is of that form.*

(ii) Totally proper solutions that belong to compact invariant sets are of the form (IV.3)-(IV.4).

(iii) The solutions of (III.5) *are uniformly globally pathwise stable.*

In the rest of this section we provide a sketch of the proof of the theorem. The full proof of the results in the paper have been omitted due to page constraints, and can be found in an extended version of the manuscript [5].

An outline of the proof of Theorem IV.1 is as follows. We define a new norm, denoted as $\mathcal{W}$-norm, and prove Theorem IV.1(iii) by showing that the distance between any two solutions in this norm is non-increasing. To prove convergence we use LaSalle's theorem with the (non-increasing) $\mathcal{W}$-distance from a saddle as the Lyapunov like function. We then characterise the invariant sets which have this distance constant and prove that there are no other invariant sets.

More precisely, to take account of the matrices K and K' we introduce the norm below.

Definition IV.2. $\|(z,p)\|_{\mathcal{W}}^2 = z^T K'^{-1} z + p^T K^{-1} p$

Pathwise stability is a consequence of the following lemma.

Lemma IV.3. *Let* (III.6) *hold, and let* $(z'(t), p'(t))$ *and* $(z(t), p(t))$ *be two solutions of* (III.5). *Then*

$$\frac{d}{dt}\|(z'(t) - z(t), p'(t) - p(t)\|_{\mathcal{W}}^2 \leq 0$$

Proof. We sketch the proof. When we differentiate using the formulation of (III.5) in remark III.2 the K and K' parts cancel. We can then use that φ is concave-convex to deduce that this quantity is non-positive, paying careful attention to the boundary. $\square$

An immediate consequence of this proof is the more general result below on gradient methods, which is not restricted to the economic network.

Corollary IV.4. *Let* (II.5) *hold. Then the solutions of the gradient method* (II.6) *are uniformly globally pathwise stable.*

A first step in the charactisation of invariant sets is the following lemma, which allows us to ignore the boundaries for these sets.

Lemma IV.5. *Let* (III.6) *hold, then all solutions of* (III.5) *that lie in a compact invariant set are proper.*

When we express solutions of (III.5) as $(\bar{z}+\tilde{z}(t), \bar{p}+\tilde{p}(t))$ for some saddle $(\bar{z}, \bar{p})$, it can be directly varified from (III.5) that the contribution of the non-linearity coming from the utility functions is zero if and only if $\tilde{z}(t) \in \ker(H)$. When this condition does not hold, computations similar to that of Lemma IV.3 show that the $\mathcal{W}$-distance from the saddle is decreasing. So all solutions which lie in compact invariant sets will have $\tilde{z}(t) \in \ker(H)$. It is however possible to find solutions which have $\tilde{z}(t_0) \in \ker(H)$ for some t_0 but which do not lie in compact invariant sets. The exact condition for this not to happen is for $\tilde{z}$ to lie in the subspace defined below.

Definition IV.6 (Subspace $\mathcal{Q}$)**.** *We define $Q = K'A^TKA$ and $\mathcal{Q}$ to be the largest subspace of $\ker(H)$ which is invariant under the action of Q. By which we mean that if $v \in \mathcal{Q}$ then $Qv \in \mathcal{Q}$.*

Remark IV.7. *It is easy to see that $\mathcal{Q}$ can equivalently be defined as*

$$\mathcal{Q} = \bigcap_{i=0}^{\infty} Q^i \ker(H)$$

The contrapositive of this is that

$$v \notin \mathcal{Q} \implies \exists i \in \mathbb{N}, \text{ such that } Q^i v \notin \ker(H).$$

Less trivially we can deduce an explicit form

$$\mathcal{Q} = \operatorname{span}\left(\{v \in \ker(H) : v \text{ is an eigenvector of } Q\}\right).$$

Remark IV.8. *There is no guarantee that $\mathcal{Q}$ is non trivial.*

To prove that $\tilde{z}(t) \in \mathcal{Q}$ if and only if the solution lies in a compact invariant set we first show that if we start a solution with initial condition $\tilde{z}(0) \in \mathcal{Q}$ then it obeys the *linear* equation $\ddot{\tilde{z}} + Q\tilde{z} = 0$. From this we can deduce the form (IV.4) from the following results on the structure of Q and its eigenvalues and eigenvectors.

Lemma IV.9. *Q has the structure (up to a reordering of the coordinates of $\tilde{z}$):*

$$Q = \operatorname{diag}(J_1, \dots, J_m)$$
$$J_j = k_j \begin{bmatrix} k'_l & \dots & k'_l \\ k'_{l+1} & \dots & k'_{l+1} \\ \vdots & \vdots & \vdots \\ k'_{l+l'} & \dots & k'_{l+l'} \end{bmatrix} \tag{IV.1}$$

where $z_l, \dots, z_{l+l'}$ are all the links connected to the supplier s_j.

The block diagonal structure of Q means that its eigenvalues are completely determined by the eigenvalues of each J_j.

Lemma IV.10. *The contribution to the eigensystem of Q from each J_j defined by* (IV.1) *is l' zero eigenvalues with eigenvectors in $\ker(A)$, and a single eigenvalue/vector pair:*

$$\begin{aligned} v^j &= [k'_l, \dots, k'_{l+l'}]^T \\ \lambda_j &= k_j(k'_l + \dots + k'_{l+l'}). \end{aligned} \tag{IV.2}$$

where the vector v_j is written out in its form as an eigenvector of J_j.

Proposition IV.11. *Let* (III.6) *hold. Then solutions of* (III.5)

$$z(t) = \bar{z} + \tilde{z}(t) \tag{IV.3}$$

where $(\bar{z}, \bar{p})$ is a saddle and $\tilde{z}(0), \dot{\tilde{z}}(0) \in \mathcal{Q}$ are of the form

$$(\tilde{z}(t))_l = (\tilde{z}(0))_l \cos(\lambda_j t) + (\dot{\tilde{z}}(0))_l \frac{1}{\lambda_j} \sin(\lambda_j t) \tag{IV.4}$$

The constants λ_j are given by (IV.2) *with $s_j \sim z_l$. In particular if $\mathcal{Q} = \{0\}$ then these solutions are saddle points.*

To prove that this exhausts all compact invariant sets comprised of totally proper solutions we consider the case where a solution has $\tilde{z}(t) \in \ker(H)$ for all time t, (a requirement for invariance), and deduce that $\tilde{z} \in \mathcal{Q}$.

To prove convergence to these solutions we use the version of LaSalle's theorem for hybrid systems given in [9] using the $\mathcal{W}$-distance from a saddle as the Lyapunov like function. This concludes the sketch of the proof of Theorem IV.1.

As a final note, we remark that it is possible to prove results on the convergence speed in specific cases.

Proposition IV.12. *Let* (III.6) *hold and all the utility functions U_i be strongly concave. Then the convergence rate of totally proper solutions of* (III.5) *to solutions of the form* (IV.3)-(IV.4) *is exponential.*

V. Modification methods for convergence

We will now look at ways of modifying the system to avoid the occurrence the pathological behaviour and cause convergence to a saddle. One such method was introduced by Arrow and Hurwicz[1], and summarised in [3], which works by transforming the constraints to compensate for the lack of strict concavity of the original Lagrangian. More precisely, the Lagrangian of (II.2) is formulated as

$$\varphi(x,y) = f(x) + \sum_{j=1}^{m} y_j \rho_j(g_j(x))$$

where $\rho_j : \mathbb{R} \to \mathbb{R}$ are strictly increasing and strictly concave functions. The resultant system of equations in the case of the economic problem are, however, not fully localised, as each user needs to also be aware of the demand of other users trading with the same supplier.

A second such method is to add penalty functions (see e.g. [4]) to the Lagrangian to penalise violated constraints. However, in the case of the economic problem the new update rules are again not fully localized as above, requiring additional information exchange.

Another approach is to modify the total utility function in a way that does not alter the maximiser, by introducing local auxiliary variables. In [12] a modification was presented, to the $K, K' = I$ case, of the form

$$U_{new}(z,z') = U_{old}(z) - \tfrac{1}{2}k(z-z')^2$$

where $z' \in \mathbb{R}^L$ is an additional vector to be maximised over, and k is a positive real constant (see also the discrete time setting in [2]). Observe that this has the same z maximisers, meaning that a maximiser of U_{old}, $\bar{z}$ corresponds to a maximiser $(\bar{z},\bar{z})$ of U_{new}, because $-|z-z'|^2 \leq 0$ and $=0$ *iff* $z' = z$. We will look at slightly different modifications of the form

$$U_{new}(z,z') = U_{old}(z) + \gamma(Bz - z')$$

where $\gamma : \mathbb{R}^a \to \mathbb{R}$ is a strictly concave function with a sole maximum $\gamma(0) = 0$, B is a constant matrix and $z' \in \mathbb{R}^a$ is an additional vector to be maximised over. This has the same z maximisers in the same way as before, with the corresponding $\bar{z}'$ equal to $B\bar{z}$. The corresponding Lagrangian and assumptions are

$$\begin{aligned} \varphi(z,z',p) = \sum_{c_i} & U_i\left(\textstyle\sum_{z_l\sim c_i} z_l\right) + \\ & + \gamma(Bz-z') + p^T(Y - Az) \\ \varphi : \mathbb{R}^L \times & \mathbb{R}^a \times \mathbb{R}^m \to \mathbb{R}, \quad \gamma : \mathbb{R}^a \to \mathbb{R} \\ \gamma \in C^2 & \text{ strictly concave with maximum } \gamma(0) = 0 \\ B \in \mathbb{R}^{a\times L} & \text{ is a constant matrix, and (III.6) holds} \end{aligned} \tag{V.1}$$

Proposition V.1. *Let* (V.1) *hold and the range of* B *be* $\mathbb{R}^a$*, then the gradient method* (II.6) *on the modified Lagrangian* (V.1) *converges to a maximum of the original problem given by* (III.1).

However the addition of L extra variables is unnecessary. We only need to add enough to make up for the non-strict concavity of the sum of $U_i\left(\sum_{c_i\sim z_l} z_l\right)$ as a function of z.

Theorem V.2. (II.6) *applied to* (V.1) *will converge to a maximum of the original problem III.1 if* $(Bv = 0 \implies v \notin (\ker(H) \setminus \{0\}))$.

An important class of matrices B is when it is local to each consumer, i.e. the components of z' split into subsets for each consumer, with each subset only depending on the links z_l connected to that consumer. This means that the modification can be performed locally. An obvious example is when B is the identity matrix, but we can also exploit theorem V.2 to reduce the number of variables.

Corollary V.3. *Let* (V.1) *hold. Split the indices of the links up into sets* $Z_i = \{1 \leq l \leq L : c_i \sim z_l\}$ *of those connected to each consumer. Then pick one index* b_i *arbitrarily from each set* Z_i *and form* $B \in \mathbb{R}^{L\times(L-n)}$ *by taking the* $L\times L$ *identity matrix and removing the* b_i*th row for each* $i = 1,\ldots,n$*. The resultant matrix* B *maps vectors in* $\mathbb{R}^L$ *to* $\mathbb{R}^{L-n}$ *by dropping each of the coordinates* b_i*. With this* B *the gradient method* (II.6) *applied to* (V.1) *will converge to a maximum of the original problem III.1.*

VI. Conclusions

We have considered the problem of distributed resource allocation in an economic network where users and suppliers trade for a single commodity. It has been discussed that due to the underlying structure of the problem the corresponding Lagrangian is not strictly concave-convex and hence convergence to an optimal solution by means of gradient type dynamics is not guaranteed. We have provided in this case a characterization of the asymptotic behaviour of trajectories, and also gave an explicit description of trajectories that lie in a compact invariant set. Modifications have also been discussed such that convergence to an optimal solution can be guaranteed. Directions for future research include applications of these results and analysis tools in specific examples, such as analysis of electricity markets and improved schemes for multipath routing in communication networks.

References

[1] Kenneth J. Arrow, Leonid Hurwicz, and Hirofumi Uzawa. *Studies in linear and non-linear programming*. Standford University Press, 1958.

[2] E. Devane and I. Lestas. Stability and convergence of distributed algorithms for the OPF problem. In *52nd IEEE Conference on Decision and Control*, December 2013.

[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] R.M. Freund. Penalty and barrier methods for constrained optimization. *Lecture notes for nonlinear programming. MIT.*, 2004.

[5] T. Holding and I. Lestas. On the emergence of oscillations in distributed resource allocation. Technical report, Cambridge University, September 2013. CUED/F-INFENG/TR.688.

[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] Hassan K. Khalil. *Nonlinear Systems*. Prentice Hall, 2002.

[9] John Lygeros, Karl Henrik Johansson, Slobodan N. Simić, Jun Zhang, and S. Shankar Sastry. Dynamical properties of hybrid automata. *IEEE Trans. Automat. Control*, 48(1):2–17, 2003.

[10] R. Srikant. *The mathematics of Internet congestion control*. Birkhauser, 2004.

[11] Boyd Stephen and Vandenberghe Lieven. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.

[12] R. Tonita and I. Lestas. Oscillations in a decentralized economic network. In *18th IFAC World Congress*, August 2011.

[13] H. R. Varian. *Microeconomic Analysis*. W.W.Norton & Company, 1996.