Nonlinear contraction tools for constrained optimization

by

Jonathan Soto

Ingenieur, Ecole Nationale Superieure de l'Aeronautique et de l'Espace (2008)

Submitted to the Department of Mechanical Engineering in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering at MIT

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2010

Author ....
Department of Mechanical Engineering
August 20, 2010

Certified by ....
Jean-Jacques Slotine
Professor of Mechanical Engineering;
Professor of Brain and Cognitive Sciences
Thesis Supervisor

Accepted by ....
David E. Hardt
Chairman, Department Committee on Graduate Theses

Nonlinear contraction tools
for constrained optimization

by

Jonathan Soto

Submitted to the Department of Mechanical Engineering
on August 20, 2010, in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering at MIT

Abstract

This thesis derives new results linking nonlinear contraction analysis, a recent stability theory for nonlinear systems, and constrained optimization theory. Although dynamic systems and optimization are both areas that have been extensively studied [21], few results have been achieved in this direction because strong enough tools for dynamic systems were not available. Contraction analysis provides the necessary mathematical background. Based on an operator that projects the speed of the system on the tangent space of the constraints, we derive generalizations of Lagrange parameters.

After presenting some initial examples that show the relations between contraction and optimization, we derive a contraction theorem for nonlinear systems with equality constraints. The method is applied to examples in differential geometry and biological systems. A new physical interpretation of Lagrange parameters is provided. In the autonomous case, we derive a new algorithm to solve minimization problems.

Next, we state a contraction theorem for nonlinear systems with inequality constraints. In the autonomous case, the algorithm solves minimization problems very fast compared to standard algorithms.

Finally, we state another contraction theorem for nonlinear systems with time-varying equality constraints. A new generalization of time varying Lagrange parameters is given. In the autonomous case, we provide a solution for a new class of optimization problems, minimization with time-varying constraints.

Thesis Supervisor: Jean-Jacques Slotine
Title: Professor of Mechanical Engineering;
Professor of Brain and Cognitive Sciences

Acknowledgments

I wish to express my deepest gratitude to my advisor, Professor Jean-Jacques Slotine. He was a constant source of support, ideas and excellent guidance.

Special thanks to the people in the Laser Interferometry Gravitational Observation laboratory and also to all my friends at MIT.

Finally, I am most grateful to my family and all my friends in France and Spain, who supported and encouraged my work.

Support for this work was provided by Laser Interferometry Gravitational Observation (LIGO) at MIT.

To my parents.

Contents

1 Preliminaries **11**

1.1 Contraction theory 11

1.2 Optimization theory 13

1.2.1 Unconstrained optimization 13

1.2.2 Constrained optimization with equality constraints 13

1.2.3 Constrained optimization with inequality constraints 14

1.3 Some examples linking both theories 15

2 Contraction theory with equality constraints **19**

2.1 Starting on the constraints:
contraction theory 19

2.1.1 Theorem 20

2.1.2 Proof 20

2.1.3 Theorem's application 23

2.1.4 An algorithm for solving minimization problems 26

2.2 Starting outside the constraints:
sliding behavior 37

2.2.1 The problem and intuition of the solution 37

2.2.2 Sliding surface behavior 39

2.2.3 Examples 40

3 Contraction theory with inequality constraints **47**

3.1 Starting on the constraints:
contraction theory 47
3.1.1 Theorem 48
3.1.2 Proof 48
3.1.3 Theorem's application 50
3.1.4 An algorithm for solving minimization problems 52
3.2 Starting outside the constraints:
one-side sliding behavior 55
3.2.1 One-side sliding condition 55
3.2.2 One side sliding surface behavior 56
3.2.3 Examples 57

4 Contraction theory with time-varying equality constraints 63
4.1 Starting on the constraints:
contraction theory 63
4.1.1 Theorem 63
4.1.2 Proof 64
4.1.3 Theorem's applications 66
4.2 Starting outside the constraints:
sliding behavior 68
4.2.1 Sliding surface behavior 68
4.2.2 Examples 70

5 Conclusion and directions for future research 73
5.1 Conclusion 73
5.2 Directions for future research 74

List of Figures

2-1 Projected speed of the dynamic system 21
2-2 Minimization of the length of a square in a circle over time 27
2-3 Minimization of the length of a square in a circle in 2D 28
2-4 Phase plane of the length minimization of a square in a circle 29
2-5 Minimum distance from a point to a sphere 31
2-6 Phase plane of the minimum distance from a point to a sphere 32
2-7 Minimum distance from a point to an ellipse 33
2-8 Maximum volume inside an ellipse 34
2-9 Minimal sum of lengths for a given volume 35
2-10 Contraction behavior 36
2-11 No contraction behavior 36
2-12 Example of minimization with two constraints 37
2-13 Minimization on a torus 38
2-14 Sliding dynamic 38
2-15 Minimization of the length of a square in a circle 40
2-16 Minimum distance from a point to a sphere 42
2-17 Maximum volume on an ellipse 43
2-18 Minimal sum of lengths for a given volume 44
2-19 Contraction of the initial and projected system 45
2-20 Example of minimization with two constraints 46

3-1 Minimization of the length of a square inside a circle with two different boundary layers 53

3-2 Solution to Kepler planimetric problem with two different boundary layers 53
3-3 Minimum distance from a point inside a sphere 54
3-4 Maximum volume on an ellipse with two different boundary layers . . 55
3-5 Minimization with two constraints with two different boundary layers 56
3-6 Minimization of the length of a square inside a circle with two different boundary layers 58
3-7 Solution to Kepler planimetric problem with two different boundary layers 58
3-8 Minimum distance from a point inside a sphere 59
3-9 Maximum volume on an ellipse with two different boundary layers . . 60
3-10 Minimization with two constraints with two different boundary layers 61

4-1 Projection with time-varying constraints 64
4-2 Minimization of the sum of the sides under a growing circle 67
4-3 Minimization of the sum of the sides under a changing ellipse 68
4-4 Using time varying constraints to solve problems with static constraints 69
4-5 Minimization of the sum of the sides under a growing circle 70
4-6 Minimization of the sum of the sides under a changing ellipse 71

Chapter 1

Preliminaries

This thesis explores the relation between nonlinear contraction analysis and optimization theory. First, we introduce the basic idea and theorem of contraction analysis. Next, we state the theorems of minimization in the three different cases - unconstrained, constrained with equality constraints and constrained with inequality constraints. The last part gives different examples in order to show that both theories are linked in some simple cases.

1.1 Contraction theory

Intuitively, contraction analysis is based on a slightly different view of what stability is. Stability is generally viewed relative to some nominal motion or equilibrium point. Contraction analysis is motivated by the elementary remark that talking about stability does not require to know what the nominal motion is: intuitively, a system is stable in some region if initial conditions or temporary disturbances are somehow "forgotten", i.e., if the final behavior of the system is independent of the initial conditions. All trajectories then converge to the nominal motion. In turn, this shows that stability can be analyzed differentially rather than through finding some implicit motion integral as in Lyapunov theory, or through some global state transformation as in feedback linearization. To avoid any ambiguity, we shall call "convergence" this form of stability.

In this section, we summarize the variational formulation of contraction analysis of [1]. It is a way to prove the contraction of a whole system by analyzing the properties of its Jacobian only. Consider a n-dimensional time-varying system of the form:

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t) \tag{1.1}$$

where $\mathbf{x} \in R^n$ and $t \in R^+$ and f is $n \times 1$ nonlinear vector function which is assumed to be real and smooth in the sense that all required derivatives exist and are continuous.

Before stating the main contraction theorem, recall first the following. The symmetric part of a matrix $\mathbf{A}$ is $\mathbf{A}_H = \frac{1}{2}(\mathbf{A} + \mathbf{A}^{*T})$. A complex square matrix A is *Hermitian* if $\mathbf{A}^\mathbf{T} = \mathbf{A}^*$, where ^{T} denotes matrix transposition and * complex conjugation. The *Hermitian part* $\mathbf{A_H}$ of any complex square matrix $\mathbf{A}$ is the Hermitian matrix $\frac{1}{2}(\mathbf{A} + \mathbf{A^{*T}})$. All eigenvalues of a Hermitian matrix are *real* numbers. A Hermitian matrix $\mathbf{A}$ is said to be *positive definite* if all its eigenvalues are strictly positive. This condition implies in turn that for any non-zero real or complex vector $\mathbf{x}$, $\mathbf{x}^{*T}\mathbf{A}\mathbf{x} > 0$. A Hermitian matrix $\mathbf{A}$ is called *negative definite* if $-\mathbf{A}$ is positive definite.

A Hermitian matrix dependent on time or state will be called *uniformly* positive definite if all its eigenvalues remain larger than strictly positive constant for all states and all $t \geq 0$. A similar definition holds for uniform negative definiteness.

The main result of contraction analysis is given by theorem 1

Theorem 1. *Denote by* $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ *the Jacobian matrix of* $\mathbf{f}$ *with respect to* $\mathbf{x}$. *Assume that there exists a complex square matrix* $\mathbf{\Theta}(\mathbf{x}, \mathbf{t})$ *such that the Hermitian matrix* $\mathbf{\Theta}(\mathbf{x}, \mathbf{t})^{*\mathbf{T}}\mathbf{\Theta}(\mathbf{x}, \mathbf{t})$ *is uniformly positive definite, and the Hermitian part* $\mathbf{F}_H$ *of the matrix*

$$\mathbf{F} = \left(\dot{\mathbf{\Theta}} + \mathbf{\Theta} \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right) \mathbf{\Theta}^{-1}$$

is uniformly negative definite. Then, all system trajectories converge exponentially to a single trajectory, with convergence rate $|\sup_{\mathbf{x},\mathbf{t}} \lambda_{\max}(\mathbf{F}_H)| > 0$. *The system is said to be* contracting, $\mathbf{F}$ *is called its* generalized Jacobian, *and* $\mathbf{\Theta}(\mathbf{x}, \mathbf{t})^{*\mathbf{T}}\mathbf{\Theta}(\mathbf{x}, \mathbf{t})$ *its contraction* metric. *The contraction rate is the absolute value of the largest eigenvalue* *(closest to zero, although still negative)* $\lambda = |\lambda_{max}(\mathbf{F}_H)|$.

1.2 Optimization theory

Optimization means finding "best available" values of some objective function given a defined domain, including a variety of different types of objective functions and different types of domains.

We are interested in the most general case of optimization, nonlinear optimization. In the following we give the sufficient conditions for the point $\mathbf{x}^*$ to be a minimum of a minimization problem. There are three main classes of problems which are stated below.[2]

1.2.1 Unconstrained optimization

In unconstrained optimization we search for the minimum over the whole range of variables. The generic form is:

$$min\ U(\mathbf{x}) \tag{1.2}$$

The second-order sufficient conditions in the unconstrained case are given by theorem 2.

Theorem 2. *Let $U \in C^2(R^n, R)$ be a function defined on a region in which the point $\mathbf{x}^*$ is an interior point. Suppose in addition that*

$$\nabla U(\mathbf{x}^*) = 0 \ \ and\ \nabla^2 U(\mathbf{x}^*)\ is\ positive\ definite$$

Then x^* is a solution of problem 1.2.

1.2.2 Constrained optimization with equality constraints

In constrained optimization, the Karush-Kuhn-Tucker conditions (also known as the Kuhn-Tucker or KKT conditions) are sufficient for a solution in nonlinear programming to be optimal, provided that some regularity conditions are satisfied. The generic form is with m constraints is:

$$min\ U(\mathbf{x}) \quad subject\ to\ (s.t) \quad \mathbf{h}(\mathbf{x}) = \mathbf{0} \tag{1.3}$$

The second-order sufficient conditions in the constrained case are given by theorem 3.

Theorem 3. *Let $U(\mathbf{x}) \in C^2(R^n, R)$ and $\mathbf{h}(\mathbf{x}) \in C^2(R^n, R^m)$ be smooth functions and $\mathbf{x}^*$ is a regular point (the columns of $\nabla h(\mathbf{x}^*)$ are linearly independent). Suppose in addition that*

$$\nabla L(\mathbf{x}^*) = 0 \ and\ \mathbf{y}'\nabla^2 L(\mathbf{x}^*)\mathbf{y} > 0$$

where $\mathbf{y}\nabla h(\mathbf{x}^*) = 0$, λ are the lagrange parameters and L is the lagrangian function defined as $L = U + \lambda\mathbf{h}$
Then $\mathbf{x}^*$ is a solution of problem 1.3.

1.2.3 Constrained optimization with inequality constraints

The generic form of the third kind of problems is:

$$min\ U(\mathbf{x}) \quad subject\ to\ (s.t) \quad \mathbf{h}(\mathbf{x}) \preceq \mathbf{0} \tag{1.4}$$

where $\mathbf{x} \preceq \mathbf{0}$ means that all the scalar components of $\mathbf{x}$ are less or equal to zero. The second-order sufficient conditions in the constrained case are given by theorem 4.

Theorem 4. *Let $U(\mathbf{x}) \in C^2(R^n, R)$ and $\mathbf{h}(\mathbf{x}) \in C^2(R^n, R^m)$ be smooth functions, $A(\mathbf{x})$ the set of active constraints at point $\mathbf{x}$ and $\mathbf{x}^*$ a regular point (the columns of $\nabla h(x^*)$ are linearly independent). Suppose in addition that*

$$\nabla L(\mathbf{x}^*) = 0\ ,\ y'\nabla^2 L(\mathbf{x}^*)y > 0 \ and\ \lambda_i^* \geq 0$$

where $y\nabla h(x^*) = 0$, λ^* are the lagrange parameters and L is the lagrangian function defined as $L = U + \lambda\mathbf{h}$. Also if $i \in A(\mathbf{x}^*)$ then $\lambda^* > 0$ and if $i \notin A(\mathbf{x}^*)$ then $\lambda^* = 0$
Then $\mathbf{x}^*$ is a solution of problem 1.4.

1.3 Some examples linking both theories

In this section we present some cases - unconstrained optimization, duality theory - in which both theories are linked. The following example in [3] is the starting point of this thesis.

Example 1.3.1.: Consider a gradient autonomous system $\dot{\mathbf{x}} = -\nabla U(\mathbf{x})$ that is contracting in an identity metric. As it is autonomous and contracting in a time independent metric then it has a unique equilibrium point because

$$\frac{d}{d\mathbf{x}}(\Theta\nabla U) = F(\Theta\nabla U)$$

which implies exponential convergence of $\dot{\mathbf{x}}$ to zero and so $\mathbf{x}$ converges to a constant, $\mathbf{x}^*$. This point has zero speed, $\nabla U(\mathbf{x}^*) = 0$. The condition of contraction at this equilibrium point is the positive definiteness of $\nabla^2 U(\mathbf{x}^*)$. These are the two sufficient conditions (theorem 2) to prove that $\mathbf{x}^*$ is a solution to problem 1.2 □

With example 1.3.1, we show that contraction and unconstrained optimization are linked through gradient systems. Another important domain in optimization is duality. It is very useful to use the dual formulation instead of the primal for a variety of reasons, sometimes the dual problem has closed form solution and others the algorithm to find the minimum is much faster. If the primal problem is dimensionally low with lots of constraints, the dual problem is dimensionally big but with few constraints. In example 1.3.2, we show that contraction and duality are related using gradient systems.

The Legendre transformation is defined as follows for $y \in R^n$:

$$U^*(\mathbf{y}) = sup_{\mathbf{x} \in R^n}(\mathbf{x}'\mathbf{y} - U(\mathbf{x})) \tag{1.5}$$

The conjugate function U^* is a convex function since it is the pointwise supremum of linear functions.

Example 1.3.2.: The system $\dot{\mathbf{x}} = -\nabla U(\mathbf{x})$ is contracting for the identity metric and also smooth (existence of derivatives) if and only if the system $\dot{\mathbf{y}} = -\nabla U^*(\mathbf{y})$ is contracting for the identity metric and also smooth.

To prove the if part we use the fact that strict convexity and smoothness are dual properties, [22].

For the only if part, we apply the first part of the proof to the function U^*. We use the fact that as U^* convex, close (the domain of U is closed) and proper (it never takes the value $-\infty$ and the set $dom_g = [x|g(x) < \infty]$ is nonempty), then $U^{**} = U$ □

Contraction theory is also related to differential geometry through example 1.3.3. [18]

Example 1.3.3.: The condition of contraction of the system 1.1 for a positive definite metric $\mathbf{g} = \mathbf{\Theta}'\mathbf{\Theta}$ that verifies a parallel propagation is that the hermitian part of :

$$\mathbf{F} = \mathbf{\Theta} D\mathbf{f}(\mathbf{x}, t)\mathbf{\Theta}^{-1} = \mathbf{\Theta}\left(\frac{\partial \mathbf{f}(\mathbf{x}, t)}{\partial \mathbf{x}} + \sum_i \mathbf{\Gamma}_i \mathbf{f}^i(\mathbf{x}, t)\right)\mathbf{\Theta}^{-1}$$

has a uniformly negative maximum eigenvalue. D is the covariant derivative and the Christoffel term is defined as $\sum_k \Gamma^k_{ij} g_{kl} = \frac{1}{2}(\frac{\partial g_{il}}{\partial q^j} + \frac{\partial g_{jl}}{\partial q^i} - \frac{\partial g_{ij}}{\partial q^l})$.

The idea is to use a parallel propagation of the tensor $\mathbf{\Theta}$, [18].

$$\dot{\mathbf{\Theta}} = \mathbf{\Theta} \sum \mathbf{\Gamma}_{ij} \dot{\mathbf{x}}_i$$

□

The geometric interpretation of the covariant derivative is the projection of the directional derivative [7] on that submanifold, thus the tangential part of the directional derivative. It can be written as:

$$D_X Y = (d_X Y)^{tang} = d_X Y - < d_X Y, \nu > \nu$$

where $d_X Y$ is the directional derivative along the X direction and ν is the normal vector to the submanifold.

When writing $\delta \dot{z} = \Theta D\mathbf{f}\Theta^{-1}\delta z$, it means that the virtual speed is constrained to belong to 'some' manifold. This idea of projection is fundamental to prove contraction for constrained dynamical systems.

In [1] combination properties of contracting systems are of great importance. In example 1.3.4 we derive a new result that cannot be achieved with classical minimization theory.

Example 1.3.4.: If there are two unconstrained minimizations which have respectively their minimum at x_1^* and x_2^*, the sum of both cost functions may not have a minimum. Using only classical minimizations theorems we cannot conclude anything when summing both systems.

If the gradient systems issued from the two precedent minimization problems are contracting. Then the system, sum of the gradient systems, is still contracting. Then the sum of the two cost functions has a minimum $\square$

Chapter 2

Contraction theory with equality constraints

This section has two main parts. In the first part, we state a theorem giving the condition of contraction for a constrained nonlinear dynamic system. We have to start on the constraints to have contraction behavior. In the second part, we allow the system to start outside the constraints. We have to use sliding techniques to conclude convergence to the constraints.

In the following all the points are regular. This means that $\nabla h(\mathbf{x})\nabla h'(\mathbf{x})$ has full rank and hence is invertible. That is the *only* hypothesis that is needed in the following analysis.

2.1 Starting on the constraints: contraction theory

We define the sets $S = [\mathbf{x} \ / \ h(\mathbf{x}) = 0]$ and $M(\mathbf{x}) = [y \ / \ y\nabla h(\mathbf{x}) = 0]$. S is the space of the constraints. M is the tangent subspace of $\mathbf{h}$ at point $\mathbf{x}$, it is a k-submanifold. (a k submanifold is defined as $\nabla h(\mathbf{x})$ having full rank using the implicit function theorem)

We define the operator [4]

$$P(\mathbf{x}) = 1 - \nabla h'(\mathbf{x})[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1}\nabla h(\mathbf{x})$$

It is a symmetric projection operator onto $M(\mathbf{x})$ because it verifies $P(\mathbf{x})P(\mathbf{x}) = P(\mathbf{x})$ and $P(\mathbf{x})y \in M(\mathbf{x})$.

2.1.1 Theorem

In this section we compute the condition of contraction for dynamic systems that are under some constraints. In the real world, there are many examples. A robot arm that has to evolve in a circle is a constrained dynamic system. The concentrations of chemical systems that have their own dynamics but cannot exceed one is another example.

Theorem 5. *The condition of contraction of the constrained dynamic system*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \ \ s.t. \ \mathbf{h}(\mathbf{x}) = 0$$

is that the hermitian part of :

$$\mathbf{F} = \left(\dot{\mathbf{\Theta}}(\mathbf{x}, t) + \mathbf{\Theta}(\mathbf{x}, t)P(\mathbf{x})\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}} + \lambda(\mathbf{x}, t)\frac{d^2\mathbf{h}}{d\mathbf{x}^2}\right)\right)\mathbf{\Theta}(\mathbf{x}, t)^{-1} \quad on \ S(\mathbf{x})$$

has a maximum eigenvalue uniformly negative. $P(\mathbf{x})$ *is a projection operator onto the tangent space* $M(\mathbf{x})$ *and* $\lambda(\mathbf{x}, t) = -\mathbf{f}(\mathbf{x}, t)\nabla h'[\nabla h \nabla h']^{-1}$. *The initial condition must verify the constraint.*

2.1.2 Proof

The proof has three parts. We project the system by adding a term related to the Lagrange parameters. We compute the condition of contraction of this new system. We then prove contraction behavior.

Projection of the dynamic system

The initial problem is

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \ s.t. \ h(\mathbf{x}) = 0$$

Physically this can be interpreted as the velocity of the system goes away from the constraints. For the system to verify the constraints, the velocity has to be tangent to the constraints at $\mathbf{x}$. We project the system onto M($\mathbf{x}$) using the projector P($\mathbf{x}$).

$$\dot{\mathbf{x}} = P(\mathbf{x})\mathbf{f}(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t) - \mathbf{f}(\mathbf{x}, t)\nabla h'(\mathbf{x})[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1}\nabla h(\mathbf{x}) = f(\mathbf{x}, t) + \lambda(\mathbf{x}, t)\nabla h(\mathbf{x})$$

We note $\lambda(\mathbf{x}, t) = -f(\mathbf{x}, t)\nabla h'(\mathbf{x})[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1}$. These are not exactly the Lagrange parameters. But if $\mathbf{f} = \nabla U(\mathbf{x})$, lambda is exactly the Lagrange parameters. This can be seen as a first generalization of the Lagrange parameters. In this case the dynamical system would be $\dot{\mathbf{x}} = \nabla U(\mathbf{x}) + \lambda(\mathbf{x})\nabla h(\mathbf{x}) = \nabla L(\mathbf{x})$ where L is the usual Lagrangian function. In the following the projected system will be known as Lagrangian dynamics.

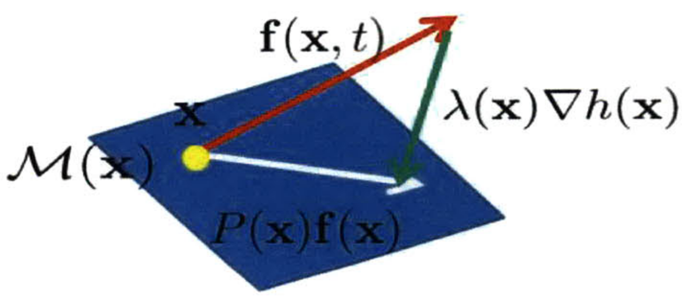

Figure 2-1: Projected speed of the dynamic system

This result gives a new insight about Lagrange parameters. The term $\lambda(\mathbf{x}, t)\nabla h(\mathbf{x})$ can be seen as a reaction force that allows the system to stay on the constraints . A Lagrange parameter is the scalar value that gives the magnitude of the force along the orthogonal direction to the constraint in order to have a tangential speed. Lagrange parameters as reaction forces is something that has already been introduced using the Lagrange equation [6] [5]. The constraints are verified at anytime because the system starts on the constraints and the velocity is always tangential to the constraints. If initially the constraints were not verified then this property would not be true.

Condition of contraction

We consider a virtual displacement $\delta\mathbf{x}$ and a positive definite $\mathbf{\Theta}(\mathbf{x},t)$. The function $\delta\mathbf{x}'\mathbf{\Theta}(\mathbf{x},t)'\mathbf{\Theta}(\mathbf{x},t)\delta\mathbf{x}$ is the distance associated to the first fundamental form of the k-submanifold $M(\mathbf{x})$. We define $\delta\mathbf{z} = \mathbf{\Theta}(\mathbf{x},t)\delta\mathbf{x}$.

We compute $\frac{d}{dt}(\delta\mathbf{z}\delta\mathbf{z}) = 2\delta\mathbf{x}\delta\dot{\mathbf{x}}$. We calculate $\delta\dot{\mathbf{z}} = \dot{\mathbf{\Theta}}(\mathbf{x},t)\delta\mathbf{x} + \mathbf{\Theta}(\mathbf{x},t)\delta\dot{\mathbf{x}}$ where $\delta\dot{\mathbf{x}} = \left(\frac{\partial f}{\partial\mathbf{x}} + \lambda(\mathbf{x},t)\frac{d^2h}{d\mathbf{x}^2} + \frac{\partial\lambda(\mathbf{x},t)}{\partial\mathbf{x}}\frac{dh}{d\mathbf{x}}\right)\delta\mathbf{x}$

We have to compute a long calculation $\frac{\partial\lambda(\mathbf{x},t)}{\partial\mathbf{x}}$.

$$\frac{\partial\lambda(\mathbf{x},t)}{\partial\mathbf{x}} = -\left(\frac{\partial f}{\partial\mathbf{x}} + \lambda(\mathbf{x},t)\frac{d^2h}{d\mathbf{x}^2}\right)\nabla h'[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1} - \dot{\mathbf{x}}\frac{d^2h}{d\mathbf{x}^2}[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1}$$

We can substitute in the precedent equation and reorganize:

$$\delta\dot{\mathbf{x}} = P(\mathbf{x})\left(\frac{df}{d\mathbf{x}} + \lambda(\mathbf{x},t)\frac{d^2h}{d\mathbf{x}^2}\right)\delta\mathbf{x} - \nabla h(\mathbf{x})'[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1}\delta\mathbf{x}\frac{d^2h}{d\mathbf{x}^2}\dot{\mathbf{x}} = \delta\dot{\mathbf{x}}^{\parallel} + \delta\dot{\mathbf{x}}^{\perp} \tag{2.1}$$

It is essential to note that the first term, $\delta\dot{\mathbf{x}}^{\parallel}$, belongs to $M(\mathbf{x})$ and the second term, $\delta\dot{\mathbf{x}}^{\perp}$, belongs to $M(\mathbf{x})^{\perp}$. In order to have contraction behavior the first term of $\delta\dot{\mathbf{x}}$ must be uniformly bounded.

$$\delta\dot{\mathbf{x}} = P(\mathbf{x})\left(\frac{\partial f}{\partial\mathbf{x}} + \lambda(\mathbf{x},t)\frac{d^2h}{d\mathbf{x}^2}\right)\delta\mathbf{x}$$

Finally

$$\frac{d}{dt}(\delta\mathbf{z}\delta\mathbf{z}) = 2\delta\mathbf{z}\left(\dot{\mathbf{\Theta}}(\mathbf{x},t) + \mathbf{\Theta}(\mathbf{x},t)P(\mathbf{x})\left(\frac{\partial f}{\partial\mathbf{x}} + \lambda(\mathbf{x},t)\frac{d^2h}{d\mathbf{x}^2}\right)\right)\mathbf{\Theta}(\mathbf{x},t)^{-1}\delta\mathbf{z}$$

Shrinking behavior

From the precedent section we have that:

$$\frac{d}{dt}(\delta\mathbf{z}\delta\mathbf{z}) = 2\delta\mathbf{z}F\delta\mathbf{z}$$

The contraction condition is that $\forall t > 0$ $2\delta\mathbf{z}F\delta\mathbf{z} \leq -\Lambda(\delta\mathbf{z}\delta\mathbf{z})$ for $\Lambda > 0$. We then conclude that $|\delta\mathbf{z}| \to 0$. As the metric is positive definite we have that $|\delta\mathbf{x}| \to 0$

2.1.3 Theorem's application

This theorem relates different areas of mathematics and also has some practical applications.

Differential geometry and projected contraction theory

When calculating the virtual speed of the system, we compute equation 2.1. The contraction behavior is given by the uniform negativity of the tangential component of the speed. In example 1.3.3, the condition of contraction is given by the covariant derivative. The definition of the covariant derivative is the tangential part of the directional derivative. In the following we achieve the same expression for the geodesic using two different theories.

Consider an autonomous system (so we can substitute the partial derivatives by absolute derivatives), one way of defining geodesics is $\ddot{\mathbf{x}}^{\|} = 0$ [7]. Using equation 2.1, we obtain equation 2.2

$$\ddot{\mathbf{x}} = \ddot{\mathbf{x}}^{\perp} = -\nabla h^{'}(\nabla h \nabla h^{'})^{-1}\dot{\mathbf{x}}\nabla^2 h\dot{\mathbf{x}} \tag{2.2}$$

Another equivalent way of defining geodesics is $\ddot{\mathbf{x}} = \omega\nabla h$, [4]. Differentiating $h(\mathbf{x}) = 0$ two times we find that $\dot{\mathbf{x}}\nabla^2 h\dot{\mathbf{x}} + \nabla h\ddot{\mathbf{x}} = 0$. Using the second definition of geodesic we calculate

$$\ddot{\mathbf{x}} = -\nabla h^{'}(\nabla h \nabla h^{'})^{-1}\dot{\mathbf{x}}\nabla^2 h\dot{\mathbf{x}} \tag{2.3}$$

Equations 2.2 and 2.3 are the same but obtained in a different way.

Control theory and projected contraction theory

The major goal of control theory is to find the control input that will make the system behave the way we want. In mathematical terms, find $\hat{\mathbf{u}}$ such that

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \mathbf{u}(\mathbf{x}, t)$$

has a specific behavior. In the case of projected contraction theory identifying $\hat{\mathbf{u}}(\mathbf{x}, t) = \lambda \nabla \mathbf{h}$, we find a control that makes the system evolve on the constraint.

Optimization and projected contraction theory

In this section we show that a gradient autonomous constrained contracting system (equation 2.4) for an identity metric has a unique equilibrium point and this equilibrium point is also the solution of a problem of minimization.

As done in example 1.3.1 an autonomous contracting dynamic system has $\dot{\mathbf{x}}$ uniformly going to zero, then $\mathbf{x}$ goes to a constant $\mathbf{x}^*$. Then the gradient autonomous dynamic system has a unique equilibrium point.

$$\dot{\mathbf{x}} = -\nabla U(\mathbf{x}) \ s.t. \ h(\mathbf{x}) = 0 \tag{2.4}$$

The projected system, which we will call Lagrangian dynamic, is $\dot{\mathbf{x}} = -\nabla L(\mathbf{x})$. It verifies $\nabla L(\mathbf{x}^*) = 0$. The dynamic system also satisfies the constraint $h(\mathbf{x}^*) = 0$ because we start on the constraint and the speed of the system is always tangential to the constraints. These two conditions ensure us that at anytime $h(\mathbf{x}) = 0$.

As the system is autonomous the condition of contraction is given by

$$\ddot{\mathbf{x}} = -P(\mathbf{x})\nabla^2 L(\mathbf{x})\dot{\mathbf{x}}$$

As the speed of the system belongs to $M(\mathbf{x})$, we have : $\dot{\mathbf{x}} = P(\mathbf{x})\dot{\mathbf{x}}$. Thus: $\ddot{\mathbf{x}} =$

$-P(\mathbf{x})\nabla^2 L(\mathbf{x})P(\mathbf{x})\dot{\mathbf{x}}$. The condition of contraction then gives

$$\mathbf{x}^{*\prime} P'(\mathbf{x}^*) \left(\frac{d^2U}{d\mathbf{x}^2} + \lambda(\mathbf{x}^*) \frac{d^2h}{d\mathbf{x}^2} \right) P(\mathbf{x}^*)\mathbf{x}^* = \mathbf{y}'\nabla^2 L(\mathbf{x}^*)\mathbf{y} > \eta > 0$$

with $\mathbf{y} = \mathbf{x}P(\mathbf{x}^*) \in M(\mathbf{x}^*)$ by definition of $P(\mathbf{x}^*)$.

These are the three conditions for the existence of a minimum to problem 1.3 . Therefore, there is a new relationship between minimization and dynamical systems using contraction theory. This is something that has been already investigated [8].

Similarity with precedent results

This result has been already extensively searched. In [1], the calculation of the derivative of the distance is given by

$$\frac{1}{2}\frac{d}{dt}(\delta\mathbf{x}^T\mathbf{M}\delta\mathbf{x}) = \delta\mathbf{x}^T \frac{\partial\mathbf{z}}{\partial\mathbf{x}}^T \frac{\partial\mathbf{f}}{\partial\mathbf{z}} \frac{\partial\mathbf{z}}{\partial\mathbf{x}}\ \delta\mathbf{x} - \delta\mathbf{x}^T \frac{\partial^2\mathbf{z}}{\partial\mathbf{x}^2}\ \mathbf{n}\ \delta\mathbf{x}$$

A similar result has been obtained in the proof of the theorem when writing equation 2.1. In both cases, the first term involves the Jacobian of the function. The second term involves the second derivative of the constraints and is along the direction perpendicular to the constraints. The final condition of contraction is on the projection of the Jacobian. The great novelty about this theorem is the introduction of the Lagrange parameters in the condition of contraction.

In [9] a similar intuition is developed. If an explicit $\mathbf{z}$ coordinates exists then we can write the dynamical system for convex constraints as

$$\dot{\mathbf{z}} = \mathbf{\Theta}\mathbf{f} - [\nabla_z\mathbf{h}\nabla_z\mathbf{h}']^{-1}\nabla_z\mathbf{h}'\nabla_z\mathbf{h}\mathbf{\Theta}\mathbf{f}$$

We recognize the generalized Lagrange parameters if $\lambda = -[\nabla_z\mathbf{h}\nabla_z\mathbf{h}']^{-1}\nabla_z\mathbf{h}'\mathbf{f}$ if $\mathbf{\Theta} = I$.

Application to biological systems

In [10], simulations of biological evolution, in which computers are used to evolve systems toward a goal, often require many generations to achieve even simple goals. In the article, the impact of temporally varying goals on the speed of evolution is studied. It is much faster when using time varying goals.

The highest speedup is found under modularly varying goals, in which goals change over time such that each new goal shares some of the subproblems with the previous goal.

This problem has strong similarities with what is developed in this thesis. We can write a non autonomous energy like function, and create a gradient system. This dynamical system searches the minimum of this function. The minimum is changing over time (time varying goal also called trajectory). Contraction theory also allows modularity.

2.1.4 An algorithm for solving minimization problems

This idea of projection creates a new dynamical system. It is a new way of solving minimization problems. This method is general, the only condition required is as stated at the beginning of chapter 2 is that $\nabla\mathbf{h}(\mathbf{x})\nabla\mathbf{h}(\mathbf{x})'$ must be invertible for $\mathbf{x} \in \mathbf{S}(\mathbf{x})$.

We give some examples. The first example is done in great detail to show how the method works.

Minimization of the length of a square in a circle

Consider :

$$min\ x + y\ s.t.\ x^2 + y^2 = 1$$

the optimal values are (from classical theory of optimization): $x^* = y^* = -\frac{1}{\sqrt{2}}$ and $\lambda^* = \frac{1}{\sqrt{2}}$. In terms of contraction the problem is $\dot{x} = -\nabla f = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$ $s.t.\ x^2+y^2 = 1$.

The projection matrix is $P(x) = \begin{pmatrix} 1-x^2 & -xy \\ -xy & 1-y^2 \end{pmatrix}$. The projected system is then :

$$\dot{x} = \begin{pmatrix} -1+x^2+xy \\ -1+y^2+xy \end{pmatrix}$$

where $\lambda = -\frac{\nabla f \nabla h'}{\nabla h \nabla h'} = -\frac{x+y}{2}$. The eigenvalue of F is $x+y$. The system is contracting in $x+y \leq 0$. As the system approaches the solution then $\lambda \rightarrow \lambda^*$

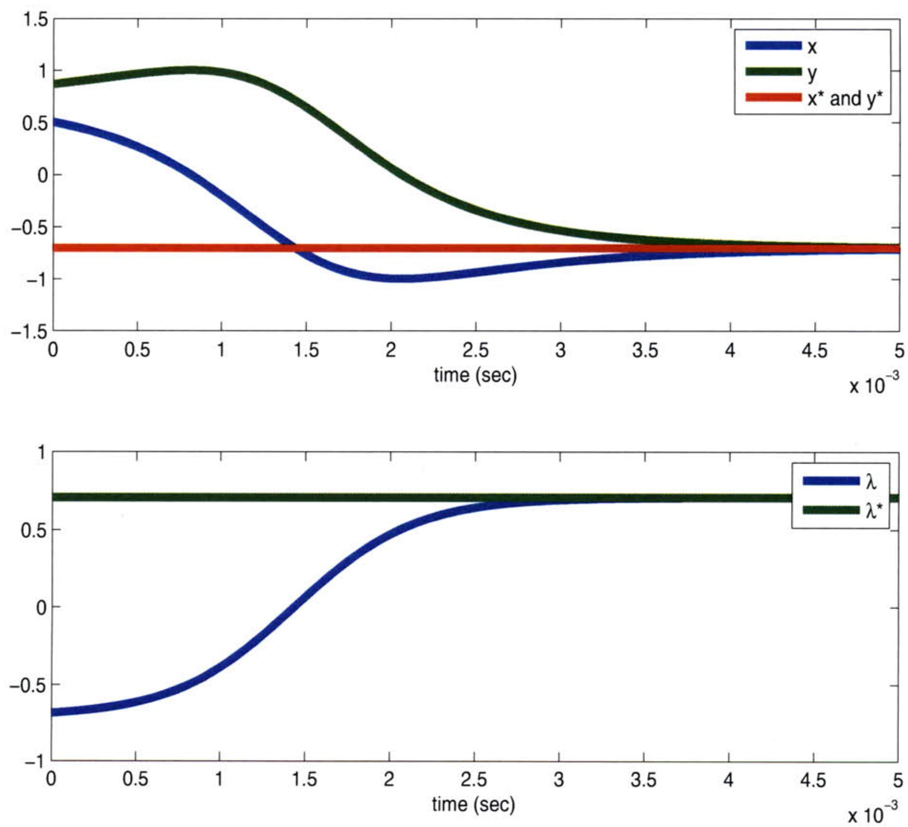

Figure 2-2: Minimization of the length of a square in a circle over time

As it can be simulated, figure $2-3$ shows that the dynamic system always stays on the constraint. The system still goes to the minimum even if it is not starting in the contraction zone as it can be seen on figure $2-4$ The contraction zone given by the precedent theorem is $x+y<0$. Looking at figure $2-4$, we see that the contraction zone can be extended to $x+y<1$.

Another important aspect is the time of computation. Using matlab normal rou-

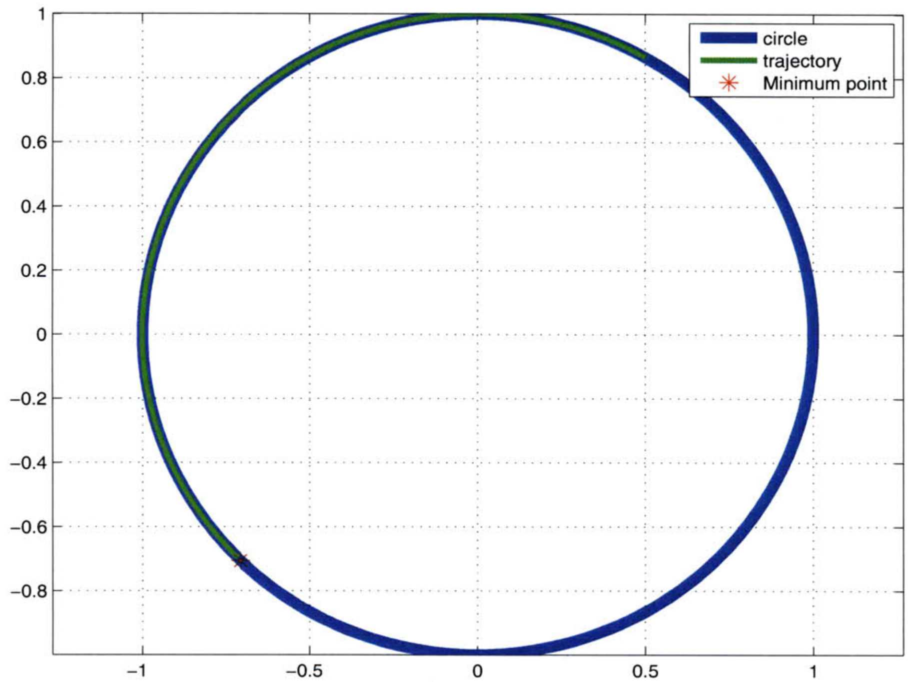

Figure 2-3: Minimization of the length of a square in a circle in 2D

tine 'fmincon' and starting at the same initial point we achieve the solution 1000 times faster than matlab. Matlab uses the sequential quadratic programming method. This method approximates at each step the actual lagrangian function by a second order approximation and solves the second order problem. Each time the system solves this problem it needs to invert two matrices [2]. In our case we only need to invert one.

Tartaglia's problem

Tartaglia's problem is a famous problem in optimization stated as : "To divide the number 8 into two parts such that the result of multiplying their product by their difference is maximal". The formulation is :

$$min\ xy(x-y)\ s.t\ x+y=8$$

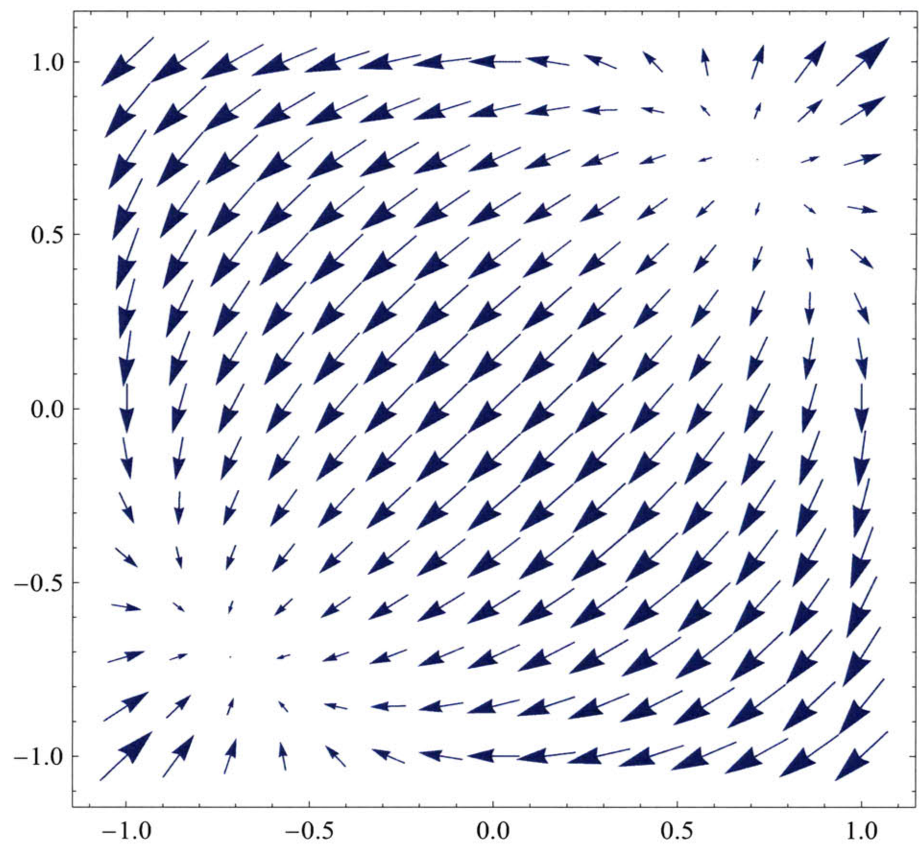

Figure 2-4: Phase plane of the length minimization of a square in a circle

The Lagrangian system is, with $\lambda = y^2 - x^2$

$$\dot{x} = \begin{pmatrix} 2xy - x^2 \\ y^2 - 2xy \end{pmatrix}$$

This algorithm starting on the constraint converges 1000 times faster than matlab (which again uses an SQP method).

Kepler's planimetric problem

The problem is how to inscribe in a given circle a rectangle of maximal area. The minimization formulation is :

$$min - xy \ s.t \ x^2 + y^2 = 1$$

In this case the Lagrangian system is

$$\dot{x} = \begin{pmatrix} y + 2\lambda x \\ x + 2\lambda y \end{pmatrix}$$

where $\lambda = -xy$. The condition of contraction gives two eigenvalues $-1-2xy, 1-2xy$. This algorithm converges to the expected solution, staying on the constraint. The convergence time is again 1000 times faster than matlab.

Minimum distance from a point to a sphere

Consider the minimization problem

$$min \ \frac{1}{2}(x^2 + y^2 + (z-2)^2) \quad s.t. \ \ h(x) = x^2 + y^2 + z^2 - 1 = 0$$

The evident solution of this problem is the projection of the point $(0,0,2)$ onto the sphere: $(0,0,1)$. We compute the usual projected system :

$$\dot{x} = \nabla U + \lambda \nabla h = \begin{pmatrix} x + \lambda x \\ y + \lambda y \\ z - 2 + \lambda z \end{pmatrix}$$

with $\lambda = 2z - 1$.

On figure $2-5$, the system evolves on the constraint. it converges towards the expected solution, which is the projection of the point $(0,0,2)$ onto the sphere.

Using mathematica we can plot the vector field. On figure $2-6$, there are two equilibrium points. The north pole is stable and the south pole is unstable. The condition of contraction is $z > 0$. But as it can be seen on the figure, we can extend it to $z > -1$. As usual the computation time, starting at the same initial point is 1000 times faster than matlab.

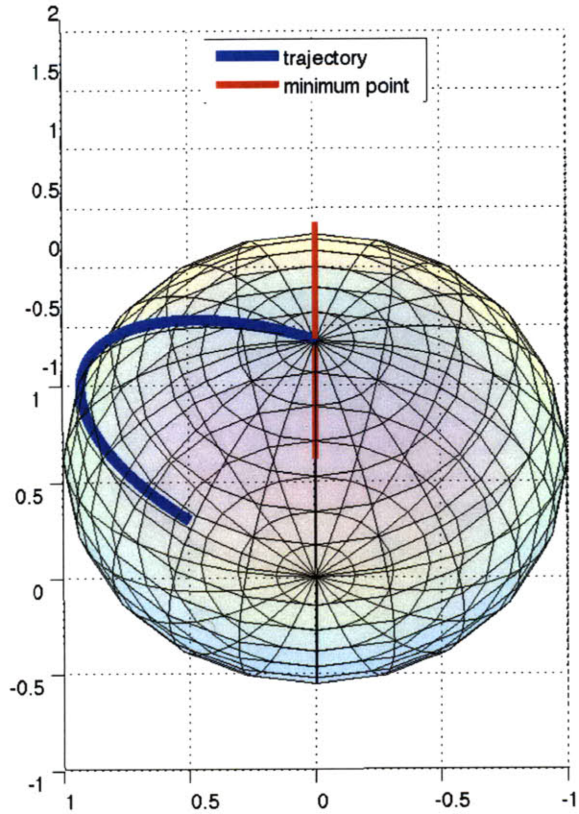

Figure 2-5: Minimum distance from a point to a sphere

Minimum distance from a point to an ellipse

This example is chosen to show that the lack of symmetry does not prevent the algorithm to converge to the right value. We want to calculate the distance from a point to the ellipse.

$$min\ \frac{1}{2}((x-x_1)^2+(y-y_1)^2+(z-z_1)^2) \quad s.t.\ \ h(x)=\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}-1=0$$

We consider $a = 15; b = 5; c = 3$; and the absolute minimum is $1, 4, 2$. On figure $2-7$ we achieve a solution which corresponds to the projection of the absolute minimum on the ellipsoid. (The red dotted line just indicates where is the projection point). As usual the convergence time is very fast compared to matlab.

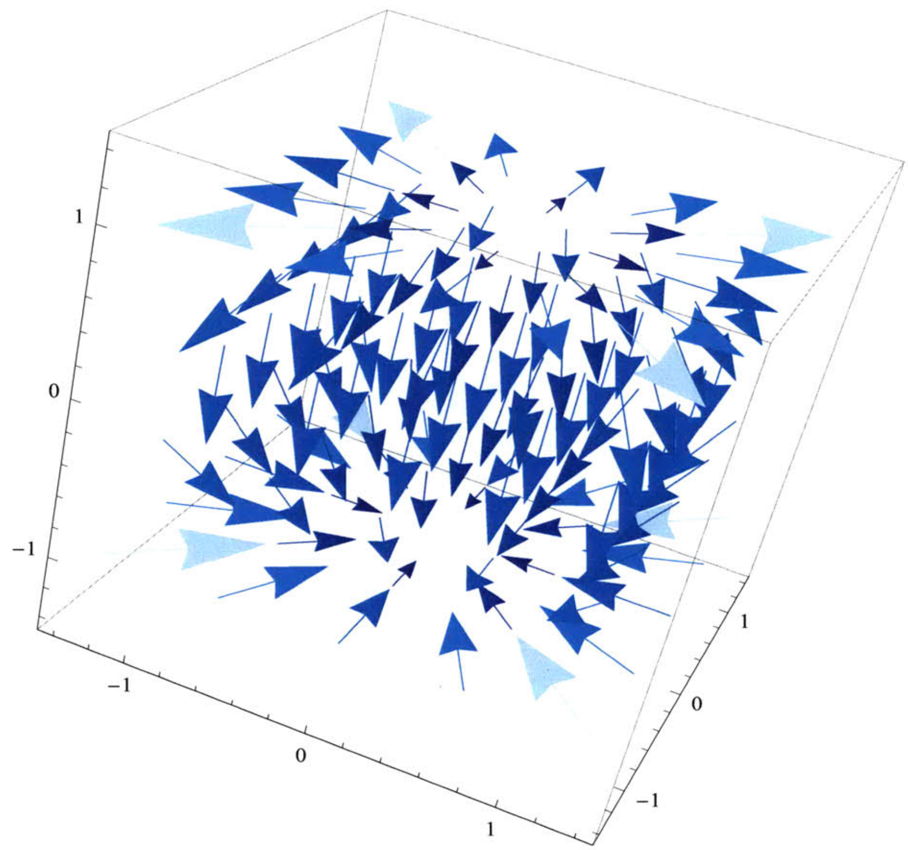

Figure 2-6: Phase plane of the minimum distance from a point to a sphere

Maximum volume inside an ellipse

The problem is to maximize the volume of a parallelipede subject to lying on an ellipse

$$min \; -xyz \quad s.t. \;\; h(x) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$$

We consider $a = 15; b = 5; c = 3;$. On figure $2-8$, we achieve the solution. As usual the convergence time is very fast compared to matlab.

Minimal sum of lengths for a given volume

The problem is to maximize the volume of a parallelipede subject to lying on ellipse

$$min \; x + y + z \quad s.t. \;\; h(x) = xyz - 1 = 0$$

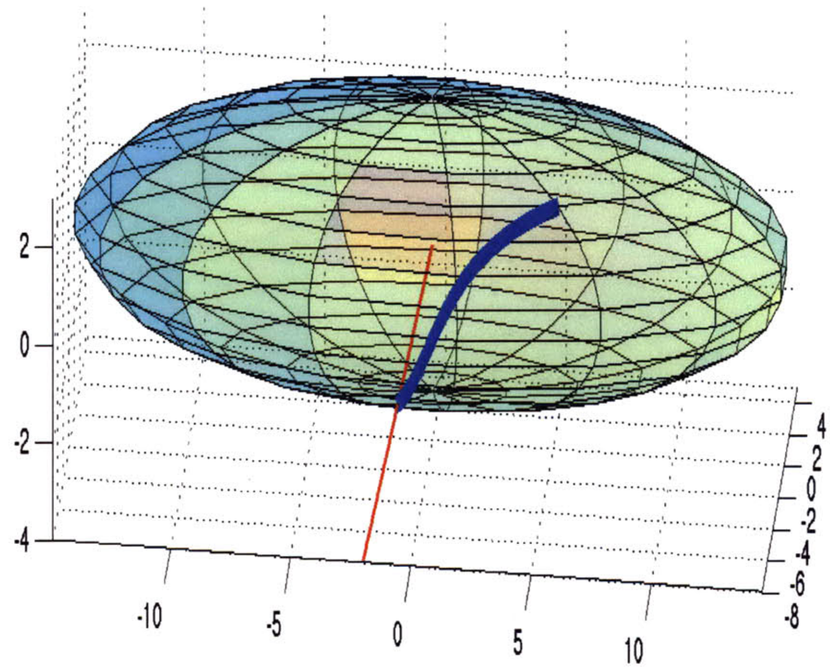

Figure 2-7: Minimum distance from a point to an ellipse

The constraints are completely nonlinear. The Lagrangian system is:

$$\dot{x} = \begin{pmatrix} 1 + \lambda yz \\ 1 + \lambda xz \\ 1 + \lambda xy \end{pmatrix}$$

where $\lambda = -\frac{yz+xz+xy}{(yz)^2+(xz)^2+(xy)^2}$. On figure $2-9$ we achieve the solution, [1 1 1]. As usual the convergence time is very fast compared to matlab.

Contraction of the initial and projected system

In this example, the initial system is not contracting and the projected system is contracting. From [11], consider the dynamic system

$$\dot{x} = [0, 0, -e^z] \; s.t. \; h(x, y, z) = \frac{x^2 + y^2}{2} - z = 0$$

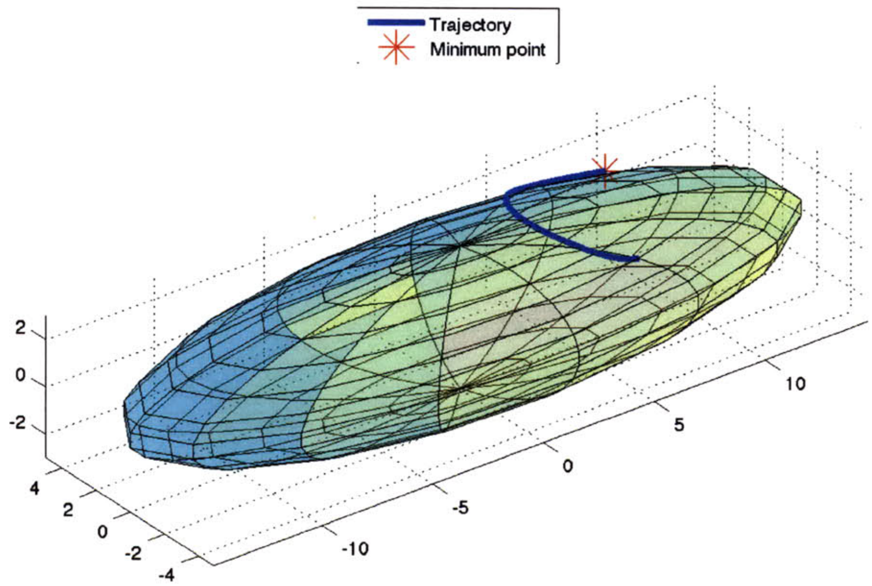

Figure 2-8: Maximum volume inside an ellipse

The constraint is a cone. We impose $(x, y) \in [-x_{min}, x_{max}] \times [-y_{min}, y_{max}]$. Lagrange parameters are $\lambda = -\frac{\nabla f \nabla h'}{\nabla h \nabla h'} = \frac{e^z}{1+x^2+y^2}$. The projection matrix is defined as

$$P(x) = \frac{1}{1+x^2+y^2} \begin{pmatrix} 1+y^2 & -xy & x \\ -xy & 1+x^2 & y \\ x & y & x^2+y^2 \end{pmatrix}$$

We calculate

$$\frac{d^2L}{d\mathbf{x}^2} = \begin{pmatrix} \frac{e^z}{1+x^2+y^2} & 0 & 0 \\ 0 & \frac{e^z}{1+x^2+y^2} & 0 \\ 0 & 0 & e^z \end{pmatrix}$$

Finally, the eigenvalues of $F = P(x)\frac{d^2L}{dx^2}$ are

$$e^{\frac{x^2+y^2}{2}} \left(\frac{1}{1+x^2+y^2}, \frac{1+x^2+x^4+y^2+2x^2y^2+y^4}{1+x^2+y^2} \right)$$

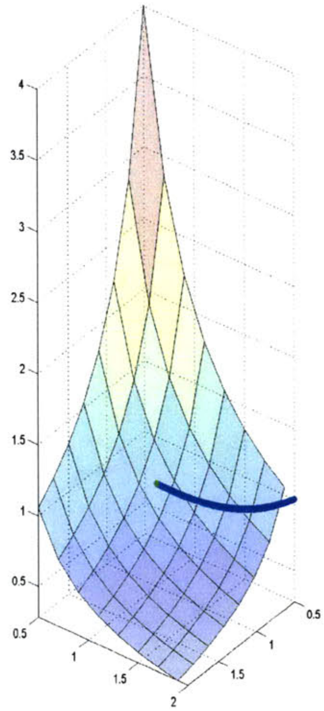

Figure 2-9: Minimal sum of lengths for a given volume

In this case $\lambda_{max} = \frac{1+x_{min}^2+x_{min}^4+y_{min}^2+2x_{min}^2y_{min}^2+y_{min}^4}{1+x_{max}^2+y_{max}^2} > 0$. This constrained dynamical system is contracting. The solution is attained in $(0,0,0)$. On figure 2-19 we observe four different trajectories with four different initial conditions. As it is an autonomous system, they all converge towards a unique equilibrium point.

Example of minimization with two constraints

Consider the minimization problem but with two constraints

$$min\ \frac{1}{2}((x-x_1)^2+(y-y_1)^2+(z-z_1)^2)\ \ s.t.\ h_1(x) = x^2+y^2+z^2-1 = 0\ \ h_2(x) = x+y+z = 0$$

The absolute minimum is again $1,4,2$. The starting point must be a point where both constraints are verified : $[0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}]$. The dynamical system follows both constraints. Each Lagrange parameter acts as a force making the dynamical system to stay on the corresponding constraint. The final point is the double projection of the absolute

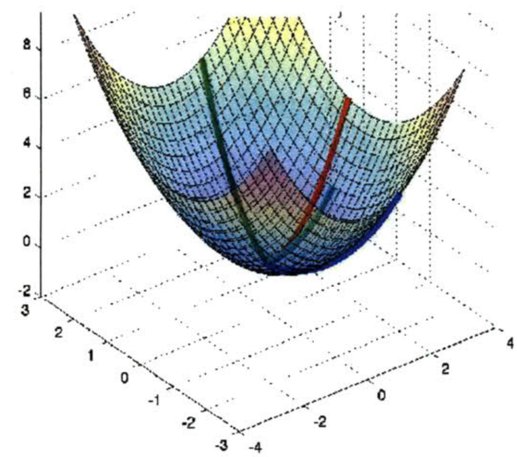

Figure 2-10: Contraction behavior

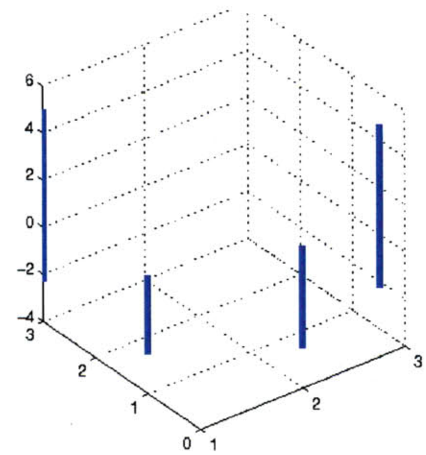

Figure 2-11: No contraction behavior

minimum, figure $(3-10)$. This method in this case is only 100 times faster than the matlab method.

Example of minimization with scaling

Consider the minimization problem

$$min \sum_{i=1}^{i=N} x_i \ \ s.t. \ h(x) = \sum x_i^2 - 1 = 0$$

For N=300, our method is 100 times faster than matlab. For N=3000, our method is 1000 times faster. This method is very robust to scaling.

Example of minimization on a Torus

Consider the minimization problem

$$min \ x \ \ s.t. \ h(x,y,z) = (R - \sqrt{x^2+y^2})^2 + z^2 - r^2 = 0$$

The minimum on the torus is on one of the sides. On figure 2-13, the system evolves on the constraint even if the shape of the surface is very complicated. R corresponds to the outer radius and r corresponds to the inner radius.

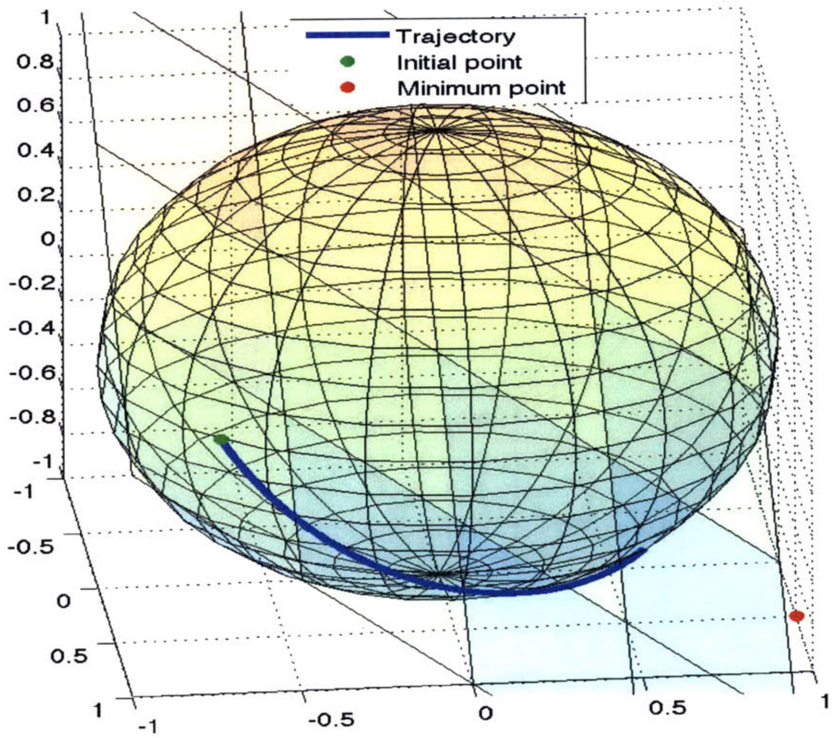

Figure 2-12: Example of minimization with two constraints

2.2 Starting outside the constraints: sliding behavior

2.2.1 The problem and intuition of the solution

The most important limitation of this problem is that we have to start on the constraints in order to achieve contraction behavior. The solution is to add the following dynamic.

$$-\sum_i c_i h_i(\mathbf{x}) \nabla h_i(\mathbf{x})$$

c_i is a constant that gives the speed of approach to the constraint i. $\nabla h_i(\mathbf{x})$ gives the direction that is normal to the surface locally and $-h_i(\mathbf{x})$ gives a sense to this direction, (figure 2-14). A priori, this property is local because far away from the surface the term in ∇h_i is not necessarily normal to the surface. Consider the modified

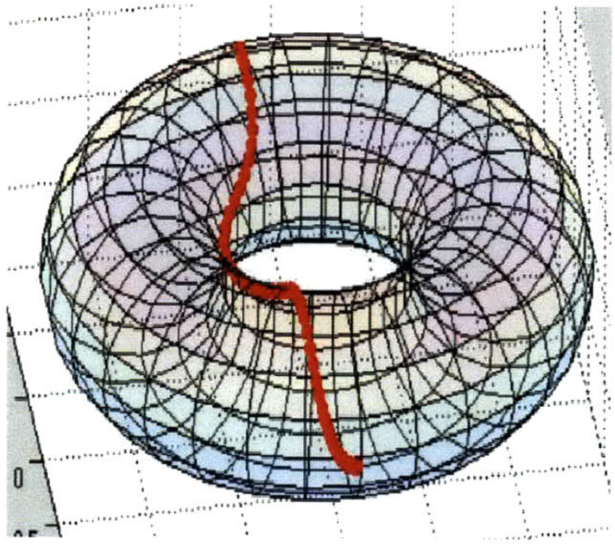

Figure 2-13: Minimization on a torus

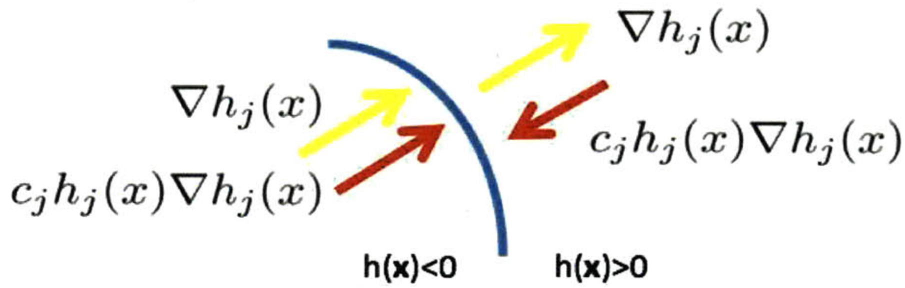

Figure 2-14: Sliding dynamic

dynamic system:

$$\dot{\mathbf{x}} = f(\mathbf{x},t)P(\mathbf{x}) - \sum_j c_j h_j(\mathbf{x})\nabla h_j(\mathbf{x}) = f(\mathbf{x},t) + \lambda \nabla h(\mathbf{x}) - \sum_j c_j h_j(\mathbf{x})\nabla h_j(\mathbf{x})$$

Integrating the precedent expression for an autonomous gradient system, we obtain

$$\dot{\mathbf{x}} = -\nabla(U(\mathbf{x}) + \lambda h(\mathbf{x}) + \frac{1}{2}\sum_j c_j h_j^2(\mathbf{x})) = -\nabla L_c$$

This term corresponds to the term that is added in optimization to construct the augmented lagrangian, L_c. This dynamic term can be understood as adding a cost when the constraints are not verified.

2.2.2 Sliding surface behavior

In this section we prove that converging to the surface when starting away from is a global property. We use sliding control techniques. We define $s = \frac{1}{2}(\sum_i c_i h_i^2(\mathbf{x}))$ where $c_i > 0$. s is by definition non negative. Its derivative is

$$\frac{ds}{dt} = \sum_i c_i h_i(\mathbf{x})\nabla h_i(\mathbf{x})\dot{\mathbf{x}} = \sum_i c_i h_i(\mathbf{x})\nabla h_i(\mathbf{x})(f(\mathbf{x},t)P(\mathbf{x}) - \sum_j c_j h_j(\mathbf{x})\nabla h_j(\mathbf{x}))$$

Because $f(\mathbf{x},t)P(\mathbf{x}) \in M(\mathbf{x})$ then $\nabla h_i(\mathbf{x})f(\mathbf{x},t)P(\mathbf{x}) = 0$. The derivative is then equal to :

$$\frac{ds}{dt} = -\sum_i c_i h_i(\mathbf{x})\nabla h_i(\mathbf{x})(\sum_j c_j h_j(\mathbf{x})\nabla h_j(\mathbf{x}))$$

$$= -\left(\sum_i c_i^2 h_i^2(\mathbf{x})\nabla h_i^2(\mathbf{x}) + 2\sum_{i<j} c_j c_i h_j(\mathbf{x})\nabla h_j(\mathbf{x})h_i(\mathbf{x})\nabla h_i(\mathbf{x})\right) = -(\sum c_i h_i \nabla h_i)^2 \leq 0$$

The second derivative can be calculated. If $\ddot{s}$ is bounded, we apply barbalat's lemma [12]. $\sum c_i h_i \nabla h_i$ converges to zero as time goes to infinity. Using the initial hypothesis, ∇h_i are linearly independent, then $c_i h_i$ goes to zero. The constraints are then verified. The technique used here is a very common technique in nonlinear control theory. h^2 can be seen as the distance to the surface, which is the same as the sliding variable s used in nonlinear control.

With one constraint we get exponential convergence towards the surface.

$$\frac{d}{dt}\frac{h^2}{2} = h(\mathbf{x})\nabla h(\mathbf{x})\dot{\mathbf{x}} = -ch^2(\mathbf{x})(\nabla h(\mathbf{x}))^2$$

There are two dynamics in this augmented system. These two dynamics can be controlled using the c parameter. The first dynamic is the so called, Lagrangian dynamic. This dynamic leads to the minimum once on the surface. The second dynamic is the so called sliding dynamic. It makes the system converge to the constraints. If the value of c is very big, we give more importance to the sliding term. The system will converge quicker to the surface. If c is small, it will take longer to converge towards the surface.

2.2.3 Examples

Minimization of the length of a square in a circle

Consider the optimization problem:

$$min\ x + y\ s.t.\ x^2 + y^2 = 1$$

The projected system is then :

$$\dot{x} = \begin{pmatrix} -1 + \frac{x^2+xy}{x^2+y^2} - c(x^2 + y^2 - 1)x \\ -1 + \frac{y^2+xy}{x^2+y^2} - c(x^2 + y^2 - 1)y \end{pmatrix}$$

If c is small, (c=0.5, figure 2−15) initially the system follows the Lagrangian dynamic, (it follows a circle that is situated at the starting position). As the system goes away from the surface, the sliding dynamic becomes more important. This algorithm is

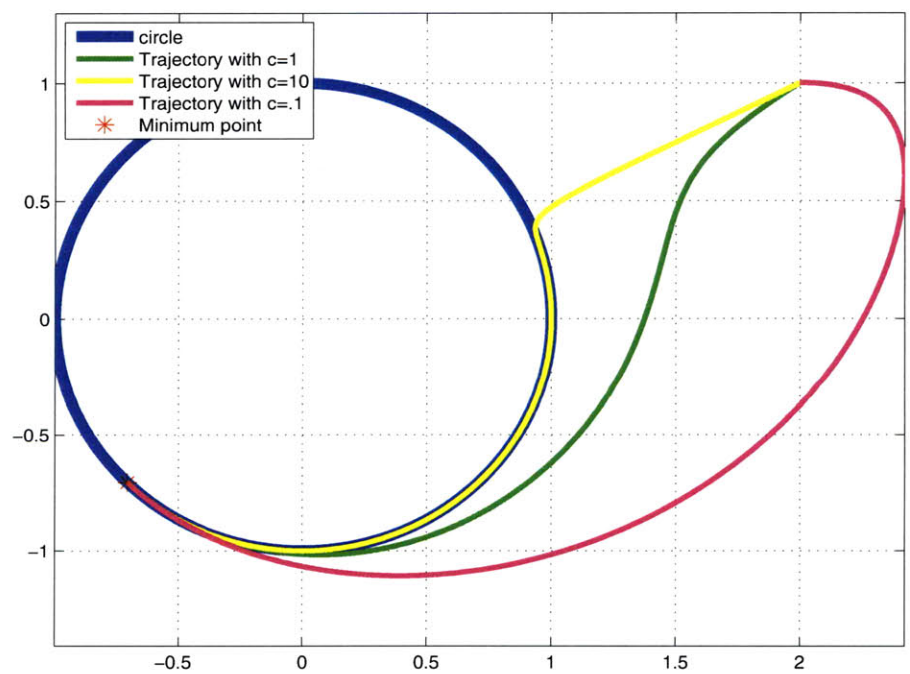

Figure 2-15: Minimization of the length of a square in a circle

1300 times faster than matlab. If c is big, (c=10, figure $2 - 15$) the system converges very quickly to the surface. Once on the surface the Lagrangian dynamic makes the system to evolve on the constraint towards the minimum. This algorithm is 134 times faster than matlab method. This method is slower because we need a smaller stepsize to make the system converge. The best trade off is to choose a middle value (c=1). We approach the surface following the shape of the constraint. The number $\frac{1}{c}$ can be seen as the radius at which the lagrangian dynamics start working.

Minimum distance from a point to a sphere

Consider the minimization problem

$$min\ \frac{1}{2}(x^2 + y^2 + (z-2)^2) \quad s.t. \ \ h(x) = x^2 + y^2 + z^2 - 1 = 0$$

The solution is $(0, 0, 1)$. The projected system adding the sliding term is:

$$\dot{x} = \nabla U + \lambda \nabla h = \begin{pmatrix} x + \lambda x + c(x^2 + y^2 + z^2 - 1)x \\ y + \lambda y + c(x^2 + y^2 + z^2 - 1)y \\ z - 2 + \lambda z + c(x^2 + y^2 + z^2 - 1)z \end{pmatrix}$$

with $\lambda = \frac{2z-1}{x^2+y^2+z^2}$. As usual the computation time, starting at the same initial point is 1000 times faster than matlab.

Maximum volume on an ellipse

The problem is to maximize the volume of a parallelipede subject to lying on an ellipse

$$min - xyz \quad s.t. \ \ h(x) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$$

In this case we consider $a = 15; b = 5; c = 3;$. On figure $2-17$ we achieve the solution. As usual the convergence time is very fast compared to matlab.

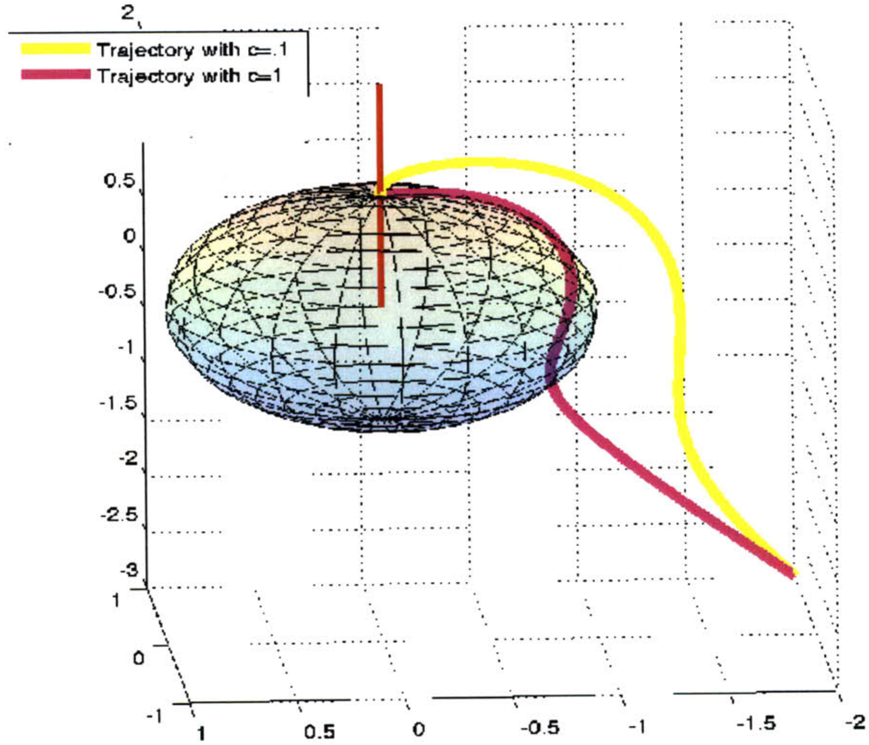

Figure 2-16: Minimum distance from a point to a sphere

Minimal sum of lengths for a given volume

The problem in this case is to maximize the volume of a parallelipede subject to lying on an ellipse

$$min\ x + y + z \quad s.t.\ \ h(x) = xyz - 1 = 0$$

The augmented Lagrangian system is:

$$\dot{x} = \begin{pmatrix} 1 + \lambda yz + c(xyz - 1)yz \\ 1 + \lambda xz + c(xyz - 1)xz \\ 1 + \lambda xy + c(xyz - 1)xy \end{pmatrix}$$

where $\lambda = -\frac{yz+xz+xy}{(yz)^2+(xz)^2+(xy)^2}$. On figure $2-18$ we achieve the solution, [1 1 1]. As usual the convergence time is very fast compared to matlab.

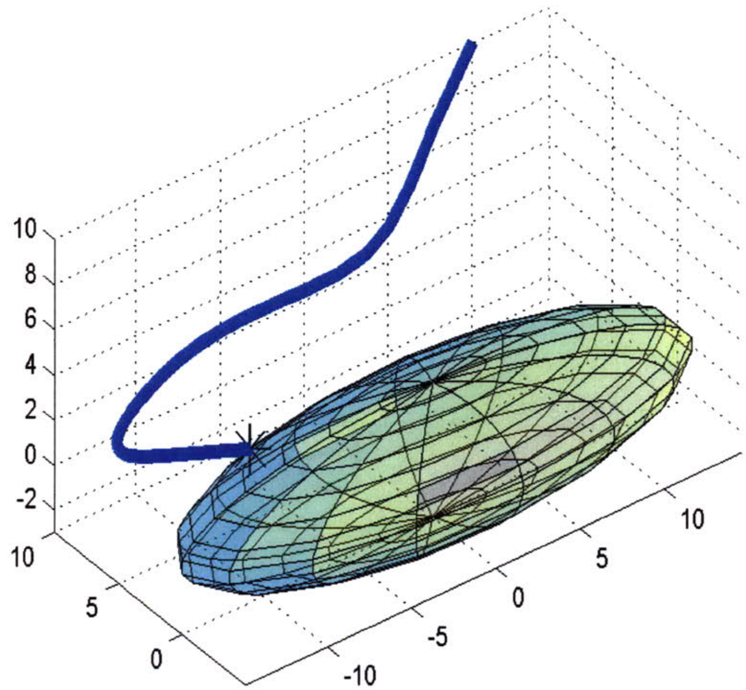

Figure 2-17: Maximum volume on an ellipse

Contraction of the initial and projected system

We consider

$$\dot{x} = [0, 0, -e^z] \; s.t. \; h(x, y, z) = \frac{x^2 + y^2}{2} - z = 0$$

Figure 2-19 shows for $c = 1$ the sliding of the system and then reaches the minimum.

Example of minimization with two constraints

Consider the minimization problem but with two constraints

$$min \, \frac{1}{2}((x-x_1)^2+(y-y_1)^2+(z-z_1)^2) \;\; s.t. \; h_1(x) = x^2+y^2+z^2-1 = 0 \; h_2(x) = x+y+z = 0$$

The final point is the double projection of the absolute minimum. The values c_1, c_2 are the gains to the two surfaces. To stay on the sphere, we increase the gain ($c_1 = 10$).

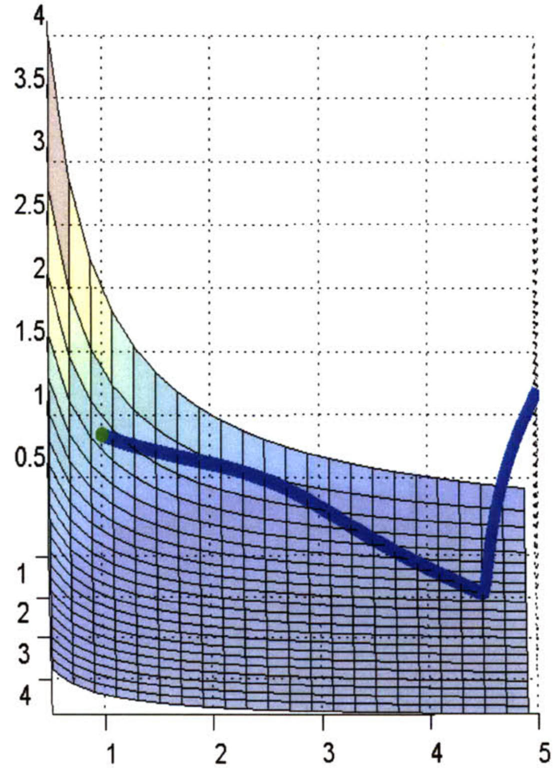

Figure 2-18: Minimal sum of lengths for a given volume

This method in this case is only 100 times faster than the matlab method.

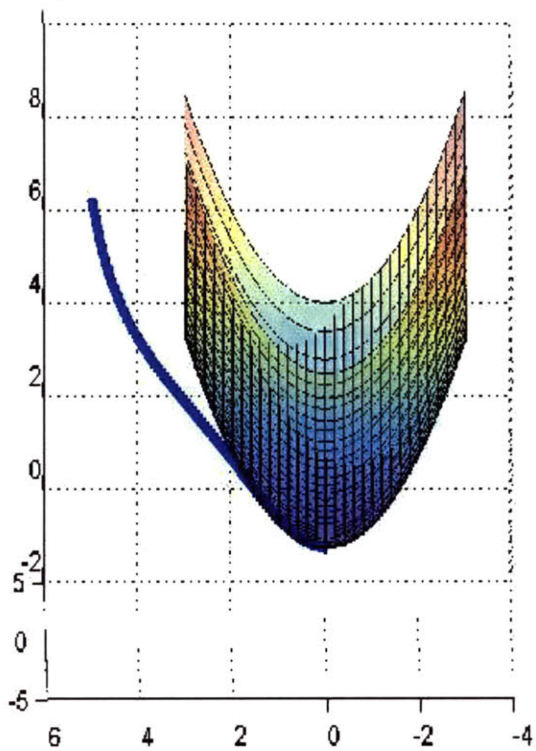

Figure 2-19: Contraction of the initial and projected system

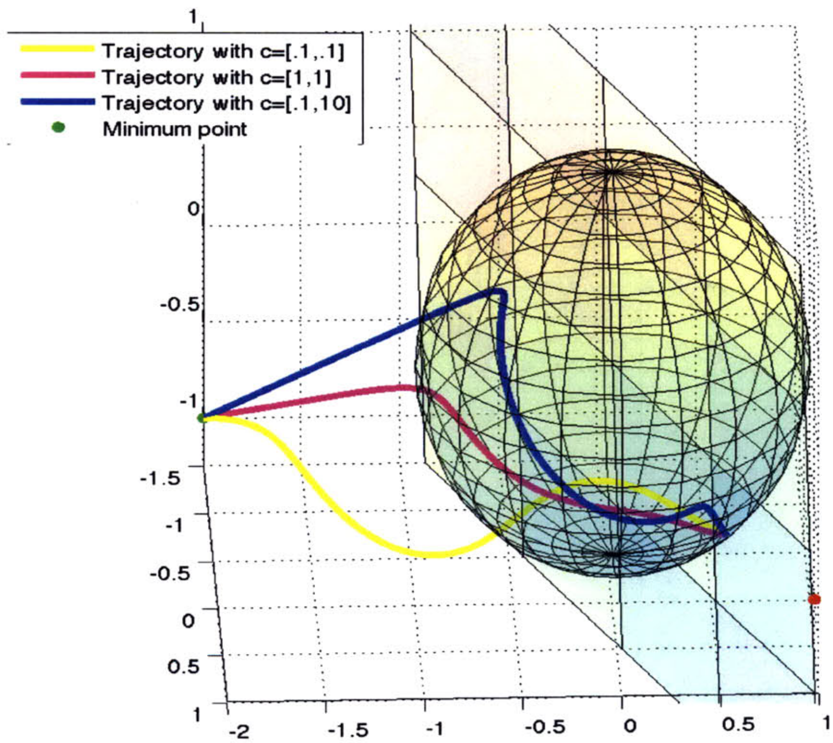

Figure 2-20: Example of minimization with two constraints

Chapter 3

Contraction theory with inequality constraints

There are two parts in this chapter. In the first one we state the theorem that gives a condition of contraction for dynamic systems constrained to inequality constraints. The second part uses a sliding condition to make the system converge to the constraints.

3.1 Starting on the constraints: contraction theory

We define the set of interior points and the boundary, $I = [\mathbf{x} \ / \ h(\mathbf{x}) \leq 0]$. We define the half saturation function as

$$h_{sat}(x) = \left\{ \begin{array}{l} 0, x \leq 0 \\ x, x \in [0,1) \\ 1, x \geq 1 \end{array} \right\} \tag{3.1}$$

and also $P_{h_{sat}}(\mathbf{x}) = 1 - \nabla\mathbf{h}'(\mathbf{x})[\nabla\mathbf{h}(\mathbf{x})\nabla\mathbf{h}(\mathbf{x})']^{-1}\nabla\mathbf{h}(\mathbf{x})\mathbf{h}_{sat}(\frac{\mathbf{h}}{\mathbf{\Phi}})$ where $\mathbf{\Phi} = [\Phi_1, ..., \Phi_m]$ are strictly positive time varying functions. These functions are boundary layers used in [12]. $P_{h_{sat}}$ is not a projector because when $x \in [0,1]$, $P_{h_{sat}} P_{h_{sat}} \neq P_{h_{sat}}$. For $x \leq 0$

then $P_{h_{sat}} = I$. For $x \geq 1$, then $P_{h_{sat}} = P$, where P is the projector defined in chapter two.

3.1.1 Theorem

In this section we compute the condition of contraction for dynamic systems that are under inequality constraints.

Theorem 6. *The condition of contraction of the system*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \ s.t. \ \mathbf{h}(\mathbf{x}) \preceq 0$$

is that the hermitian part of :

$$\mathbf{F} = \left(\dot{\mathbf{\Theta}}(\mathbf{x}, t) + \mathbf{\Theta}(\mathbf{x}, t)\left(P_{h_{sat}}\frac{\partial \mathbf{f}}{\partial \mathbf{x}} + P\lambda(\mathbf{x})h_{sat}\frac{d^2\mathbf{h}(\mathbf{x})}{d\mathbf{x}^2}\right)\right)\mathbf{\Theta}(\mathbf{x}, t)^{-1} \ on \ I(\mathbf{x})$$

has a maximum eigenvalue uniformly negative. $\lambda(\mathbf{x}) = -\mathbf{f}(\mathbf{x}, t)\nabla\mathbf{h}'[\nabla\mathbf{h}\nabla\mathbf{h}']^{-1}$. *The initial condition must verify the constraint.*

3.1.2 Proof

This proof has two parts.

Projection of the dynamic system

The initial problem is

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \ s.t. \ \mathbf{h}(\mathbf{x}) \preceq 0$$

Physically, three cases can be distinguished for each constraint h_i.
First of all, if the current point verifies $h_i(\mathbf{x}(t)) < 0$, then the system can evolve freely in any direction. The system remains the same $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$
If the current point verifies the constraint, $h_i(\mathbf{x}(t)) = 0$, and has with respect to the active constraint an inward speed, $\dot{\mathbf{x}}\nabla h_i' = \mathbf{f}(\mathbf{x}, t)\nabla h_i' < 0$, the system will verify the constraint. The system is $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$.

If the current point verifies the constraint, $h_i(\mathbf{x}(t)) = 0$, and has with respect to the active constraint an outward speed, $\dot{\mathbf{x}}\nabla h_i' = \mathbf{f}(\mathbf{x}, t)\nabla h_i' > 0$,the system will escape from the constraint. In this case we use the projection operator to avoid that. The system is $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \lambda_i(\mathbf{x}, t)\nabla h_i(\mathbf{x})$.
There is a switching between systems. This is a problematic approach. To address this problem we use $P_{h_{sat}}$. Using this function there is no more switching, the system is continuous.

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) + \lambda(\mathbf{x}, t)h_{sat}(\frac{\mathbf{h}}{\mathbf{\Phi}})\nabla h(\mathbf{x}) = \mathbf{f}(\mathbf{x}, t) + \sum_{i=0}^{i=m} \lambda_i h_{sat}(\frac{h_i}{\Phi_i})\nabla h_i$$

The idea when introducing this boundary layer is to relax the constraint $h_i = \mathbf{0}$ to $-\Phi_i \leq h_i \leq \Phi_i$. We associate being on the constraint with an inward speed to $-\Phi_i \leq h_i \leq 0$ and being on the constraint with an outward speed of the dynamic system to $0 \leq h_i \leq \Phi_i$.

Condition of contraction

As done in the precedent theorem, we define $\delta\mathbf{z} = \mathbf{\Theta}(\mathbf{x}, t)\delta\mathbf{x}$. All calculations are similar except $\delta\dot{\mathbf{x}}$. We compute it,

$$\delta\dot{\mathbf{x}} = \left(\frac{\partial f}{\partial \mathbf{x}} + \lambda(\mathbf{x}, t)\frac{d^2h}{d\mathbf{x}^2}h_{sat} + \frac{\partial\lambda(\mathbf{x}, t)}{\partial \mathbf{x}}\frac{dh}{d\mathbf{x}}h_{sat} + \lambda(\mathbf{x}, t)\frac{dh}{d\mathbf{x}}\frac{dh_{sat}}{d\mathbf{x}}\right)\delta\mathbf{x}$$

We compute $\lambda(\mathbf{x}, t)\frac{dh}{d\mathbf{x}}\frac{dh_{sat}}{d\mathbf{x}} = \sum_i \frac{\lambda_i}{\Phi}\nabla h_i \nabla h_i' \leq 0$ because when $\mathbf{x} \in [0, \Phi]$ then $\lambda_i \leq 0$
Using equation 2.1, we compute $\frac{\partial\lambda(\mathbf{x},t)}{\partial\mathbf{x}}$

$$\delta\dot{\mathbf{x}} = \left(\frac{\partial f}{\partial \mathbf{x}}P_{h_{sat}}(\mathbf{x}) + P(\mathbf{x})h_{sat}\lambda(\mathbf{x}, t)\frac{d^2h}{d\mathbf{x}^2}\right)\delta\mathbf{x}$$

Finally

$$\frac{d}{dt}(\delta\mathbf{z}\delta\mathbf{z}) = 2\delta\mathbf{z}\left(\dot{\Theta}(\mathbf{x}, t) + \Theta(\mathbf{x}, t)\left(P_{h_{sat}}(\mathbf{x})\frac{\partial f}{\partial \mathbf{x}} + P(\mathbf{x})h_{sat}\lambda(\mathbf{x}, t)\frac{d^2h}{d\mathbf{x}^2}\right)\right)\Theta(\mathbf{x}, t)^{-1}\delta\mathbf{z}$$

Shrinking behavior can be showed using what is done in chapter 2.

3.1.3 Theorem's application

This theorem has a main application to minimization theory. But has also an application to adaptive control.

Contraction analysis and adaptive control

It is possible to use the theorem of contraction for inequalities in order to conclude convergence of the adaptive control. The equations of adaptive control are given in [12]: $\dot{s} = -ks + Y\tilde{a}$ and $\dot{\hat{a}} = -Y's$ The virtual system in $[y_1, y_2]$: $\dot{y}_1 = -ky_1 + Yy_2$ $\dot{y}_2 = -Y'y_1$ The virtual displacement of this system:

$$\begin{pmatrix} \delta\dot{y}_1 \\ \delta\dot{y}_2 \end{pmatrix} = \begin{pmatrix} -k & Y \\ -Y' & 0 \end{pmatrix} \begin{pmatrix} \delta y_1 \\ \delta y_2 \end{pmatrix}$$

The symmetric part of the jacobian is $\begin{pmatrix} -k & 0 \\ 0 & 0 \end{pmatrix}$ We can conclude contraction for an identity metric for y_1 but not for y_2. Intuitively adding constraints gives some more knowledge about this second parameter. We consider now the constraints $h(y_2) = 0$. We project the second equation

$$\dot{y}_2 = -Y'y_1 + \lambda\nabla h(y_2)$$

where $\lambda = -Y'y_1\nabla h'[\nabla h\nabla h']^{-1}$ We calculate the associated virtual system:

$$\delta\dot{y}_2 = -Y'\delta y_1 - \lambda\frac{d^2h(y_2)}{dy_2{}^2}P(y_2)\delta y_2$$

The symmetric part of the jacobian is

$$\begin{pmatrix} -k & 0 \\ 0 & -\lambda\frac{d^2h(y_2)}{dy_2{}^2}P(y_2) \end{pmatrix}$$

The projection matrix is positive definite. If the constraint is convex, then $\frac{d^2h(y_2)}{dy_2{}^2} > 0$. Then we can conclude contraction of y_1 and y_2 if and only if the Lagrange parameters are strictly positive. Similar results with convex constraints have already been achieved by [17]. This result is more general because it allows many constraints.

Optimization and projected contraction theory

As done in [1] and in the precedent chapter, an autonomous system has a unique equilibrium point.

$$\dot{\mathbf{x}} = -\nabla U(\mathbf{x}) \ s.t. \ h(\mathbf{x}) \preceq 0$$

This system has a unique equilibrium point. Three cases can be distinguished for each constraint h_i

First if $h_i(\mathbf{x}^*) \leq -\Phi_i$, then it verifies $\nabla U(\mathbf{x}^*) = 0$. The condition of contraction is $\nabla^2 U(\mathbf{x}^*) > 0$.

If $-\Phi_i \leq h_i \leq 0$, then $\nabla U(\mathbf{x}^*) = 0$. The condition of contraction is $\nabla^2 U(\mathbf{x}^*) > 0$.We also have $\lambda_i < 0$.

If $0 \leq h_i \leq \Phi_i$, then $\nabla L(\mathbf{x}^*) = 0$. The condition of contraction is $\mathbf{x}' P \nabla^2 L(\mathbf{x}^*) P \mathbf{x} > 0$. We also have $\lambda_i < 0$.

The dynamic system also satisfies the constraint $h_i(\mathbf{x}^*) \leq 0$ because we start on the constraint and the speed of the system is always towards the interior of the constraints or tangential to them. These two conditions ensure us that at anytime $h_i(\mathbf{x}) \leq 0$.

In each case, these are the conditions for the existence of a minimum for the problem 1.4 needed by the Karush-Kuhn-Tucker theorem in order to have a minimum solution. Therefore, there is a new relationship between minimization and dynamical systems using contraction theory.

Similarity with precedent results

In [9] the dynamic system has to evolve in a hypercube. This corresponds to linear inequalities. The definition of the dynamical projected system is very similar to the one that we have done. The projected dynamical system is defined as

$$\dot{\mathbf{x}} = \Pi_{\Omega}(\mathbf{x}, \mathbf{f}(\mathbf{x}, t))$$

where Π_{Ω} is the point projection operator on Ω. Intuitively when inside the constraints, this operator does not do anything. When on the constraint with an outward speed, we apply the projection operator.

3.1.4 An algorithm for solving minimization problems

This projection operator gives us a new dynamic system that can be implemented.

Minimization of the length of a square inside a circle

Consider :

$$min \;\; x + y \; s.t. \; x^2 + y^2 = 1$$

As it can be simulated, if we make the boundary layer too small then the system converges very quickly to constraint.

Kepler's planimetric problem

The problem is how to inscribe in a given circle a rectangle of maximal area. The minimization formulation is :

$$min \;\; xy \; s.t \; x^2 + y^2 = 1$$

As it can be simulated, if we make the boundary layer too small then the system converges very quickly to constraint.

Minimum distance from a point inside a sphere

Consider the minimization problem

$$min \frac{1}{2}(x^2 + (y - .2)^2 + (z - .4)^2) \quad s.t. \;\; h(x) = x^2 + y^2 + z^2 - 1 = 0$$

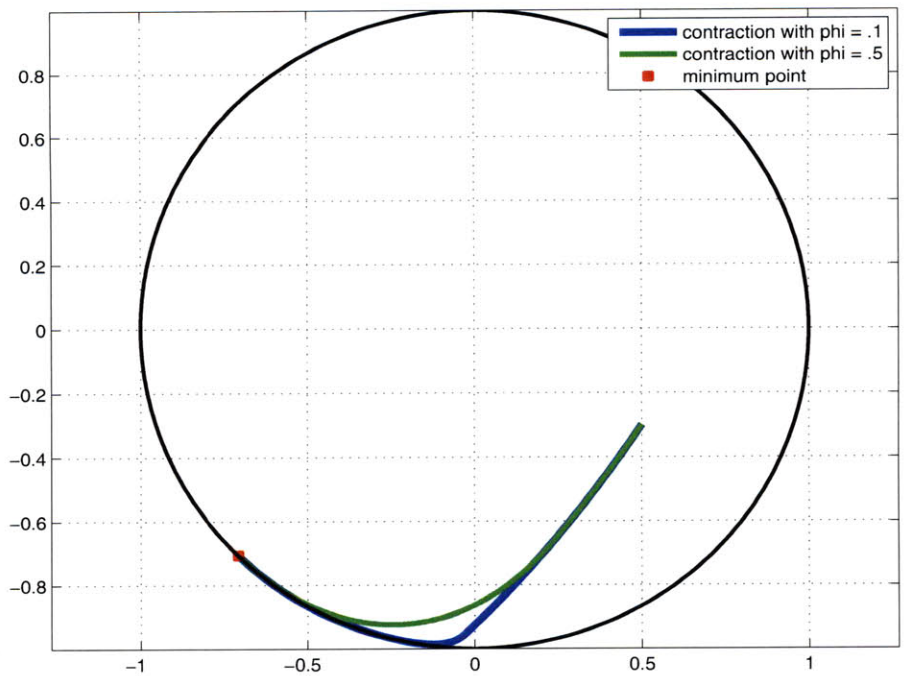

Figure 3-1: Minimization of the length of a square inside a circle with two different boundary layers

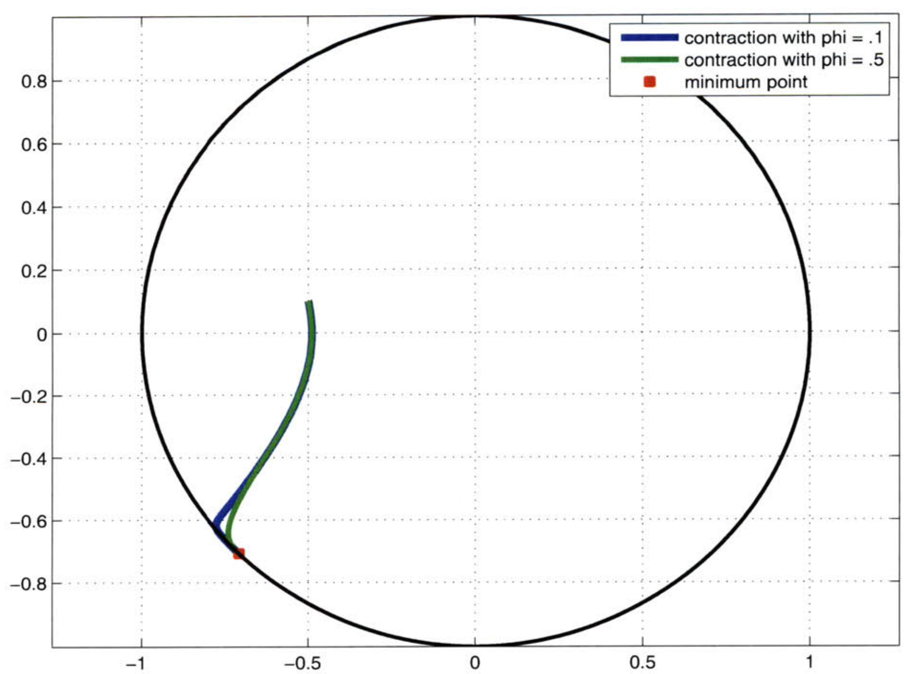

Figure 3-2: Solution to Kepler planimetric problem with two different boundary layers

The minimum is not on the boundary. The system stills achieve the solution because inside the constraint we have a gradient dynamic.

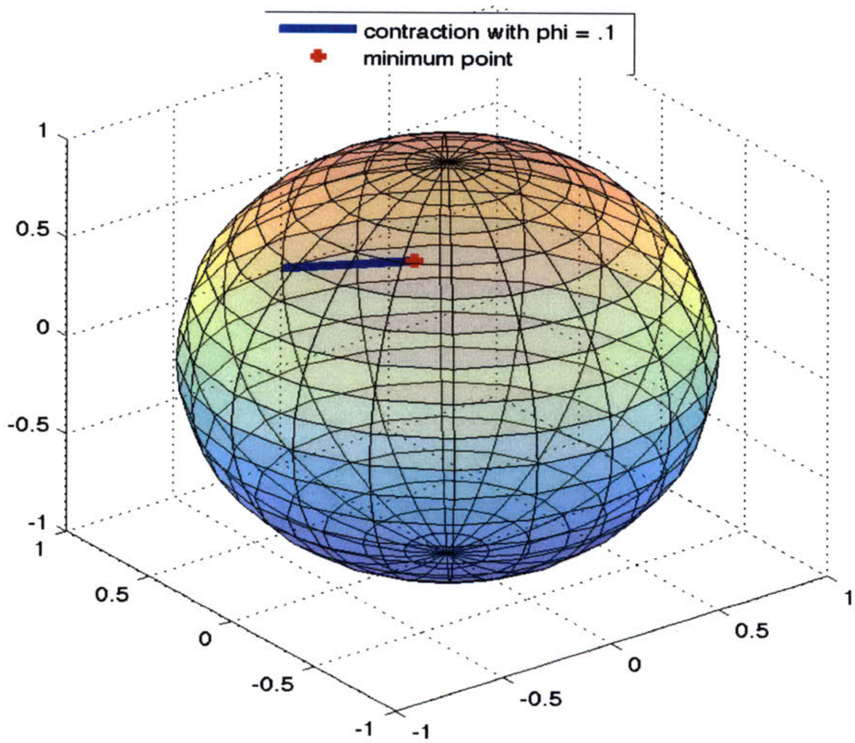

Figure 3-3: Minimum distance from a point inside a sphere

Maximum volume on an ellipse

The problem is to maximize the volume of a parallelipede subject to lying on an ellipse

$$min \ -xyz \quad s.t. \ h(x) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$$

Example of minimization with two constraints

Consider the minimization problem but with two constraints

$$min \ \frac{1}{2}((x-x_1)^2+(y-y_1)^2+(z-z_1)^2) \ \ s.t. \ h_1(x) = x^2+y^2+z^2-1 = 0 \ h_2(x) = x+y+z = 0$$

As it can be simulated, if we make the boundary layer too small then the system converges very quickly to constraint.

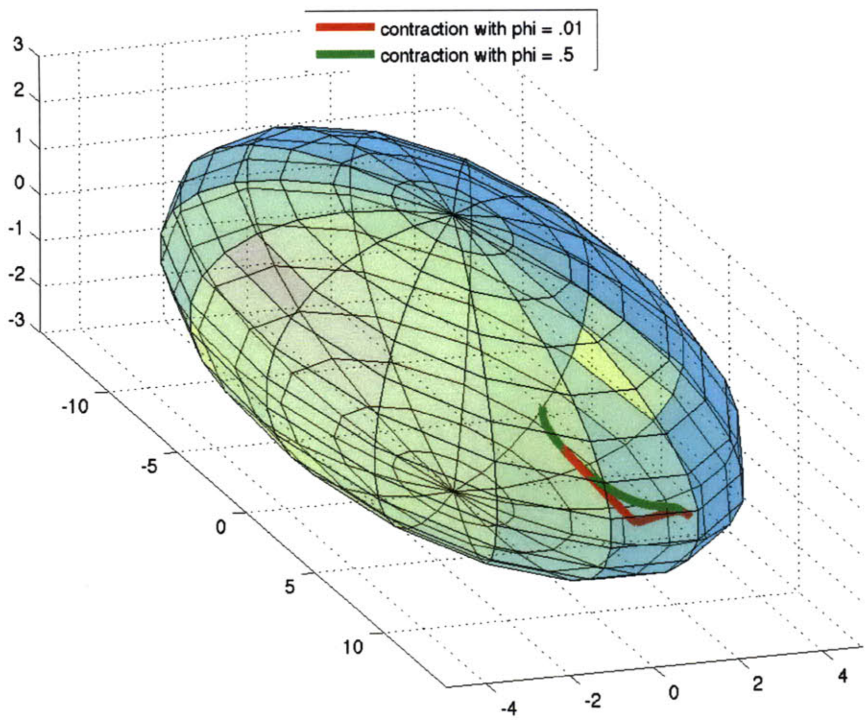

Figure 3-4: Maximum volume on an ellipse with two different boundary layers

3.2 Starting outside the constraints: one-side sliding behavior

3.2.1 One-side sliding condition

We need to introduce a modified concept of sliding. In all precedent applications, the convergence from both sides to the sliding surface was very important. With inequalities, $h_i \leq 0$ is a feasible region, we only need to slide along the side corresponding to $h_i > 0$. We introduced the idea of boundary layer earlier. If the system starts outside the boundary layer for the constraint i , $h_i \geq \Phi_i$, we want to add a dynamic that will make the system go inside the boundary layer. $\dot{h}_i < \dot{\Phi}_i$ As $h_i \geq \Phi_i$ we multiply by h_i the precedent inequality. The one side sliding condition is

$$\mathbf{h} \succeq \Phi \;\Rightarrow \frac{1}{2}\frac{d}{dt}\left(\sum_{i=0}^{i=m} h_i^2\right) \leq \sum_{i=0}^{i=m}(\dot{\Phi}_i - \eta_i)|h_i(t)| \tag{3.2}$$

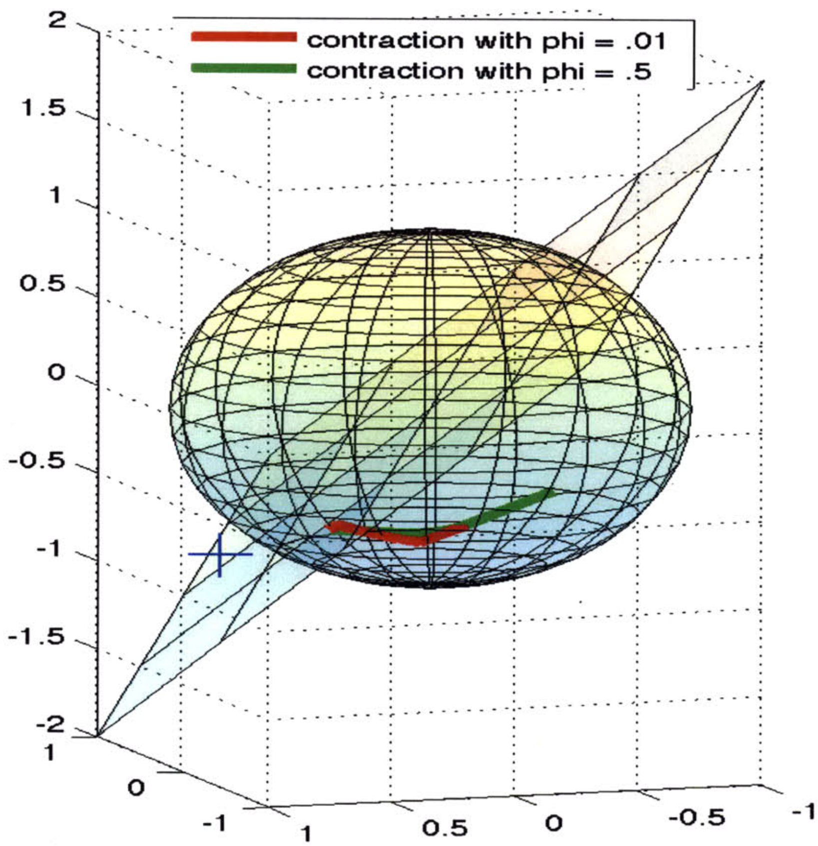

Figure 3-5: Minimization with two constraints with two different boundary layers

where m is the number of constraints. This condition is very similar to the one developed in [12].

3.2.2 One side sliding surface behavior

In order to have this one slide sliding behavior we need to add a dynamic. The term is the following

$$-\sum_i c_i \nabla h_i(\mathbf{x}) h_{sat}\left(\frac{h_i}{\Phi}\right)$$

The modified dynamic system is:

$$\dot{\mathbf{x}} = f(\mathbf{x},t)P_{h_{sat}}(\mathbf{x}) - \sum_j c_j \nabla h_j(\mathbf{x}) h_{sat} = f(\mathbf{x},t) + \left(\lambda \nabla h - \sum_j c_j h_j \nabla h_j(\mathbf{x})\right) h_{sat}$$

This system verifies the condition of one side sliding. Its derivative:

$$\frac{1}{2}\frac{d}{dt}\left(\sum_i h_i^2\right) = \sum_i h_i(\mathbf{x})\nabla h_i(\mathbf{x})\dot{\mathbf{x}} = \sum_i c_i h_i(\mathbf{x})\nabla h_i(\mathbf{x})\left(f(\mathbf{x},t)P_{h_{sat}}(\mathbf{x}) - \sum_j c_j \nabla h_j(\mathbf{x}) h_{sat}\right)$$

Because for $\mathbf{h} \geq \mathbf{\Phi}$, we have $h_{sat} = 1$ then $f(\mathbf{x},t)P_{h_{sat}}(\mathbf{x}) = f(\mathbf{x},t)P(\mathbf{x}) \in M(\mathbf{x})$ then $\nabla h_i(\mathbf{x}) f(\mathbf{x},t)P(\mathbf{x}) = 0$.

$$\frac{1}{2}\frac{d}{dt}\left(\sum_i h_i^2\right) = -\mathbf{c}\nabla\mathbf{h}\nabla\mathbf{h}'\mathbf{h}'$$

where $\mathbf{c} = [c_1, ..., c_m] = [\eta_1 - \Phi_1, ..., \eta_m - \Phi_m] = \eta - \mathbf{\Phi}$. As $\nabla h' \nabla h$ is of full rank then $x'\nabla h'\nabla h x \geq 0$. We add a very mild hypothesis, $x'\nabla h'\nabla h x \geq \gamma$ we achieve one side sliding contraction behavior

$$\frac{1}{2}\frac{d}{dt}\left(\sum_i h_i^2\right) \leq -\gamma \sum c_i h_i(t) = \gamma \sum (\dot{\Phi}_i - \eta_i) h_i(t)$$

3.2.3 Examples

We apply the sliding technique to the precedent examples.

Minimization of the length of a square inside a circle

Consider :

$$min\ x + y\ s.t.\ x^2 + y^2 = 1$$

Convergence to the minimum is achieved even starting outside the constraints.

Kepler's planimetric problem

The problem is how to inscribe in a given circle a rectangle of maximal area. The minimization formulation is :

$$min\ xy\ s.t\ x^2 + y^2 = 1$$

Convergence to the minimum is achieved even starting outside the constraints.

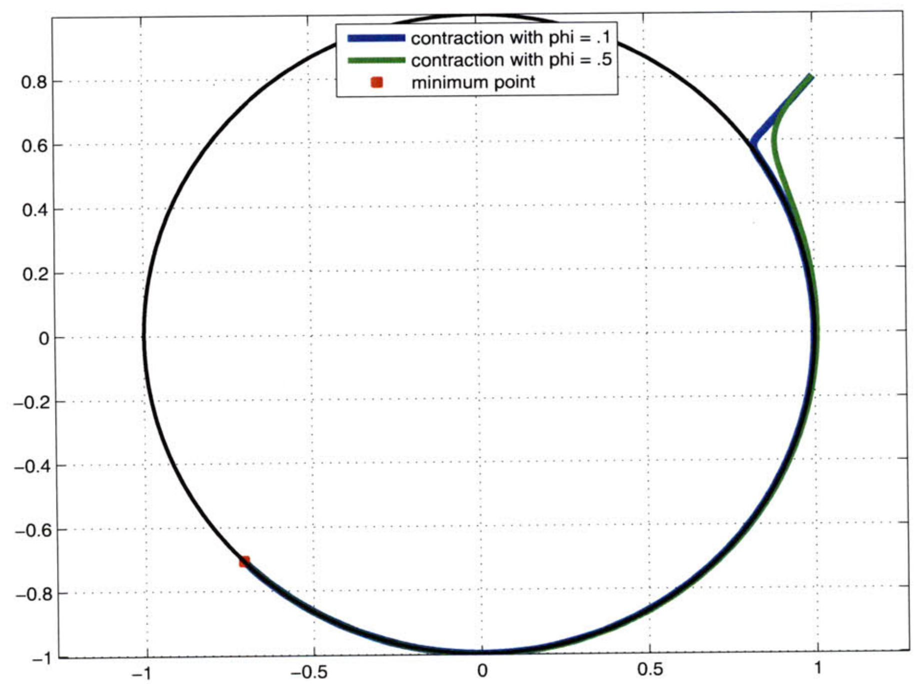

Figure 3-6: Minimization of the length of a square inside a circle with two different boundary layers

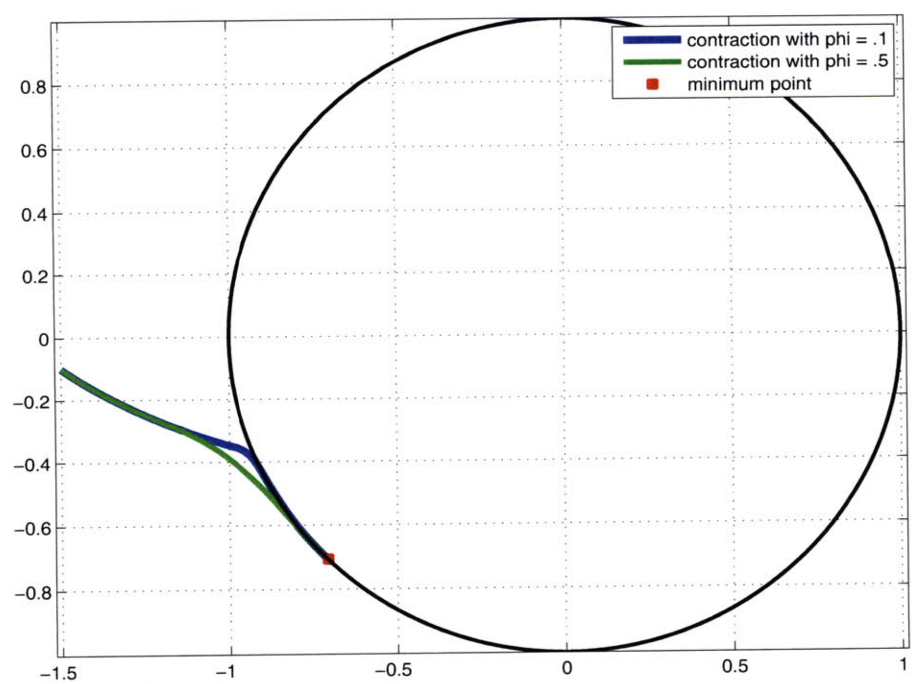

Figure 3-7: Solution to Kepler planimetric problem with two different boundary layers

Minimum distance from a point inside a sphere

Consider the minimization problem

$$min\ \frac{1}{2}(x^2 + (y - .2)^2 + (z - .4)^2) \quad s.t.\ \ h(x) = x^2 + y^2 + z^2 - 1 = 0$$

Convergence to the minimum is achieved even starting outside the constraints.

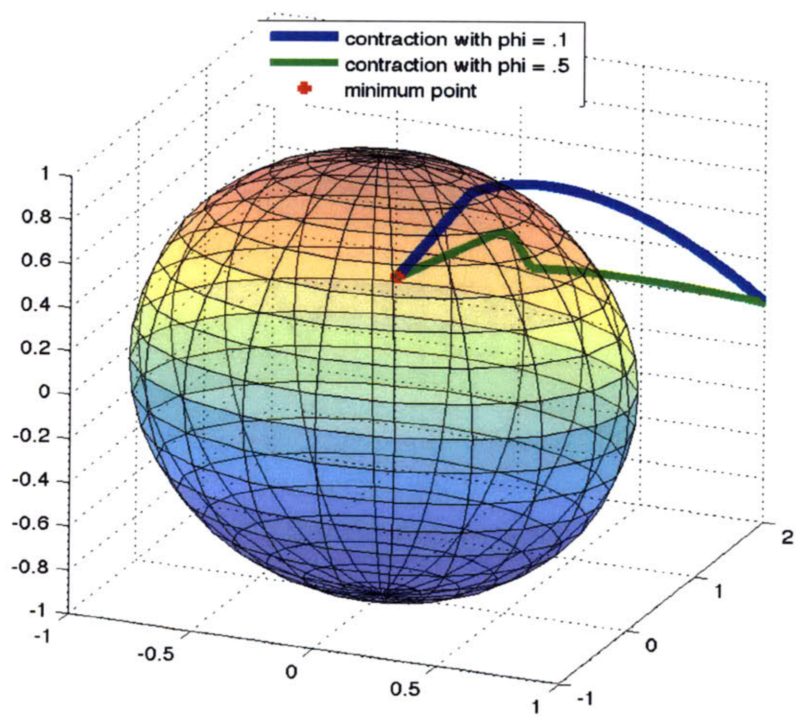

Figure 3-8: Minimum distance from a point inside a sphere

Maximum volume on an ellipse

The problem in this case is to maximize the volume of a parallelipede subject to lying on an ellipse

$$min\ -xyz \quad s.t.\ \ h(x) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0$$

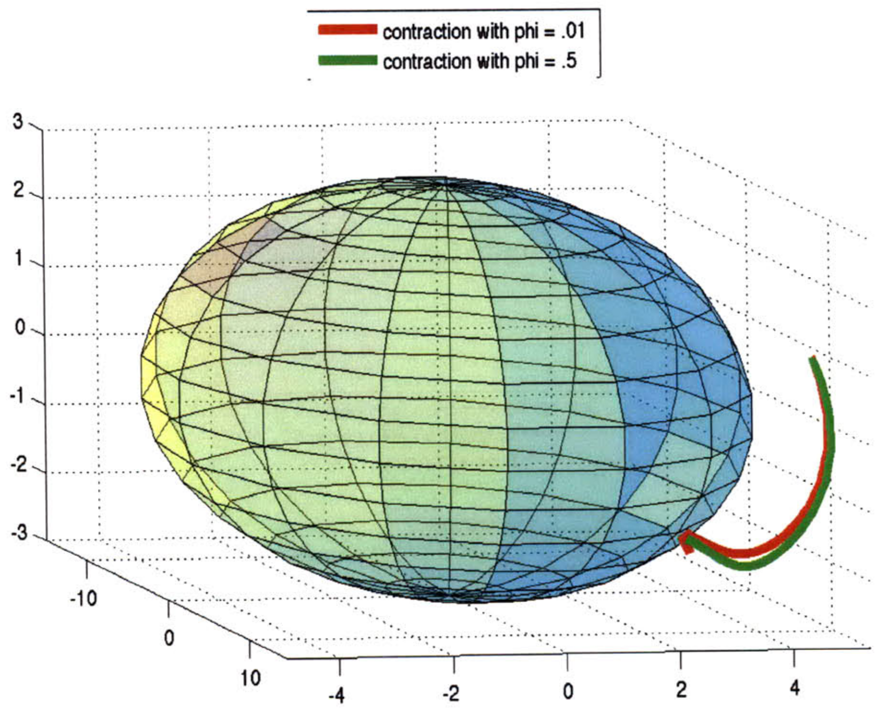

Figure 3-9: Maximum volume on an ellipse with two different boundary layers

Example of minimization with two constraints

Consider the minimization problem but with two constraints

$$min\ \frac{1}{2}((x-x_1)^2+(y-y_1)^2+(z-z_1)^2)\ \ s.t.\ h_1(x) = x^2+y^2+z^2-1 = 0\ h_2(x) = x+y+z = 0$$

Convergence to the minimum is achieved even starting outside the constraints.

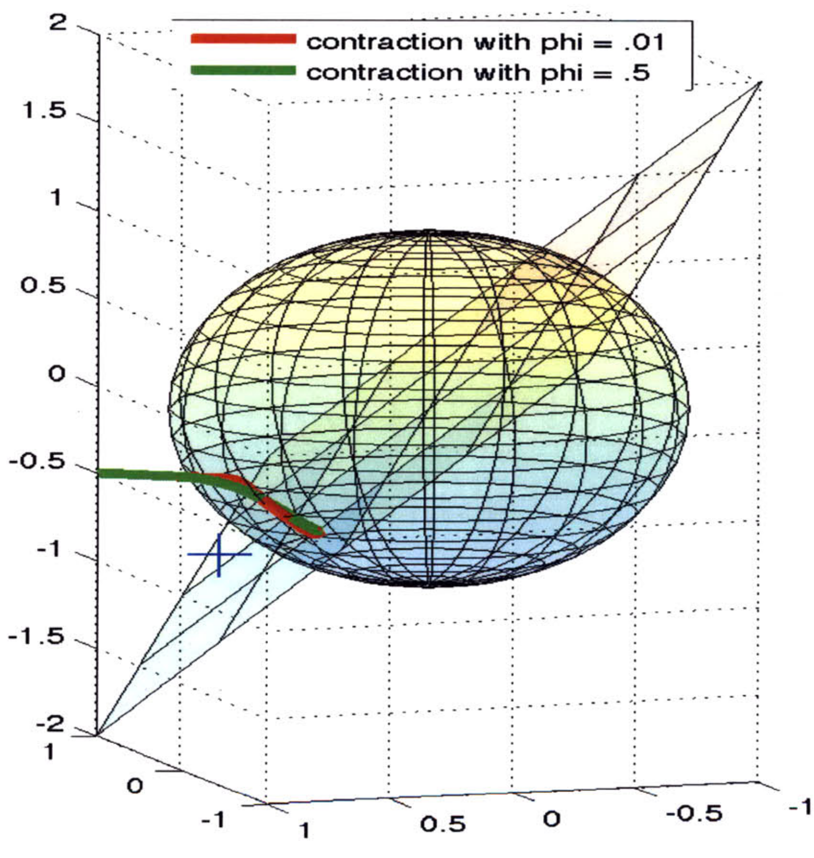

Figure 3-10: Minimization with two constraints with two different boundary layers

Chapter 4

Contraction theory with time-varying equality constraints

This section has two main parts. In the first part we state a theorem giving the condition of contraction for a time varying constrained nonlinear dynamic system. We have to start on the constraints to have contraction behavior. In the second part, we allow the system to start outside the constraints. We have to use sliding techniques to conclude convergence to the constraints.

4.1 Starting on the constraints: contraction theory

We define the set $S(\mathbf{x}, t) = [\mathbf{x} \ / \ \mathbf{h}(\mathbf{x}, t) = 0]$.

4.1.1 Theorem

In this section we compute the condition of contraction for dynamic systems that are under time-varying constraints.

Theorem 7. *The condition of contraction of the system*

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \ s.t. \ h(\mathbf{x}, t) = 0$$

is that the hermitian part of :

$$\mathbf{F} = \left(\dot{\mathbf{\Theta}}(\mathbf{x},t) + \mathbf{\Theta}(\mathbf{x},t)P(\mathbf{x})\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}} + \lambda(\mathbf{x},t)\frac{\partial^2 \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}^2}\right)\right)\mathbf{\Theta}(\mathbf{x},t)^{-1} \quad on\ S(\mathbf{x},t)$$

has a maximum eigenvalue uniformly negative. $P(\mathbf{x})$ *is a projection operator onto the tangent space* $S(\mathbf{x},t)$ *and* $\lambda(\mathbf{x},t) = -(\mathbf{f}(\mathbf{x},t)\nabla h + \frac{\partial h}{\partial t})[\nabla h \nabla h']^{-1}$. *The initial condition must verify the constraint.*

4.1.2 Proof

Projection of the dynamic system

The initial problem is

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},t)\ s.t.\ h(\mathbf{x},\mathbf{t}) = 0$$

Physically to remain on the constraints, we need to project the speed and we need to take into account the speed of the surface, figure 4-1. To calculate this last term, we

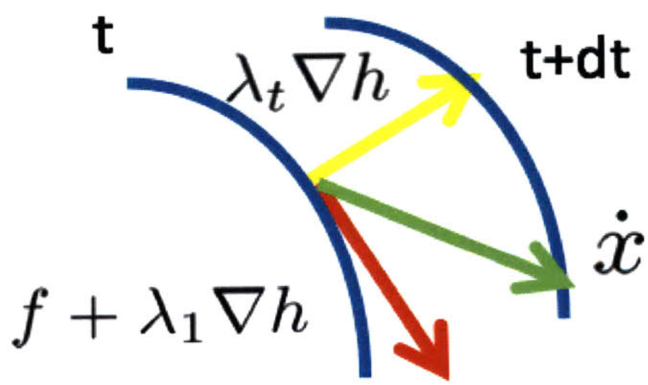

Figure 4-1: Projection with time-varying constraints

use the fact that the constraints are always verified, $0 = \frac{dh}{dt} = \nabla h \dot{x}' + \frac{\partial h}{\partial t}$ As we have the choice on $\dot{x}'$, we add a term to ensure that this equation is verified. This term is :

$$-\frac{\partial h}{\partial t}[\nabla h \nabla h']^{-1}\nabla h = \lambda_2 \nabla h$$

We can see this new parameter as a force that will make the system stay on the constraint. It is a new generalization of Lagrange parameters. Hence the total system is:

$$\dot{x} = \mathbf{f}(\mathbf{x},t) + \lambda_1 \nabla h - \frac{\partial h}{\partial t}[\nabla h \nabla h']^{-1}\nabla h = \mathbf{f}(\mathbf{x},t) + \lambda_1 \nabla h + \lambda_2 \nabla h = f(x,t) + \lambda \nabla h$$

where $\lambda = \lambda_1 + \lambda_2$. The constraints are verified at anytime because the system starts on the constraint and the velocity is such that it remains on the constraints.

Condition of contraction

We compute $\delta\dot{\mathbf{x}} = \left(\frac{\partial f}{\partial \mathbf{x}} + \lambda(\mathbf{x})\frac{\partial^2 h}{\partial \mathbf{x}^2} + \frac{\partial \lambda_1(\mathbf{x})}{\partial \mathbf{x}}\frac{\partial h}{\partial \mathbf{x}} + \frac{\partial \lambda_2(\mathbf{x})}{\partial \mathbf{x}}\frac{\partial h}{\partial \mathbf{x}}\right)\delta\mathbf{x}$. $\frac{\partial \lambda_1(\mathbf{x})}{\partial dx}$ can be computed.

$$\frac{\partial \lambda_1(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial f}{\partial \mathbf{x}} + \lambda_1 \frac{\partial^2 h}{\partial \mathbf{x}^2}\right) - \nabla h'[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1} - \left(\frac{\partial f}{\partial \mathbf{x}} + \lambda_1 \frac{\partial^2 h}{\partial \mathbf{x}^2}\right)\frac{\partial^2 h}{\partial \mathbf{x}^2}[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1}$$

We have to compute also $\frac{\partial \lambda_2(\mathbf{x})}{\partial \mathbf{x}}$.

$$\frac{\partial \lambda_2(\mathbf{x})}{\partial \mathbf{x}} = \left(\frac{\partial h}{\partial t} - \lambda_2 \frac{\partial^2 h}{\partial \mathbf{x}^2}\right)\nabla h'[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1} - \lambda_2 \frac{\partial^2 h}{\partial \mathbf{x}^2}[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1}$$

We can substitute in the precedent equation and reorganize:

$$\delta\dot{\mathbf{x}} = P(\mathbf{x})\left(\frac{\partial f}{\partial \mathbf{x}} + \lambda(\mathbf{x})\frac{\partial^2 h}{\partial \mathbf{x}^2}\right)\delta\mathbf{x} - \nabla h(\mathbf{x})'[\nabla h(\mathbf{x})\nabla h(\mathbf{x})']^{-1}\left(\delta\mathbf{x}\frac{\partial^2 h}{\partial \mathbf{x}^2}\dot{\mathbf{x}} + \frac{\partial h}{\partial t}\delta\mathbf{x}\right)$$

It is essential to note that the first term belongs to $M(\mathbf{x})$ and the second term belongs to $M(\mathbf{x})^{\perp}$. To have contraction behavior the first term of $\delta\dot{\mathbf{x}}$ must be uniformly bounded.

$$\delta\dot{\mathbf{x}} = P(\mathbf{x})\left(\frac{\partial f}{\partial \mathbf{x}} + \lambda(\mathbf{x})\frac{\partial^2 h}{\partial \mathbf{x}^2}\right)\delta\mathbf{x}$$

Finally

$$\frac{d}{dt}(\delta\mathbf{z}\delta\mathbf{z}) = 2\delta\mathbf{z}\left(\dot{\Theta}(\mathbf{x},t) + \Theta(\mathbf{x},t)P(\mathbf{x})\left(\frac{\partial f}{\partial \mathbf{x}} + \lambda(\mathbf{x})\frac{\partial^2 h}{\partial \mathbf{x}^2}\right)\right)\Theta(\mathbf{x},t)^{-1}\delta\mathbf{z}$$

4.1.3 Theorem's applications

One of the most important application is optimization under varying constraints.

Minimization of the sum of the sides under a growing circle

The problem is :

$$min\ x + y\ s.t.\ x^2 + y^2 = t^2$$

The solution using KKT theorem is given by $x^* = y^* = -\frac{t}{\sqrt{2}}$. Using the precedent theorem we can create the 'extended' lagrangian system.

$$\dot{x} = -\begin{pmatrix} 1 - \frac{(x+y)}{t^2}x - \frac{x}{t} \\ 1 - \frac{(x+y)}{t^2}y - \frac{y}{t} \end{pmatrix}$$

The condition of contraction is given by $\nabla^2 f + (\lambda_1 + \lambda_2)\nabla^2 h = -\frac{x+y+t}{t^2}I > 0$. Then the final condition is $x + y + t < 0$. On figure $4-2$, we achieve the solution after some transient time

Minimization of the sum of the sides under a changing ellipse

The problem is :

$$min\ x + y\ s.t.\ tx^2 + y^2 = 1$$

The solution using KKT theorem is given by $x^* = -\frac{1}{\sqrt{1+t^2}}$, $y^* = -\frac{t}{\sqrt{1+t^2}}$.

$$\dot{x} = -\begin{pmatrix} 1 - \frac{(tx+y)}{t^2x^2+y^2}tx - \frac{x^2}{2t^2x^2+2y^2}xt \\ 1 - \frac{(tx+y)}{t^2x^2+y^2}y - \frac{x^2}{2t^2x^2+2y^2}y \end{pmatrix}$$

The condition of contraction is given by $\nabla^2 f + (\lambda_1 + \lambda_2)\nabla^2 h = -\frac{x^2+tx+y}{t^2x^2+y^2}I > 0$. Then the final condition is $x^2 + tx + y < 0$. On figure $4-3$ we achieve the solution after some transient time

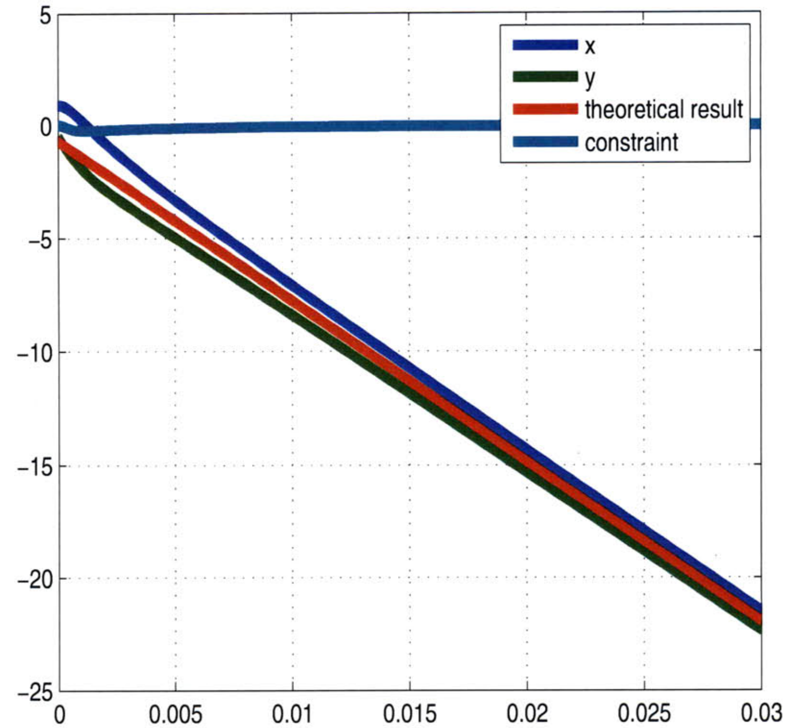

Figure 4-2: Minimization of the sum of the sides under a growing circle

Using time varying constraints to solve problems with static constraints

In order to make a smoother approach to a certain constraint h_0, we can create a time varying constraint such that when $t \to \infty$ then $h(x,t) \to h_0(x)$. Consider the minimum system:

$$min\ x + y \quad s.t.\ x^2 + y^2 = e^{\frac{2}{t}}$$

The dynamical system associated to that is :

$$\dot{x} = - \begin{pmatrix} 1 - \frac{(x+y)}{e^{\frac{2}{t}}} x - \frac{x}{t^2} \\ 1 - \frac{(x+y)}{e^{\frac{2}{t}}} y - \frac{y}{t^2} \end{pmatrix}$$

On figure $4-4$, we observe that we achieve the solution after some transient time This is the solution achieved by the minimization of the sum of the lengths subject to a circle.

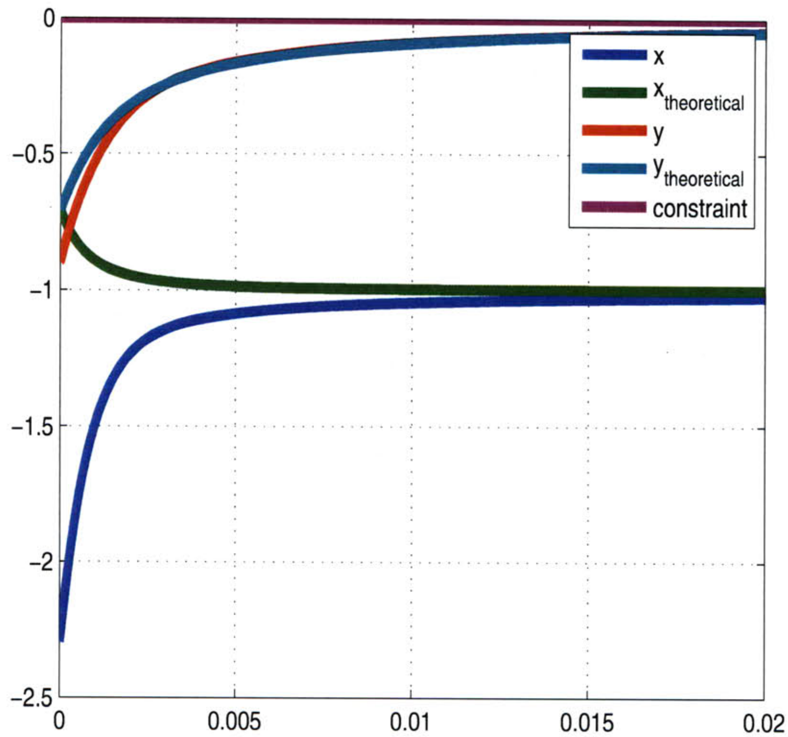

Figure 4-3: Minimization of the sum of the sides under a changing ellipse

4.2 Starting outside the constraints: sliding behavior

4.2.1 Sliding surface behavior

As done in the precedent chapter, we add a sliding term. The new system is :

$$\dot{x} = \mathbf{f}(x,t) + \lambda \nabla \mathbf{h} - \frac{\partial \mathbf{h}}{\partial t}[\nabla \mathbf{h} \nabla \mathbf{h}']^{-1} \nabla \mathbf{h} - \sum_j c_j h_j(x) \nabla h_j(x)$$

We have global convergence towards the surface.We define $s = \frac{1}{2}(\sum_i c_i h_i^2(x))$ where $c_i > 0$. s is by definition strictly positive. Its derivative:

$$\frac{ds}{dt} = C'\mathbf{h}' \left(\frac{\partial \mathbf{h}}{\partial x} \dot{x}' + \frac{\partial \mathbf{h}}{\partial t} \right)$$

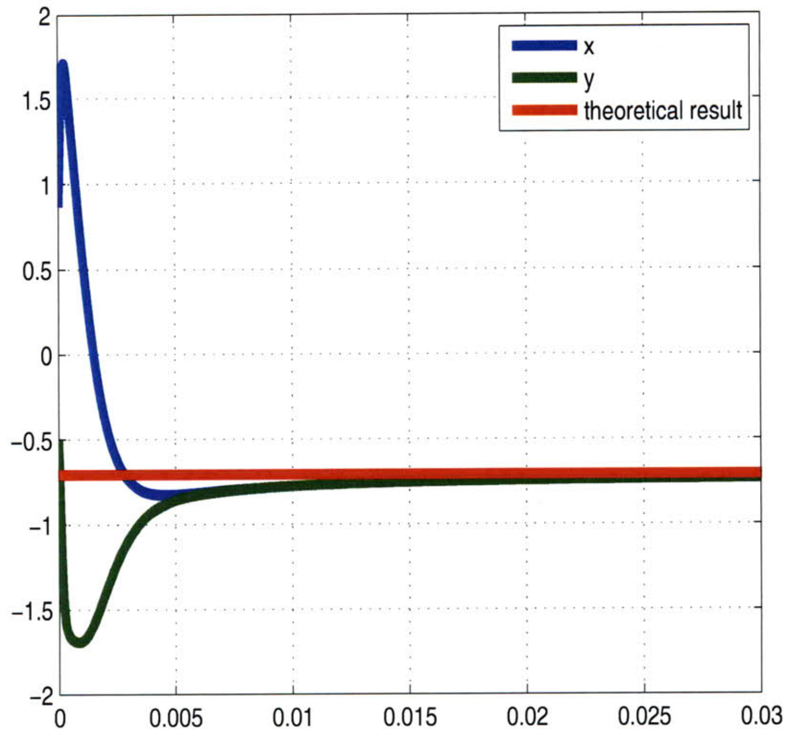

Figure 4-4: Using time varying constraints to solve problems with static constraints

$$= C'\mathbf{h}' \left(\frac{\partial \mathbf{h}}{\partial x} \left(P(x)'\mathbf{f}(x,t)' - \nabla \mathbf{h}' [\nabla \mathbf{h} \nabla \mathbf{h}']^{-1} \frac{\partial \mathbf{h}}{\partial t} - \sum_j c_j h_j(x) \nabla h_j(x) \right) + \frac{\partial \mathbf{h}}{\partial t} \right)$$

Because $\mathbf{f}(x,t)P(x) \in M(x)$ then $\nabla h_i(x)\mathbf{f}(x,t)P(x) = 0$. As $\nabla \mathbf{h} \nabla \mathbf{h}' [\nabla \mathbf{h} \nabla \mathbf{h}']^{-1} \frac{\partial \mathbf{h}}{\partial t} = \frac{\partial \mathbf{h}}{\partial t}$. The derivative is then equal to :

$$\frac{ds}{dt} = -(\sum c_i h_i \nabla h_i)^2 \leq 0$$

The second derivative is smooth, then it is bounded. Applying barbalat's lemma, we conclude that $\sum c_i h_i \nabla h_i$ converges to zero as time goes to infinity. As ∇h_i are linearly independent then $c_i h_i$ goes to zero. The constraints are then verified.

4.2.2 Examples

To show how this sliding term works, we use the precedent examples but starting on a point outside the surface.

Minimization of the sum of the sides under a growing circle

The problem is :

$$min\ x + y\ s.t.\ x^2 + y^2 = t^2$$

The new system is

$$\dot{x} = -\begin{pmatrix} 1 - \frac{(x+y)}{x^2+y^2}x - \frac{x}{t} + chx \\ 1 - \frac{(x+y)}{x^2+y^2}y - \frac{y}{t} + chy \end{pmatrix}$$

On figure $4-5$ we achieve the solution after some transient time even starting outside of the dynamic constraint.

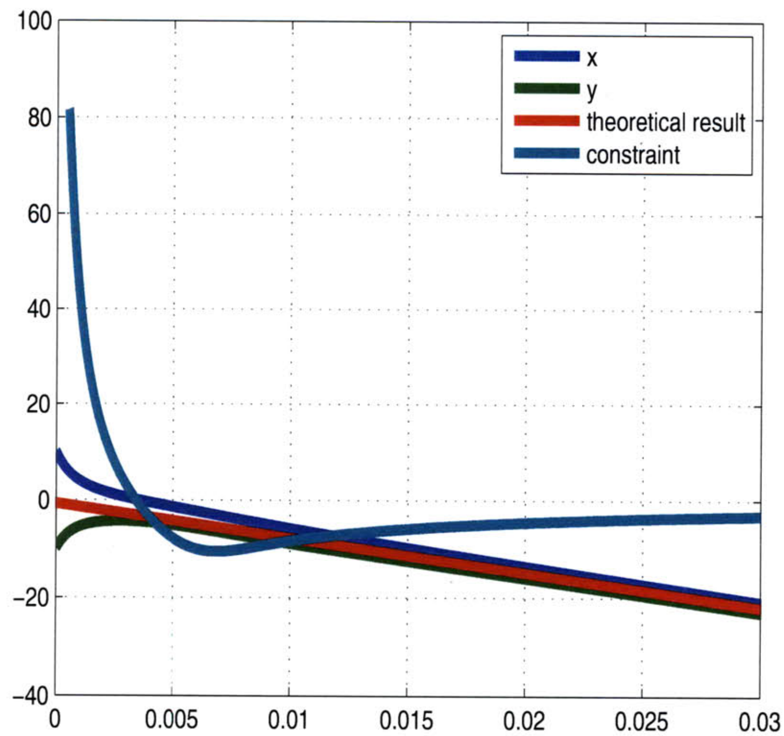

Figure 4-5: Minimization of the sum of the sides under a growing circle

Minimization of the sum of the sides under a changing ellipse

The problem is :

$$min\ x + y\ s.t.\ tx^2 + y^2 = 1$$

The new system is

$$\dot{x} = -\begin{pmatrix} 1 - \frac{(tx+y)}{t^2x^2+y^2}tx - \frac{x^2}{2t^2x^2+2y^2}xt + chxt \\ 1 - \frac{(tx+y)}{t^2x^2+y^2}y - \frac{x^2}{2t^2x^2+2y^2}y + chy \end{pmatrix}$$

On figure $4-6$ we achieve the solution after some transient time even starting outside of the dynamic constraint.

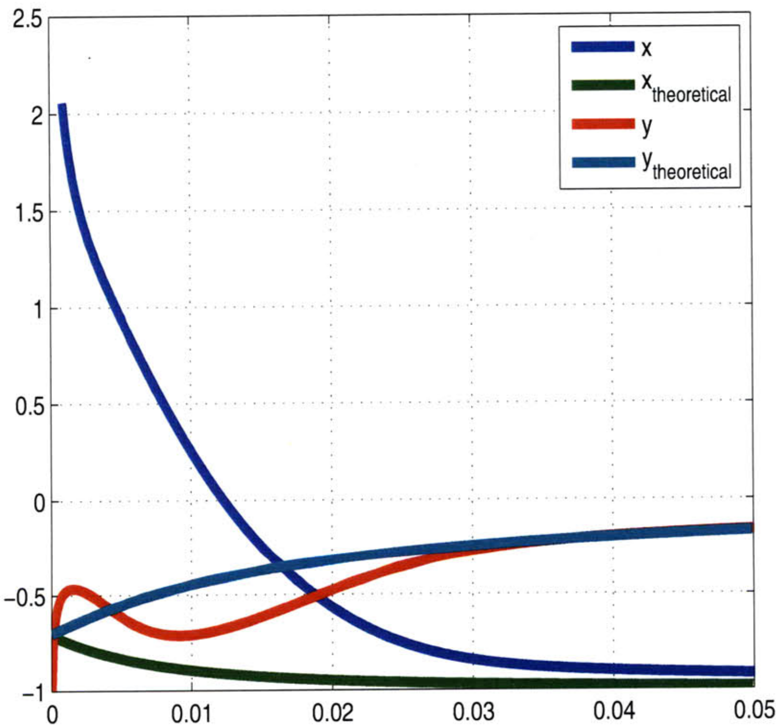

Figure 4-6: Minimization of the sum of the sides under a changing ellipse

Chapter 5

Conclusion and directions for future research

5.1 Conclusion

This thesis has four main new contributions.

The first and more important is the condition of contraction (a condition of convergence) for three different dynamic systems, constrained with equalities, constrained with inequalities and constrained with time varying inequalities.

The second contribution is, in the particular case of a gradient autonomous contracting in a metric identity system, a fast algorithm to find the minimum. The contraction conditions are only partially understood. In this very particular case, we achieve such a quick algorithm to find the minimum. May be a more general metric will allow to solve more complicated problems.

The third contribution is the understanding that adding a term $\frac{h^2}{2}$ in the cost function corresponds to dynamic term that makes the dynamic system converge towards the surface

The fourth and last contribution is on Lagrange parameters. This thesis gives a new physical approach of Lagrange parameters. They can be understood as a scalar value that ensure the system to have a tangential speed to the constraints. Also it gives three generalizations of Lagrange parameters, with a time varying cost function $\lambda = -\nabla U(\mathbf{x},t)\nabla h'[\nabla h \nabla h']^{-1}$, with a time varying general vector $\lambda = -\mathbf{f}(\mathbf{x},t)\nabla h'[\nabla h \nabla h']^{-1}$ and with a a time varying general vector with time varying constraints $\lambda = -\left(\mathbf{f}(\mathbf{x},t)\nabla h' + \frac{\partial h}{\partial \mathbf{x}}\right)[\nabla h \nabla h']^{-1}$. This last case is the most general because it contains the other two.

5.2 Directions for future research

There are many direction for the future research.

We compare the speed of the minimization algorithm to the matlab method fmincon. In [13], a new algorithm that adapts its code to the particular minimization problem, has been developed. Comparing both methods would give some more conclusive results about the speed of this new algorithm

Theorem 6 is very complicated and does not make clear links with optimization theory. Many ideas can be explored to simplify it, creating a second order system ($\ddot{\mathbf{x}}$) or creating hierarchical system or feedback systems.

We have always used an identity metric. It is interesting to explore what happens with constant metrics and time varying metrics. In the gradient autonomous case, it can may be include some information of the constraints making the minimization faster.

In chapter 1, there is an example with duality theory. There are many other links to be explored.

When having many different goals, Pareto optimality does not have a general theorem to find the best solution. May be some conclusive theorem can be found exploring Pareto optimality using contraction theory.

Contraction theory has very interesting combination properties. In chapter one an example using parallel combination has been presented. It is interesting to try to use feedback and hierarchical combination.

The Hamilton-Jacobi-Bellman (HJB) equation is a partial differential equation which is central to optimal control theory. It is the solution of a minimization problem subject to a dynamic system. As chapter 4 has time varying constraints, it is reasonable to explore this problem.

Bibliography

[1] Lohmiller, W., and Slotine, J.J.E., On Contraction Analysis for Nonlinear Systems, Automatica, 34(6), 1998.

[2] Bertsekas, Dimitri P., Athena Scientific, Nonlinear programming, 1995

[3] J.J. Slotine, Modular Stability Tools for Distributed Computation and Control J. Adaptive Control and Signal Processing, 2002

[4] David G. Luenberger, The Gradient Projection Method Along Geodesics, Management science, 1972.

[5] Personal Communication

[6] Gignoux, Claude, Silvestre-Brac, Bernard, Solved Problems in Lagrangian and Hamiltonian Mechanics, Grenoble Sciences, 2004.

[7] Wolfgang Kuhnel, Differential Geometry: Curves - Surfaces - Manifolds, AMS, 2006

[8] P.-A. Absil, K. Kurdyka, On the stable equilibrium points of gradient systems , Systems and control letters. 2006

[9] Lohmiller, W., and Slotine, J.J.E. (2000b). Nonlinear Process Control Using Contraction Theory, A.I.Ch.E. Journal.

[10] Nadav Kashtan, Elad Noor, and Uri Alon, Varying environments can speed up evolution, PNAS, 2007

[11] Constantin Udriste, Convex Functions And Optimization Methods On Riemannian Manifolds, Springer, 1994.

[12] Slotine, Jean-Jacques and Li, Weiping, Prentice Hall,Applied Nonlinear Control, 1990

[13] Jacob Mattingley and Stephen Boyd, Real-Time Convex Optimization in Signal Processing, IEE Signal Processing magazine, May 2010

[14] Luenberger, D.G. (1979). Introduction to dynamic systems, Wiley.

[15] Arnold, V.I. (1978). Mathematical Methods of Classical Mechanics, Springer Verlag.

[16] Zuo, L., and Slotine, J.J.E. (2003). Robust vibration isolation using frequency shaped sliding variables, MIT Nonlinear Systems Laboratory, Report MIT-NSL-030201, submitted.

[17] Petros A. Ioannou, Bari Fidan, Adaptive control tutorial, SIAM, 2006

[18] Lohmiller Winfried, J.J. Slotine, Exact Modal Decomposition of Nonlinear Hamiltonian Systems Hamiltonian decomposition, American Institute of Aeronautics and Astronautics

[19] Amari S.I. Natural Gradient Works Efficiently in Learning, RIKEN Frontier Research Program, Saitama 351-01, Japan

[20] Goldstein H., Classical Mechanics, Addison Wesley, 1980.

[21] Uwe Helmke (Author), John B. Moore (Author), R. Brockett, Optimization and Dynamical Systems, Communications and Control Engineering.

[22] On the duality of strong convexity and strong smoothness: Learning applications and matrix regularization, Sham M. Kakade and Shai Shalev-Shwartz and Ambuj Tewari, Toyota Technological Institute Chicago, USA