Lyapunov Convergence for Lagrangian Models of Network Control

Nigel Walker Ben Strulo Marc Wennink

BT Research

MoN June 07

Network Control

- Rate control
- Routing
- DSL Access (Spectrum Management)
- Wireless power control
- Overload control

Network Control

- Rate control
  Kelly, Maulloo & Tan [1998], Jin, Wei & Low (FAST TCP) [2004]
- Routing
  Griffin, Sheperd & Wilfong [2002], Walker & Wennink [2005]
- DSL Access (Spectrum Management)
  Cendrillon, Huang, Chiang & Moonen [2007]
- Wireless power control
  Hande, Rangan & Chiang [2006], ...
- Overload control
  Wennink, Williams, Walker & Strulo [2007]

Network Control

- Routing and congestion control
  Paganini [2006]
- Routing, congestion control, and MAC scheduling
  Chen, Low, Chiang, & Doyle [2006]

- Layering
  Chiang, Low, Calderbank & Doyle [2007]
  'Layering as optimization decomposition: A mathematical theory of network architectures'

Lagrangian Models

- State (or formulate) control problem as objective and constraints in the language of mathematical optimisation theory, eg.
  - control problem: routing
  - objective: minimise cost of flow
  - constraints: maintain flow balance at nodes
- Combine objective and constraints into single function, the Lagrangian. Each constraint introduces a dual variable (Lagrange multiplier). Optimisation problem becomes a saddle point problem.

Lyapunov Convergence

- Decompose Lagrangian into a collection of subproblems.
- Different parts of network then own different variables.
- Interaction between subproblems specifies a distributed algorithm, or dynamic system.
- Lyapunov function is a certificate that this algorithm does find the saddle point, as intended.

Example - Flow control

Network Utility Maximisation (NUM) formulation

$$\max_{x_i} \sum_i U_i(x_i)$$

$$\sum_{i \in j} x_i < K_j$$

- User determined flows, x_i
- Concave utility of flows $U_i(x_i)$
- Network resource capacities, K_j
- $i \in j$ if flow i uses resource j

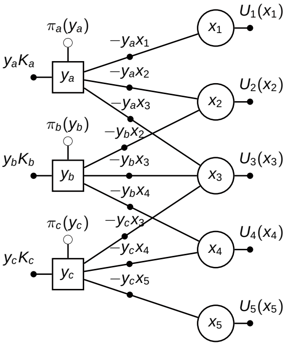
Flow control

Lagrangian formulation

$$L(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} U_i(x_i) + \sum_{j=1}^{m} \Big(y_j \big(K_j - \sum_{l \in j} x_l\big) - \pi_j(y_j) \Big)$$

- $y_j > 0$ Lagrange multiplier associate with resource j
- $\pi_j(y_j)$ Barrier function representing queue behaviour or congestion costs.

Graphical presentation of Lagrangian



dual variables	primal variables
prices	flows
owned by network	owned by user

Flow control dynamics, (Kelly, Maulloo, Tan, 1998)

Primal algorithm (User control)

$$\dot{x}_i = \kappa\left(w_i - x_i \sum_{k \in i} y_k\right)$$

Dual algorithm (Network control)

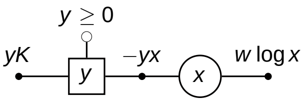
$$\dot{y}_j = \nu\left(\sum_{l \in j} x_l - \left(K_j - \pi_j'(y_j)\right)\right)$$

- Why these two candidates?
- Do they reach equilibrium?
- If so, is the equilibrium the saddle point of L?
- Are there other possibilities?

Illustration, single flow, single resource

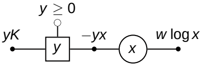
Utility $U(x) = w\log(x)$

$$\begin{array}{ll} \max & w\log x \\ \text{s.t.} & x \le K \end{array} \quad \Rightarrow \quad \min_{y\ge 0}\max_{x} \; w\log x - y(x-K)$$



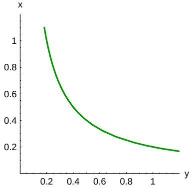
Saddle point conditions

$$L(x;y) = w \log x - yx + yK \,, \quad y \geq 0$$



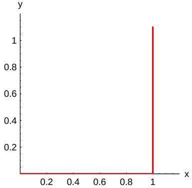
given y : $\max\limits_{x} \; L(x;y)$

$$\frac{\partial L}{\partial x} = 0 \Leftrightarrow \frac{w}{x} - y = 0 \Rightarrow x = \frac{w}{y}$$



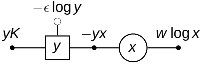
given x : $\min\limits_{y \geq 0} \; L(x;y)$

$$y \begin{cases} = 0 & \text{if } x < K \\ = +\infty & \text{if } x > K \\ \geq 0 & \text{if } x = K \end{cases}$$



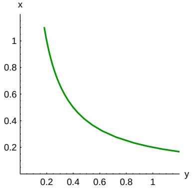
Saddle point conditions

$$L(x;y) = w\log x - yx + yK - \epsilon \log y$$



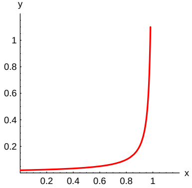
given y : $\max_x L(x;y)$

$$\frac{\partial L}{\partial x} = 0 \Leftrightarrow \frac{w}{x} - y = 0 \Rightarrow x = \frac{w}{y}$$

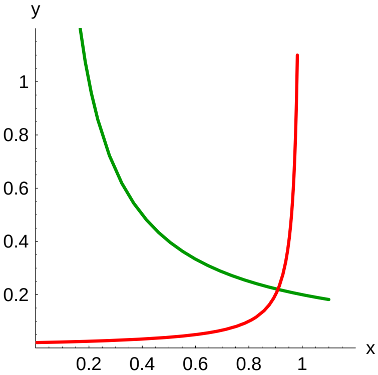


given x : $\min_y L(x;y)$

$$\frac{\partial L}{\partial y} = 0 \Rightarrow y = \frac{\epsilon}{K - x}$$



The saddle point



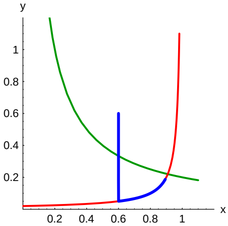
Primal algorithm

We assume y updates instantaneously, maintaining $y = \frac{\epsilon}{K-x}$.

Then x performs a gradient search:

$$\frac{dx}{dt} = \lambda \frac{\partial L}{\partial x} = \lambda\left(\frac{w}{x} - y\right)$$

(Increase x if $U'(x) > y$, decrease if $U'(x) < y$.)



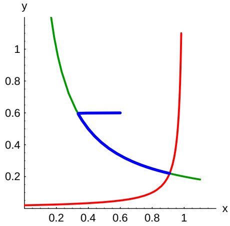
Dual algorithm

We assume x updates instantaneously, maintaining $x = \frac{w}{y}$.

Then y performs a gradient search:

$$\frac{dy}{dt} = -\mu \frac{\partial L}{\partial y} = \mu(x - K + \frac{\epsilon}{y})$$

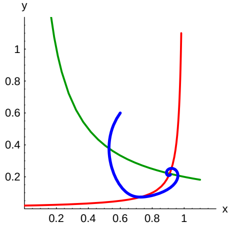
(Increase y if $x > K$, decrease if $x \ll C$.)



Combined primal-dual algorithm

Both x and y perform a gradient search:

$$\frac{dx}{dt} = \lambda \frac{\partial L}{\partial x} \qquad \frac{dy}{dt} = -\mu \frac{\partial L}{\partial y}$$



Our result, informally

Given

- concave-convex $L(\mathbf{x}, \mathbf{y})$
- concave $F(\mathbf{x})$, convex $G(\mathbf{y})$

Then trajectories with

$$-\frac{d}{dt}\nabla F = \frac{\partial L}{\partial x} \qquad -\frac{d}{dt}\nabla G = \frac{\partial L}{\partial y}$$

converge on the saddle point of L.

Flow control dynamics, (Kelly, Maulloo, Tan, 1998)

Primal algorithm (User control)

$$F(x) = -\frac{1}{\kappa}\sum_i \log x_i$$

Dual algorithm (Network control)

$$G(y) = \frac{1}{2\nu}\sum_j y_j^2$$

- Automatic convergence proof
- Can combine primal and dual algorithms

Proof is by Lyapunov function

A function $\phi(x(t), y(t))$ such that

- $\phi \geq 0$
- $d\phi/dt < 0$ everywhere except at equilibrium

acts as a certificate of stability, or convergence.

Our Lyapunov function

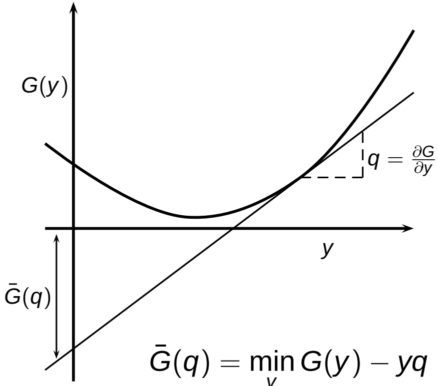
$$\phi(\mathbf{x}, \mathbf{y}) = \overline{G}(\mathbf{q}(\mathbf{y})) - \overline{F}(\mathbf{p}(\mathbf{x}))$$

where

$$\overline{G}(\mathbf{q}) \longleftrightarrow G(\mathbf{y})$$
$$\overline{F}(\mathbf{p}) \longleftrightarrow F(\mathbf{x})$$

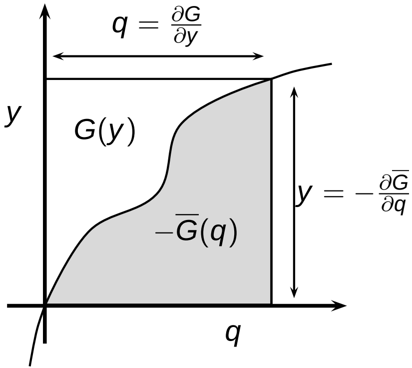
are related by Legendre transform.

The Legendre transform - first visualisation

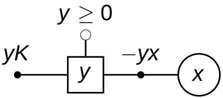


$$\bar{G}(q) = \min_y G(y) - yq$$

The Legendre transform - second visualisation



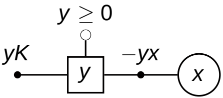
Flow Balance example



capacity constraint:

$$K \geq x$$

Flow Balance example



capacity constraint:

$$K \geq x$$

Intuition:

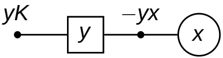
y reacts to flow imbalance

- increases ($\rightarrow +\infty$) when $x > K$
- decreases ($\rightarrow 0$) when $x < K$

is a signal which should lead to reduction of imbalance

- distance label in routing; congestion price in flow control

Flow Balance example



flow balance constraint:

$$K = x$$

Intuition:

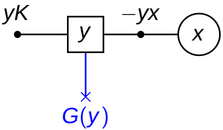
y reacts to flow imbalance

- increases ($\rightarrow +\infty$) when $x > K$
- decreases ($\rightarrow -\infty$) when $x < K$

is a signal which should lead to reduction of imbalance

- distance label in routing; congestion price in flow control

Flow Balance example



flow imbalance:

$$K \approx x$$

Intuition:

y reacts to flow imbalance

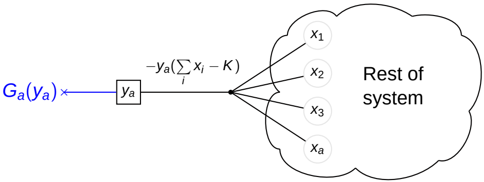
- increases ($\rightarrow +\infty$) when $x > K$
- decreases ($\rightarrow -\infty$) when $x < K$

is a signal which should lead to reduction of imbalance

- distance label in routing; congestion price in flow control

$G(y)$ specifies dynamic response of y to imbalance.

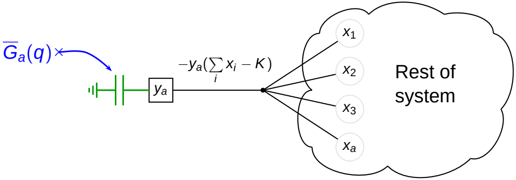
Behaviour of y_j



Rather than define y_a (or $\frac{dy_a}{dt}$) directly in terms of the imbalance (and somehow via G) . . .

. . . consider the dynamics in terms of accumulated imbalance

Behaviour of y_j

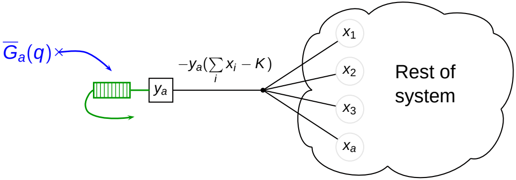


Intuition: consider charge stored in capacitor

Rather than define y_a (or $\frac{dy_a}{dt}$) directly in terms of the imbalance (and somehow via G) . . .

. . . consider the dynamics in terms of accumulated imbalance

Behaviour of y_j

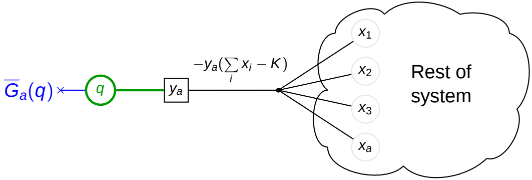


Intuition: consider charge stored in capacitor or packets stored in queue

Rather than define y_a (or $\frac{dy_a}{dt}$) directly in terms of the imbalance (and somehow via G) . . .

. . . consider the dynamics in terms of accumulated imbalance

Behaviour of y_j



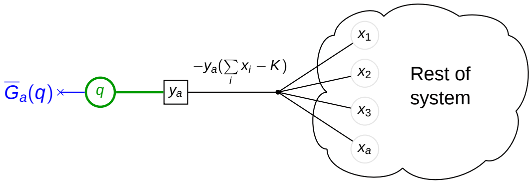
Intuition: consider charge stored in capacitor or packets stored in queue

Rather than define y_a (or $\frac{dy_a}{dt}$) directly in terms of the imbalance (and somehow via G) . . .

. . . consider the dynamics in terms of accumulated imbalance

In fact we use an abstract intermediate variable which integrates this imbalance

Behaviour of y_j



Consider a small time period

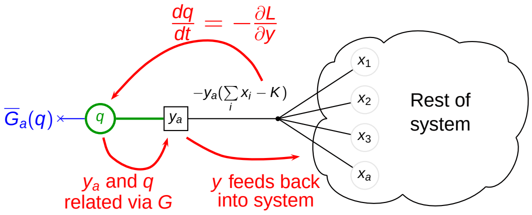
$$\delta q - (\sum_i x_i - K)\delta t = 0$$

Or for a general Lagrangian L

$$\frac{dq}{dt} = -\frac{\partial L}{\partial y}$$

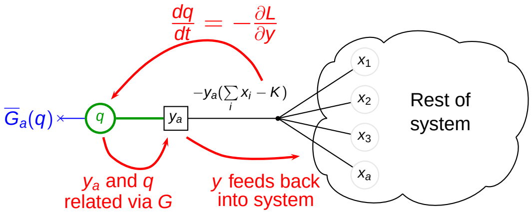
Here q is determined by the dynamics of the rest of the system

Behaviour of y



Now if we can define the behaviour of y_a in terms of q then we have defined the process as we require

Behaviour of y



We use G and its Legendre Transform to define the relationship between q and y:

$$q = \frac{\partial G}{\partial y} \qquad y = \frac{\partial \overline{G}}{\partial q}$$

Now the flow balance equation will give us dynamics for y, and the behaviour of $\overline{G}$ gives us a Lyapunov function.

Proof Outline

Eliminating q in the flow balance equation

$$\frac{dq}{dt} = -\frac{\partial L}{\partial y}$$

gives us dynamics for y

$$\frac{d}{dt}\left(\frac{\partial G}{\partial y}\right) = -\frac{\partial L}{\partial y}$$

and a Lyapunov function

$$\phi(y) = \overline{G}(q(y))$$

Proof Outline

The Lyapunov function is decreasing because

$$
\begin{aligned}
\frac{d}{dt}\phi(y) &= \frac{d}{dt}\,\overline{G}(q(y)) && \\
&= \frac{\partial \overline{G}}{\partial q}\,\frac{dq}{dt} && \text{chain rule} \\
&= -y\,\frac{d}{dt}\left(\frac{\partial G}{\partial y}\right) && \text{Legendre transform } \times 2 \\
&= -y\,\frac{\partial L}{\partial y} && \text{dynamic equation} \\
&\leq 0 && \text{convexity of } L
\end{aligned}
$$

Main Result

$$-\frac{d}{dt}\nabla F \in \partial_x L$$

$$-\frac{d}{dt}\nabla G \in \partial_y L$$

- Primal-dual case
- More than 1 dimension
- General coupled energy functions
- Global asymptotic convergence
- Arbitrary equilibrium (away from origin)
- Non-differentiable L (via sub-gradients)

In this way we directly integrate the convex optimisation statement with the formulation of the dynamic system.

Non-Strict Lagrangians

Strictness can be side-stepped by using the LaSalle Theorem

- Convergence to $\dot{\phi} = 0$ (i.e. in strict dimensions) may be enough.
  For example, if all primals converge we may not care about the duals.
- Alternatively, and typically, convergence in the strict dimensions may imply convergence in the non strict ones since the limit set is the largest invariant set inside $\dot{\phi} = 0$.

Conclusion

- Advocate a methodology for network control based on optimisation theory
  - formal process from specification to implementation
  - certifying good behaviour along the way
  - "design for provability"
- We provide a clean and general result in support of this methodology
- Striking integration of concepts linking optimisation with dynamics
- More such results are required
  - discrete time, state space
  - delay

References

- Ben Strulo, Nigel Walker and Marc Wennink, ‘Lyapunov Convergence for Lagrangian models of Network Control’ T. Chahed and B. Tuffin (Eds.): NET-COOP 2007, LNCS 4465, pp.168-177, 2007.
- F. Kelly, A. Maulloo and D. Tan, ‘Rate control in communication networks: shadow prices, proportional fairness and stability.’, In Journal of the Operational Research Society, Volume 49 (1998).