

RECEDING HORIZON COVARIANCE CONTROL

A Thesis

by

ERIC DUONG BA WENDEL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2012

Major Subject: Aerospace Engineering

RECEDING HORIZON COVARIANCE CONTROL

A Thesis

by

ERIC DUONG BA WENDEL

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Raktim Bhattacharya
Committee Members,	Colleen Robles Igor Zelenko
Head of Department,	Dimitris Lagoudas

August 2012

Major Subject: Aerospace Engineering

ABSTRACT

Receding Horizon Covariance Control. (August 2012)

Eric Duong Ba Wendel, B.S., University of California Berkeley

Chair of Advisory Committee: Dr. Raktim Bhattacharya

Covariance assignment theory, introduced in the late 1980s, provided the only means to directly control the steady-state error properties of a linear system subject to Gaussian white noise and parameter uncertainty. This theory, however, does not extend to control of the transient uncertainties and to date there exist no practical engineering solutions to the problem of directly and optimally controlling the uncertainty in a linear system from one Gaussian distribution to another. In this thesis I design a dual-mode Receding Horizon Controller (RHC) that takes a controllable, deterministic linear system from an arbitrary initial covariance to near a desired stationary covariance in finite time.

The RHC solves a sequence of free-time Optimal Control Problems (OCP) that directly control the fundamental solution matrices of the linear system; each problem is a right-invariant OCP on the matrix Lie group \mathbf{GL}_n of invertible matrices. A terminal constraint ensures that each OCP takes the system to the desired covariance. I show that, by reducing the Hamiltonian system of each OCP from $T^*\mathbf{GL}_n$ to $\mathfrak{gl}_n^* \times \mathbf{GL}_n$, the transversality condition corresponding to the terminal constraint simplifies the two-point Boundary Value Problem (BVP) to a single unknown in the initial or final value of the costate in \mathfrak{gl}_n^* .

These results are applied in the design of a dual-mode RHC. The first mode repeatedly solves the OCPs until the optimal time for the system to reach the desired covariance is less than the RHC update time. This triggers the second mode, which applies covariance assignment theory to stabilize the system near the desired

covariance. The dual-mode controller is illustrated on a planar system. The BVPs are solved using an indirect shooting method that numerically integrates the fundamental solutions on \mathbb{R}^4 using an adaptive Runge-Kutta method. I contend that extension of the results of this thesis to higher-dimensional systems using either indirect or direct methods will require numerical integrators that account for the Lie group structure. I conclude with some remarks on the possible extension of a classic result called Lie's method of reduction to receding horizon control.

DEDICATION

*To my parents,
Dennis and Chien Wendel
for their constant encouragement and support
and for knowing me better than I know myself.*

ACKNOWLEDGMENTS

I am grateful to my advisor and chair, Dr. Raktim Bhattacharya, for his tutelage, mentorship and advice, and for introducing me to the wonderful worlds of real-time trajectory optimization and uncertainty propagation. I am also grateful to my committee members Dr. Colleen Robles and Dr. Igor Zelenko for many interesting suggestions and valuable comments on my thesis, and for their patient guidance throughout my graduate career.

In this thesis I had a rare opportunity in which to blend my separate interests in receding horizon control, Lie theory and optimal control theory, and I am extremely fortunate to have had Dr. Bhattacharya, Dr. Robles and Dr. Zelenko serve on my committee as experts in these areas. I am also indebted to Abhishek Halder for pointing me to the recent papers by Roger Brockett on covariance control. It was ultimately through those papers that I was able to see how everything might eventually come together.

It would be remiss of me not to also mention Dr. Aaron Ames, who introduced me to control systems theory when I was just an undergraduate 6 years ago in Berkeley. That was a rare and formative experience, and it set me on my current path. Dr. Ames also taught me how to share my ideas and how to alchemize coffee into readable papers as a graduate student in College Station. I am grateful to Karen Knabe and the Aerospace Engineering graduate advising office for always being willing to help and for their seeming illimitable patience with me as I navigated academia's dual world of final defense forms and thesis approval forms.

NOMENCLATURE

$\mathbb{R}^{m \times n}$	The vector space of $m \times n$, \mathbb{R} -valued matrices.
$Df(x)$	The derivative of a smooth map $f : M \rightarrow N$ between manifolds; a linear map $T_x M \rightarrow T_{f(x)} N$, where $x \in M$ and $f(x) \in N$.
$Df(x) \cdot z$	The derivative of f evaluated on a tangent vector $z \in T_x M$.
$\mathcal{X}(M)$	The infinite-dimensional Lie algebra of smooth vector fields on a smooth manifold M .
\mathbf{G}	A matrix Lie group.
$T_H R_X, T_H L_X$	The derivatives of right- and left-multiplication on \mathbf{G} ; for all $X, H \in \mathbf{G}$, $T_H R_X = DR_X(H) : T_H \mathbf{G} \rightarrow T_{HX} \mathbf{G}$ and $T_H L_X = DL_X(H) : T_H \mathbf{G} \rightarrow T_{XH} \mathbf{G}$.
\mathbf{GL}_n	The general linear group of $n \times n$, real-valued invertible matrices.
\mathfrak{gl}_n	The general linear Lie algebra of \mathbf{GL}_n , $\mathfrak{gl}_n = \mathbb{R}^{n \times n}$.
\mathbf{PD}_n	Set of $n \times n$ covariance (positive-definite symmetric) matrices.
\mathbf{Skew}_n	Set of $n \times n$ skew-symmetric matrices.
\mathcal{A}	The set of covariance matrices S that are solutions to the Lyapunov equation $0 = (A+BK)S + S(A+BK)^T$, where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and $K \in \mathbb{R}^{m \times n}$.
λ	The action of a Lie group \mathbf{G} on a manifold M , a smooth map $\lambda : \mathbf{G} \times M \rightarrow M$.
$\text{Orb}_\lambda(S)$	The orbit of the action λ through the point $S \in M$.
\mathbf{H}_S^λ	The isotropy group of the action λ at $S \in M$.
\mathfrak{h}_S^λ	The isotropy algebra of the action λ at $S \in M$.

TABLE OF CONTENTS

	Page
ABSTRACT	iii
DEDICATION	v
ACKNOWLEDGMENTS	vi
NOMENCLATURE	vii
TABLE OF CONTENTS	viii
LIST OF FIGURES	ix
1. INTRODUCTION	1
2. STEADY-STATE COVARIANCE ASSIGNMENT	5
2.1 Uncertainty Propagation	5
2.2 Linear Systems	9
2.3 Covariance Assignment Theory	10
2.3.1 Application to a Planar Linear System	14
3. OPTIMAL COVARIANCE CONTROL	17
3.1 Actions of Lie Groups	18
3.2 Invariant Optimal Control Problems	25
3.2.1 Application to a Planar Linear System	32
4. RECEDING HORIZON COVARIANCE CONTROL	39
4.1 Dual-Mode Receding Horizon Control Strategy	41
4.1.1 Application to a Planar Linear System	44
4.2 Lie's Method of Reduction	46
5. CONCLUSION	48
REFERENCES	51

LIST OF FIGURES

FIGURE	Page
3.1 The running cost along optimal trajectories. The optimal cost at the final time is $\int_0^{1.3493} \frac{1}{2} \langle K, K \rangle dt = 0.759$	36
3.2 The Hamiltonian of a free-time problem should be equal to 0. The error in the $H(P)$ is on the order of 1e-11, an order of magnitude <i>smaller</i> than the integration tolerance.	36
3.3 The eigenvalues of the costate in $\mathbf{g}\mathbf{t}_n^*$ over time. It is apparent that the trajectories are isospectral up to plotting accuracy.	37
3.4 The evolution of the covariance associated with the system. The desired final covariance is reached.	37
3.5 The time-varying entries of the matrix $K(t) \in \mathbb{R}^{1 \times 2}$	38
4.1 The evolution of the covariance under the dual-mode receding horizon controller $\mathcal{RH}(0.05)$. The first mode of the controller brings the covariance as close as possible to the desired and the second applies ensures the covariance remains close to the desired covariance, shown dashed.	45
4.2 The controls corresponding to the dual-mode receding horizon controller $\mathcal{RH}(0.05)$. Note the large discontinuity in the controls in the switch from the first to the second mode of the controller.	46

1. INTRODUCTION

In many engineering applications, system parameters or initial conditions are not known exactly and can only be measured with limited statistical certainty. Examples include aircraft control applications [1] and chemical process control [2], where in both cases the complexity of the physical processes (aero- and fluid-dynamics) and the inability to directly measure the values of key states require a statistical approach to quantifying system behavior. It is therefore necessary to design control systems that can account for the modeled and estimated uncertainty in the states and parameters of a control system; the controllers that can handle uncertainty while satisfying necessary performance and stability requirements are called *robust control systems*.

A typical approach in the theory of robust linear and nonlinear control theory is to design an optimal control problem [3] that simultaneously minimizes a cost function while maximizing a measure that is inversely proportional to the largest magnitude system disturbances. This min-max design is exemplary of a philosophy where one accounts for the uncertainty in the system by designing a controller to handle worst-case behavior. There has been a shift in paradigm in recent years away from such worst-case designs towards risk-aware designs, wherein the actual uncertainty is modeled and indirectly controlled. A popular approach is to minimize the expected value of a suitable cost function in such a way that its minimization implies the minimization of the uncertainty in the system [4].

This thesis adopts a risk-aware design philosophy but differs from most robust controllers in that here we are able to *directly* control the uncertainty in the system. This is accomplished by utilizing a basic result from the field of uncertainty propagation, that the integral curves of a deterministic dynamical system induce a semigroup of so-called *Perron-Frobenius operators*, which are operators on the space of probability distribution functions. It is easy to find simple differential equations

This thesis follows the style of *IEEE Transactions on Automatic Control*.

for the evolution of probability distributions along trajectories of a dynamical system by specializing to the case of linear dynamical systems with normally-distributed uncertainty. In particular, we desire controllers that take the first and second moments of a Gaussian distribution to prescribed values. This thesis focuses on the optimal control of the second moment, or covariance, of a Gaussian distribution. The covariance evolves on the space of symmetric, positive-definite matrices.

Early work by Anthony Hotz and Robert E. Skelton [5] showed that the steady-state solutions of the covariance differential equation can be tackled using basic facts about solutions of linear matrix equations. Their contribution represents the first approach to the direct control of uncertainty in linear systems. Although restricted to the steady-state case, their work paved the way for a characterization of all stabilizing feedback control gains [6] and is the primary inspiration for this thesis. The contributions of this theory may be summarily described as a theory of covariance assignment.

Our objective is to complete this work to a theory of covariance control by designing controllers that take a linear system from an arbitrary distribution of uncertain system states and parameters to a stationary distribution representing desired root-mean-square performance requirements and measures of statistical cross-correlation error. To this end, I extend the work of Roger Brockett [7] and John Baillieul [8] on optimal control of Lie groups and numerically solve an optimal control problem on the fundamental matrix solutions of a linear system. Of particular importance in this effort is the fact that the evolution of the system covariance is obtained by the congruence action of the fundamental solutions on an initial covariance.

The main contribution of this thesis is the implementation of this optimal control problem in a receding horizon fashion. Historically, receding horizon control has been implemented in the chemical process industry [2] as a pragmatic, heuristic approach to obtain a stabilizing controller for systems with high degrees of model and parameter uncertainty. In this context, receding horizon control is sometimes called

model predictive control, because the repeated, real-time solution of an optimization problem at discrete time intervals provides a prediction of the future behavior of the system, based on the given model. As the width of the discrete time intervals approaches zero one obtains an optimal control problem solved at every instant of time, and thus receding horizon control can be viewed as an approximation of a feedback controller. To my knowledge this thesis designs the first receding horizon controller on a matrix Lie group.

A secondary contribution of this thesis is the reduction of the Hamiltonian system of equations from $T^*\mathbf{GL}_n$ to $\mathfrak{gl}_n^* \times \mathbf{GL}_n$. The resulting costate differential equation is the infinitesimal generator of the coadjoint action of \mathbf{GL}_n on \mathfrak{gl}_n^* . In [7] this result is obtained essentially by hand, *i.e.* without recourse to the formal reduction of the cotangent bundle. However, the formal reduction allows us to identify two so-called equations of Lie type appearing in the optimal covariance control problem, the first being the covariance differential equation and the second the costate differential equation. Equations of Lie type are differential equations induced by the action of a Lie group on a manifold. The formal reduction of the cotangent bundle also allows us to effectively describe the transversality condition corresponding to the terminal constraint on the covariance in the optimal control problem. In particular, the transversality condition for the Hamiltonian system on $\mathfrak{gl}_n^* \times \mathbf{GL}_n$ reduces the number of scalar unknowns in the boundary value problem from $2n^2$ to n^2 or $n(n+1)/2$. I show that if it is stable to integrate backwards in time, the boundary value problem has $n(n+1)/2$ many unknowns; integrating forward in time implies n^2 many unknowns.

This thesis makes a rare connection between the work of Skelton, Brockett and Baillieul with the literature on robust optimal control of linear systems. I argue that the successful extension of these results to more general linear systems will require Lie group numerical integrators for solving boundary value problems using either direct (collocation-based integration) or indirect (time-stepping integration) methods. I

conclude with some remarks on a classic technique, called Lie's method of reduction, that has potential in either case to assist in the receding horizon implementation.

The organization of the thesis is as follows: in chapter 2, I briefly review uncertainty propagation, the covariance assignment theory and its limitations. In chapter 3, I review concepts in Lie groups, algebras and Lie group actions, and formulate the necessary conditions for optimality of solutions to the free-time, right-invariant optimal control problem with a constraint on the terminal covariance. In the last chapter 4, I discuss the implementation of the optimal control problem from chapter 3 in a receding horizon fashion and, in the conclusion, how to extend the results to a complete theory of covariance control.

2. STEADY-STATE COVARIANCE ASSIGNMENT

The objective of this chapter is to review concepts necessary for understanding the effects of uncertainty in deterministic dynamical systems [9]. The propagation of uncertainty associated with a dynamical system is defined via a time-varying family of transformations of probability distributions, called the *semigroup of Perron-Frobenius operators*. These operators describe how a continuous transformation of points in a space induce a transformation in the probability distribution over the space in which those points belong, essentially by means of a change-of-variables formula.

More precisely, associated with any 1-parameter family of smooth maps φ_t (more generally, with any 1-parameter family of continuous transformations of M) is a 1-parameter *semi*-group of operators on the space of probability densities. The underlying intuitive notion is that instead of studying transformations of points (initial conditions) in a state space, one may study instead transformations of *point clouds* or *distributions* of initial conditions. This enables the shift in paradigm away from controlling one possible configuration of a system to controlling a distribution of configurations.

By specializing these constructions to the case of linear, time-invariant control systems subject to linear state feedback, we arrive at equations for the first and second moments of a Gaussian probability distribution evolving along trajectories of the system. Then, I review the classic results of steady-state covariance assignment theory, which studies the steady-state covariance equation describing stationary distributions of uncertainty in linear systems.

2.1 Uncertainty Propagation

This section reviews basic concepts in dynamical systems and the theory of uncertainty propagation. There is a neat correspondence between flows of dynamical

systems, their fundamental solutions and a certain semigroup of transfer operators on the space of probability distributions.

Dynamical systems: Let M be a smooth manifold, TM its tangent bundle, and $\pi : TM \rightarrow M$ the canonical projection. A deterministic, smooth *dynamical system* is the pair (M, f) , where $f : M \rightarrow TM$ is a *vector field*, *i.e.*, a smooth assignment of a vector in TM to every point in M . For the sake of simplicity, for the remainder of this section we assume that M is an embedded submanifold of \mathbb{R}^n . An *integral curve* $x(t)$ of f is a trajectory in M tangent to the vector field at all points along the curve, *i.e.* $f(x(t)) = \dot{x}(t)$. For every *initial condition* $x_0 \in M$ there exists an $\epsilon > 0$ and a neighborhood U of x_0 such that x_0 is taken into an integral curve $x(t)$ by the *flow*, which is the map given by $\varphi : (-\epsilon, \epsilon) \times U \rightarrow M$. We adopt the shorthand $\varphi(t, x_0) =: \varphi_t(x_0) = x(t)$. It is known that $\varphi_t : U \rightarrow \varphi_t(U)$ is a diffeomorphism onto its image satisfying

- $\varphi_0(x) = x$,
- $\varphi_{t+s}(x) = \varphi_t \circ \varphi_s(x) = \varphi_s \circ \varphi_t(x)$,
- $\varphi_t^{-1}(x) = \varphi_{-t}(x)$,

for all $t, s \in (-\epsilon, \epsilon)$. In this way we obtain a local *1-parameter family of diffeomorphisms* $\{\varphi_t\}$ of M , $t \in (-\epsilon, \epsilon)$.

Fundamental solutions: As above, let $x(t) = \varphi_t(x_0)$. The second derivative of $x(t)$ obeys the differential equation $\ddot{x}(t) = Df(x(t)) \cdot \dot{x}(t)$, where $Df(x) = \partial f(x)/\partial x$ is the Jacobian matrix of $f(x)$. The *fundamental solutions* of this linear, time-varying differential equation are invertible $n \times n$ matrices denoted $X(t, t_0)$, where $[t_0, t] \subset (-\epsilon, \epsilon)$. The fundamental solutions satisfy the same differential equation as for \ddot{x} ,

$$\dot{X}(t, t_0) = \frac{\partial f(x(t))}{\partial x} X(t, t_0).$$

The solutions of this differential equation are called *fundamental solutions* and live, as we will see, in the matrix Lie group of invertible matrices. Fundamental solutions also satisfy the following properties [10, 11]:

$$(P1) \quad X(t, t) = I_n, \text{ for all } t \in (-\epsilon, \epsilon),$$

$$(P2) \quad X(t + s, t_0) = X(t, s)X(s, t_0),$$

$$(P3) \quad D\varphi_t(x_0) = X(t, t_0), \quad D\varphi_t^{-1}(x_0) = X^{-1}(t, t_0).$$

In the engineering literature $X(t, t_0)$ is called the *state transition matrix* [12, 13], and due to property (P3) is often used as a measure of the sensitivity of the system flow to perturbations in its initial conditions.

Perron-Frobenius operators: Associated with the flow $\varphi_t(x_0)$ and fundamental solution $X(t, t_0)$ is a 1-parameter semigroup of operators on the space of probability distributions.

Definition 2.1.1: The *Perron-Frobenius operator* $P_t : \mathcal{D} \rightarrow \mathcal{D}$ (where \mathcal{D} is the space of probability distributions on the manifold M) associated with the diffeomorphism $\varphi_t : M \rightarrow M$ is a so-called *transfer operator* on the space of probability distributions, defined by the change-of-variables formula

$$P_t \rho = \int_M \rho(\varphi_t^{-1}(x)) \det(D\varphi_t^{-1}(x)) \mu(dx), \quad (2.1)$$

where $\mu(dx)$ is Lebesgue measure on M and $\rho \in \mathcal{D}$ is arbitrary. Furthermore, the family of operators $\{P_t\}_{t \in [0, \epsilon]}$ is a *1-parameter semigroup of Perron-Frobenius operators* satisfying

- $P_0 \rho = \rho,$
- $P_{t+s} \rho = P_t \circ P_s \rho = P_s \circ P_t \rho,$

and describes the time-evolution of ρ induced by the flow.

Remark 2.1.2: The Perron-Frobenius operator describes how some initial distribution of system uncertainty $\rho \in \mathcal{D}$ is propagated forward in time by the system dynamics. The spectral properties of the operator are also key to the identification and global analysis of the invariant sets of a dynamical system; see [9, 14] for discussion and examples of this approach to uncertainty propagation. This thesis, on the other hand, takes a local approach in the sense that we attempt to control the evolution of uncertainty along particular integral curves of (M, f) .

The following proposition characterizes infinitesimal changes in the system uncertainty. A proof can be found in [9].

Proposition 2.1.1 (Lasota & Mackey): Denote the semigroup of Perron-Frobenius operators associated with the flow φ_t of a deterministic dynamical system by $\rho_t(x) = \rho(t, x) = P_t \rho_0(x)$, where $\rho_0 \in \mathcal{D}$ is fixed and given. The *Liouville equation*

$$\frac{\partial}{\partial t} \rho_t(x) + \sum_{i=1}^n \frac{\partial}{\partial x_i} (\rho_t(x) f(x)) = 0, \quad (2.2)$$

is the unique *infinitesimal generator* of $\rho_t(x)$.

Remark 2.1.3: Although this thesis is only concerned with deterministic dynamics, it is worth noting that the infinitesimal generator of the semigroup of Perron-Frobenius operators when the dynamical system is *stochastic* is a partial differential equation called the *Fokker-Planck* equation. It is the counterpart of the Liouville equation [9], and describes the propagation of system uncertainty when the system evolution is itself random.

Throughout the remainder of this thesis we will be concerned with the analysis and control of solutions to (2.2) when $f(x)$ is a linear control system on \mathbb{R}^n .

2.2 Linear Systems

Consider the differential equation $\dot{x} = Ax + Bu$. The parameter $u = u(t) \in \mathbb{R}^m$ is a column vector of functionals called the *controls*, $A \in \mathbb{R}^{n \times n}$ is the *drift matrix*, and $B \in \mathbb{R}^{n \times m}$ is a full column-rank *control matrix*. This so-called linear time-invariant control system has solution [15]

$$x(t) = \varphi_{t-t_0}(x_0) = X(t, t_0)x_0 + \int_{t_0}^t X(t, \tau) B u(\tau) d\tau, \quad (2.3)$$

where $X(t, t_0)$ is the fundamental matrix solution. We set $t_0 = 0$ without loss of generality and, for notational convenience, write $X(t) := X(t, 0)$ and $\mathcal{I} = [0, t] \subset [0, \infty)$.

Linear feedback control structure: We are interested in the case where the controls depend linearly on the states, *i.e.* $u = Kx$, where $K : \mathcal{I} \rightarrow \mathbb{R}^{m \times n}$ is a matrix of time-varying *control gains*. Having fixed this *closed-loop feedback structure* the dynamics become

$$\dot{x} = (A + BK)x, \quad (2.4)$$

and the fundamental solution $X(t)$ satisfies the differential equation $\dot{X} = (A + BK)X$. Equation (2.3) for the integral curves of (2.4) simplifies to $x(t) = \varphi_t(x_0) = X(t)x_0$.

This feedback structure admits simple expressions for the time evolution of the moments of the propagated uncertainty. Recall that a Gaussian probability distribution function is of the form

$$\rho(x) = \frac{1}{\sqrt{(2\pi)^n \det(S)}} \int_M \exp\left(-\frac{1}{2}(x - \mu)^T S^{-1}(x - \mu)\right) \mu(dx),$$

where $\mu(dx)$ is Lebesgue measure on M , S is the covariance and μ is the mean of the distribution. Let us also assume that φ_t , $t \in \mathcal{I}$, is defined for all $x \in M$. We have the following general result on the moments of Gaussian probability distributions transformed by the flow.

Proposition 2.2.1: Let $\rho_0(x)$ be a Gaussian probability distribution with mean μ_0 and covariance S_0 and let $\rho(x, t) = P_t\rho_0(x)$ be the probability distribution induced by the flow φ_t of the linear system (2.4). Then, $\rho(x, t)$ is Gaussian, with mean $\mu(t) = X(t)\mu_0$ and covariance $S(t) = X(t)S_0X^T(t)$ satisfying the differential equations

$$\begin{aligned}\dot{\mu}(t) &= (A + BK)\mu(t) \\ \dot{S}(t) &= (A + BK)S(t) + S(t)(A + BK)^T.\end{aligned}\tag{2.5}$$

Proof. Note that $\varphi_t^{-1}(x) = X(t)^{-1}x$ and $\det(D\varphi_t(x)) = \det X(t)$, for all $x \in M$. Then equation (2.1) becomes

$$P_t\rho_0(x) = \int_M \frac{\exp\left(-\frac{1}{2}(x - X(t)\mu_0)^T X(t)^{-T} S_0^{-1} X(t)^{-1} (x - X(t)\mu_0)\right)}{\det(X(t)) \sqrt{(2\pi)^n \det(S_0)}} \mu(dx)$$

Define $S(t) := X(t)S_0X^T(t)$, $\mu(t) = X(t)\mu_0$ and differentiate directly to obtain equations (2.5). The result follows by noting that $\det(S(t)) = \det(X(t))^2 \det(S_0)$ and substituting these expressions into the above equation. \square

Remark 2.2.1: It is not always necessary to solve the Liouville equation (2.2) in order to compute the propagation of uncertainty along integral curves of a dynamical system. In the derivation of Proposition 2.2.1 we implicitly relied on the fact that any continuous linear transformation induces a Frobenius-Perron operator that preserves normal distributions. In the case of arbitrary nonlinear systems, computing $P_t\rho_0(x)$ for any ρ_0 along integral curves is possible using Liouville's formula [16, pg. 34]; see [17] for examples and applications.

2.3 Covariance Assignment Theory

This section reviews the main results of covariance assignment theory, as introduced by Anthony Hotz and Robert E. Skelton in their seminal 1987 paper [5]. The objective of covariance assignment theory is to develop an effective means of

controlling linear system (2.4) to achieve specific steady-state performance requirements in the presence of uncertainty. For example, the diagonal entries of a so-called *assignable* covariance matrix may be interpreted as root-mean-square values of the system states, while off-diagonal entries indicate the amount of statistical cross-correlation between states. As a consequence, an appropriate choice of covariance matrix specifies steady-state system behavior. More precisely, it specifies the steady-state distribution of system states.

It is well-known that a given matrix $A \in \mathbb{R}^{n \times n}$ is *stable* (it has eigenvalues with negative real part) if and only if there exists a matrix $S \in \text{PD}_n$ satisfying the *Lyapunov equation* $0 = AS + SA^T + Q$, where $Q \in \text{PD}_n$ is given. There exist many efficient numerical methods [18–20] that search for a matrix S given A and Q .

The classic steady-state covariance control problem, first solved by Anthony Hotz and Robert E. Skelton in [5], considers a Lyapunov equation that has special meaning for linear systems subject to stochastic dynamics. Specifically, [5, 6] consider the problem of finding the unknown matrices $K \in \mathbb{R}^{m \times n}$ and $S \in \text{PD}_n$ that solve

$$0 = (A + BK)S + S(A + BK)^T + Q,$$

where $Q \in \text{PD}_n$ is the given covariance of a white Gaussian process modeling stochastic disturbances and noise; the stochastic effects are intended to account for unmodeled and unknown deterministic forces and sensor noise.

This thesis is only interested in the analysis of deterministic (smooth) dynamical systems. We obtain a statement of the steady-state covariance control problem for smooth dynamical systems by omitting the covariance matrix from the Lyapunov equation.

Definition 2.3.1: The *covariance assignment problem* is the problem of simultaneously finding a constant control gain $K \in \mathbb{R}^{m \times n}$ and a symmetric, positive-definite matrix $S \in \text{PD}_n$ solving the (deterministic) *steady-state covariance equation*

$$0 = (A + BK)S + S(A + BK)^T, \quad (2.6)$$

where A is a drift matrix and B a control matrix of a linear control system.

Remark 2.3.2: Note that $B \in \mathbb{R}^{n \times m}$ and $K \in \mathbb{R}^{m \times n}$ and so $BK = \sum_{i=1}^n B_i K^i$ is clearly a sum of rank-1 matrices, where B_i is the i^{th} column of B and K^i is the i^{th} row of K . Consequently, the covariance assignment problem as defined is the problem of finding the set of allowable rank-1 updates to the drift matrix A such that a quadratic form defined by $S \in \text{PD}_n$ is preserved along trajectories of the linear system.

To see this, define the function $q_S : \mathbb{R}^n \rightarrow \mathbb{R}$ by $q_S(x) = x^T S^{-1} x$. If the differential of q_S is 0 along all trajectories of the linear system $\dot{x} = (A + BK)x$ then

$$0 = \frac{d}{dt} q_S(x) = x^T (A + BK)^T S^{-1} x + x^T S^{-1} (A + BK) x$$

for all $x \in \mathbb{R}^n$ yields the equivalent condition $(A + BK)^T S^{-1} + S^{-1} (A + BK) = 0$, which can be rearranged into (2.6) by multiplying on both sides by S .

Although we confine our attention to deterministic dynamical systems, we may still use classic steady-state covariance control theory to describe the set of all feedback gains K solving (2.6). We obtain the following corollary by omitting the covariance of the white Gaussian process from the statement of Theorem 3.1 in [6].

First, recall that we defined the control matrix $B \in \mathbb{R}^{n \times m}$ to have full column rank, so $\text{rank}(B) = m$. Let B^+ denote its Moore-Penrose pseudoinverse, a map from the range of B to its nullspace, $B^+ : \mathcal{R}(B) \rightarrow \mathcal{N}(B)$. Then, $\Pi^\perp := BB^+$ and $\Pi := I - BB^+$ are orthogonal projections onto $\mathcal{R}(B)$ and $\mathcal{N}(B)$, respectively, where I is the $n \times n$ identity matrix.

Corollary 2.3.1 (Ohara & Kitamori): Any covariance matrix $S \in \text{PD}_n$ belonging to the convex subset

$$\mathcal{A} := \{S \in \text{PD}_n : \Pi(AS + SA^T)\Pi = 0\}, \quad (2.7)$$

is a solution to (2.6), and moreover assignable by the constant feedback gain

$$K = -B^+(AS + SA^T)(I - \frac{1}{2}\Pi^\perp)S^{-1} - B^+QS^{-1}, \quad (2.8)$$

where $Q = \Pi^\perp Q = \Pi^\perp Q \Pi^\perp \in \text{Skew}_n$ is any skew-symmetric matrix preserving $\mathcal{R}(B)$.

The set of assignable covariances $\mathcal{A} \subset \text{PD}_n$ is convex [21] by virtue of $\Pi(AS + SA^T)\Pi = 0$ being linear in S . Because we seek a steady-state covariance matrix representing desirable steady-state behavior, in typical applications it is not sufficient to simply choose any symmetric, positive-definite matrix from \mathcal{A} . We address this issue by breaking the covariance assignment problem into two separate steps:

- Step 1. Find a $S \in \mathcal{A}$ satisfying steady-state performance requirements expressed as linear inequality and equality constraints on the matrix entries of S .
- Step 2. Specify the control gain (2.8) assigning $S \in \mathcal{A}$ by selecting a matrix $W \in \text{Skew}_n$ that preserves $\mathcal{R}(B)$ (such that $\mathcal{R}(B)$ is an invariant subspace of the linear transformation W).

It is natural to attempt to solve Step 1 using *semidefinite programming* [22], because \mathcal{A} is a convex set and such programs often have fast and efficient numerical implementations. Although other approaches certainly exist [21, 23], semidefinite programming is still a natural choice because it is widely applicable in linear systems theory [20]. The minimizing argument, $S \in \mathcal{A}$, of the following semidefinite program

is the assignable covariance that is closest in the sense of *relative entropy* to a given ideal system covariance matrix $\hat{S} \in \text{PD}_n$ that might not be assignable.

$$\begin{aligned} \min_{S \in \mathcal{A}} \quad & \text{tr}(\hat{S}^{-1} S) - \log \det(S) \\ \text{s. t.} \quad & QS \leq \alpha, \quad CS = 0. \end{aligned} \tag{2.9}$$

The matrices Q and C represent linear inequality and equality constraints on the entries of S , in addition to constraint (2.7).

The relative entropy of two normal distributions is not a distance metric but a *divergence measure* on the set of symmetric, positive-definite matrices, and plays a significant role in information geometry [24] and Markov process stability theory [9, Ch. 9], and is ubiquitous in semidefinite programming [22, 25]. It is of course possible to replace the objective in (2.9) with any convex measure of distance on PD_n . This solves Step 1. To address the problem of choosing a control gain assigning $S \in \mathcal{A}$, it is sufficient to choose a constant skew-symmetric matrix W such that $W = \Pi^\perp W = \Pi^\perp W \Pi^\perp$. This matrix can be solved for algebraically or with a root-finding program (such as Matlab's `fsolve`).

2.3.1 Application to a Planar Linear System

As a brief illustration of covariance assignment, let us consider the planar linear system $\dot{x} = (A + BK)x$, where

$$\begin{aligned} A &= \begin{pmatrix} -1.417 & 1 \\ 2.86 & -1.183 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ -3.157 \end{pmatrix}, \\ \Pi^\perp &= BB^+ = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \Pi = I - \Pi^\perp = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned} \tag{2.10}$$

We seek the assignable covariance $S \in \mathcal{A}$ that is closest in the sense of relative entropy to the ideal covariance $\hat{S} = \frac{1}{10}I \notin \mathcal{A}$. The ideal covariance represents a desired

steady-state variance in each state variable of 0.1 units with no cross-correlation. Setting $Q = C = 0$ in the semidefinite program (2.9), we solve using `cvx` [18] and obtain

$$S = \begin{pmatrix} 0.033246 & 0.047109 \\ 0.047109 & 0.16675 \end{pmatrix}.$$

For this system, the control matrix B is such that there is no nontrivial matrix $W \in \text{Skew}_n$ satisfying $W = \Pi^\perp W \Pi^\perp$, so $K = (2.4947, -0.82357)$ is the only solution to $0 = (A + BK)S + S(A + BK)^T$.

I extend this example in the next two chapters, and show that a receding horizon controller can bring the system from an arbitrary covariance close to a desired. It is not possible to extend the methodology of [5, 6] to the transient control of the covariance; a different approach is needed in order to tackle the complete problem of optimal covariance control. Of course, the optimal steady-state covariance control problem was considered in [21], but the main contribution of that work was the extension of [5] with the analytical solution of a Procrustes matrix nearness problem [26], and also unfortunately does not extend to the transient case.

A geometric formulation of the problem has the potential for a complete perspective of the covariance control problem. For example, consider the following interpretation of the covariance assignment problem afforded by the fundamental solutions of the linear system. If the linear system $\dot{x} = (A + BK)x$ is stationary at a desired covariance S_d , then $S(t) = X(t) S_d X^T(t) = S_d$ for all t , and differentiating this expression yields (2.6). Because $X(t)$ is an invertible matrix, it is immediate that the fundamental solutions assigning the desired covariance evolve in the so-called *isotropy group* of S_d , a subgroup of the Lie group of invertible matrices. The isotropy group is an example of a quadratic matrix Lie group and equation (2.6) is the defining condition for the closed-loop matrix $A + BK$ to be in its Lie algebra. See [27] for definitions and discussion. From this geometric perspective we may consider an optimal covariance assignment problem using Pontryagin's Maximum Principle and

the language of Lie groups and Lie algebras. This is the approach taken in [28] for a driftless system on several classic Lie groups.

In the next chapter we adopt this geometric perspective for the transient covariance control problem, providing optimal, time-varying control gains that take a linear system from some initial arbitrary covariance to a desired covariance. Future work will consider the stationary case as an optimal control problem in a quadratic matrix Lie group.

3. OPTIMAL COVARIANCE CONTROL

In this chapter I solve the problem of optimally stabilizing a linear system from some arbitrary covariance to a desired covariance. This problem is naturally formulated in the language of Lie group actions on manifolds and, as a consequence, we obtain two separate examples of so-called *equations of Lie type*. These are differential equations induced by the action of fundamental solutions of an invariant vector field. The first equation of Lie type is the covariance differential equation, induced by the left action by congruence on \mathbf{PD}_n . The second is the costate differential equation on \mathfrak{gl}_n^* , induced by the left coadjoint action of the fundamental solutions. These facts have many possible implications for the numerical solution of the boundary value problem arising from solution of the optimal controls, and I discuss the limitations and advantages of some indirect and direct methods.

The starting point for the results of this chapter are classic results on optimal control in Lie groups and coset spaces [8,29]. The connection between optimal control of fundamental solutions and covariance control was first made in [7]. This chapter extends [7] with transversality conditions that numerical experimentation shows are necessary for either direct and indirect numerical methods of solving the optimal control problem to succeed. Finally, I show that formulating the optimal control problem on $\mathfrak{gl}_n^* \times \mathbf{GL}_n$ allows the resulting boundary value problem to be reduced to the determination of a single initial condition for the system costate. If a stable backwards numerical integrator is available, the problem can be even further reduced from a problem with n^2 unknowns to $n(n+1)/2$ unknowns.

To my knowledge, this chapter represents the first attempt at a practical engineering solution to the covariance control problem while recasting it in modern geometric language.

3.1 Actions of Lie Groups

Lie groups and Lie algebras: A *matrix Lie group* \mathbf{G} is a subset of $\mathbb{R}^{n \times n}$ endowed with the structure of both a topological group and a smooth manifold, meaning that the group operations of multiplication and inversion are smooth. Matrix Lie groups appear frequently in engineering applications. The special orthogonal group \mathbf{SO}_3 is often used in spacecraft attitude control [30] and the special Euclidean group \mathbf{SE}_3 appears frequently in robotics [31]. Both of these Lie groups are subgroups of the *general linear group*,

$$\mathbf{GL}_n = \{X \in \mathbb{R}^{n \times n} : \det(X) \neq 0\},$$

a n^2 -dimensional manifold with two connected components, $\mathbf{GL}_n = \mathbf{GL}_n^+ \sqcup \mathbf{GL}_n^-$. The component containing the identity is $\mathbf{GL}_n^+ := \{X \in \mathbf{GL}_n : \det(X) > 0\}$, while $\mathbf{GL}_n^- := \{X \in \mathbf{GL}_n : \det(X) < 0\}$. As we will see, due to property (P1) of fundamental matrix solutions, optimal trajectories of covariance matrices are always generated by curves in \mathbf{GL}_n^+ .

For any $X, Y \in \mathbf{G}$, denote right- and left- matrix multiplication by $R_X(Y) = YX$ and $L_X(Y) = XY$. The maps L_X and R_X are diffeomorphisms of \mathbf{G} and their derivatives linear isomorphisms of the tangent spaces of \mathbf{G} : $T_H L_X(Z) = XZ$ and $T_H R_X(Z) = ZX$ for all $H \in \mathbf{G}$, $Z \in T_H \mathbf{G}$.

Let $\mathcal{X}(\mathbf{G})$ denote the set of all smooth vector fields on \mathbf{G} . The *Lie bracket* of two vector fields $F, G \in \mathcal{X}(\mathbf{G})$, or *Jacobi-Lie bracket* [27], is given by

$$[F, G](X) := DG(X) \cdot F(X) - DF(X) \cdot G(X), \quad (3.1)$$

where $DG(X)$ is the Jacobian matrix or derivative of the vector field G and $DG(X) \cdot Z$ is the Jacobian evaluated on a vector $Z \in T_X \mathbf{G}$. A vector field $F \in \mathcal{X}(\mathbf{G})$ is called *right-invariant* if it satisfies $T_X R_H F(X) = F(XH)$ and *left-invariant* if it satisfies

$T_X L_H F(X) = F(HX)$. Denote the set of all left-invariant vector fields on \mathbf{G} by $\mathcal{X}_L(\mathbf{G})$, and the set of all right-invariant vector fields by $\mathcal{X}_R(\mathbf{G})$.

Remark 3.1.1: The fundamental solutions of a nonlinear dynamical system are integral curves of a right-invariant vector field on \mathbf{GL}_n^+ . To see this, recall from section 2.1 that if $\dot{x} = f(x)$ is a vector field on \mathbb{R}^n with flow $\varphi_t(x_0)$ then $F(X(t)) = A(t)X(t)$ is the vector field for the fundamental matrix solution $X(t)$, an invertible $n \times n$ matrix, where $A(t) := Df(\varphi_t(x_0))$. We readily verify that, for any $Y_0 \in \mathbf{GL}_n^+$, $T_X R_{Y_0} F(X) = A(t)XY_0 = F(XY_0)$. More to the point, with $X(t)$ an integral curve of F through the identity at $t = 0$ (due to property (P1)), $X(t)Y_0$ is the integral curve of F through Y_0 at $t = 0$.

The tangent space at the identity, $T_I \mathbf{G}$, is isomorphic to $\mathcal{X}_L(\mathbf{G})$, with the isomorphism provided by mapping left-invariant vector fields to their value at the identity and mapping vectors at the identity to their left-translate in \mathbf{G} . For example, if $F(X) \in \mathcal{X}_L(\mathbf{G})$ such that $F(I) = Z \in T_I \mathbf{G}$, then the map $F \mapsto Z$ has inverse mapping $Z \mapsto T_I L_X Z = F(X)$.

An *abstract Lie algebra* is a vector space V equipped with a bilinear, skew-symmetric operator $[\cdot, \cdot] : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathfrak{g}$, satisfying the Jacobi identity $[Y, [Z, W]] + [Z, [W, Y]] + [W, [Y, Z]] = 0$, for all $Y, Z, W \in T_I \mathbf{G}$. The *Lie algebra \mathfrak{g} of the Lie group \mathbf{G}* is defined to be the abstract Lie algebra on the space $T_I \mathbf{G}$ equipped with a Lie bracket defined by the Jacobi-Lie bracket of left-invariant vector fields evaluated at the identity, as follows: let $Z, W \in T_I \mathbf{G}$ and $F, G \in \mathcal{X}_L(\mathbf{G})$ such that $Z = F(I)$ and $W = G(I)$. Then,

$$[Z, W] := [F, G](I) = DG(I) \cdot F(I) - DF(I) \cdot G(I). \quad (3.2)$$

Note that the Lie bracket of $F, G \in \mathcal{X}_L(\mathbf{G})$ is again left-invariant, *e.g.* if $F(X) = XA$ and $G(X) = XB$ then $[F(X), G(X)] = X[A, B]$. On the other hand, if $F, G \in \mathcal{X}_R(\mathbf{G})$ and $F(X) = AX$, $G(X) = BX$ then $[F(X), G(X)] = -[A, B]X$. For future reference

we introduce the *exponential map* $\exp : \mathfrak{g} \rightarrow \mathbf{G}$ which, when \mathbf{G} is a matrix Lie group, is simply the matrix exponential $\exp(A) = \sum_{i=0}^{\infty} \frac{1}{i!} A^i$, $A \in \mathfrak{g}$.

For example, the Lie algebra of the general linear group is the *general linear algebra*

$$\mathfrak{gl}_n = T_I \mathbf{GL}_n = \mathbb{R}^{n \times n}.$$

Applying equation (3.2), the Lie bracket of left-invariant vector fields $A, B \in \mathfrak{gl}_n$ is simply the matrix commutator $[A, B] = AB - BA$.

The fundamental correspondences between matrix Lie groups and their Lie algebras are often referred to as *Lie's theorems*. Concise descriptions of these results can be found in [32, Ch. 1] and [27].

Actions of Lie groups on manifolds: Lie group actions are central to an understanding of the role of symmetry in differential equations. We use Lie group actions in this thesis to describe the solutions of the covariance differential equation (2.5) and to understand the extremal solutions of invariant optimal control problems on Lie groups in their proper context.

Definition 3.1.2: The *left action* of a Lie group \mathbf{G} on a manifold M is a smooth map $\lambda : \mathbf{G} \times M \rightarrow M$ such that $\lambda(I, S) = S$, where $I \in \mathbf{G}$ is the identity, and $\lambda(X, \lambda(Y, S)) = \lambda(XY, S)$. If, instead, $\lambda(X, \lambda(Y, S)) = \lambda(YX, S)$, λ is called a *right action*. We often adopt the shorthand $\lambda_X(S) := \lambda(X, S)$.

A left- or right-action is said to be:

- *transitive* if for all $R, S \in M$ there exists a $X \in \mathbf{G}$ such that $\lambda(X, R) = S$;
- *effective* if $\lambda(X, S) = S$ for all $S \in M$ implies that $X = I$;
- *free* if $\lambda(X, S) = S$ for any $S \in M$ implies that $X = I$, *i.e.* the map $\lambda(X, \cdot) : M \rightarrow M$ has no fixed points.

The *orbit* of λ through $S \in M$ is the subset $\text{Orb}_\lambda(S) = \{\lambda(X, S) : X \in \mathbf{G}\} \subset M$. A *homogeneous space* is a manifold M with a transitive action λ ; that is, it is a space with only one orbit.

Definition 3.1.3: The *action by congruence* of GL_n on PD_n is the transitive, left Lie group action $\Theta : \mathrm{GL}_n \times \mathrm{PD}_n \rightarrow \mathrm{PD}_n$ defined by $\Theta_X(S) := \Theta(X, S) = X S X^T$.

The transitivity is a consequence of Sylvester's Law of Inertia [19], which states that any two real symmetric matrices are congruent provided they have the same eigenvalue signature. PD_n is therefore a homogeneous space; see [32] for details.

Recall from section 2.1, Proposition 2.2.1 that the covariance of a linear dynamical system with linear feedback controls $u = Kx$ is given by $S(t) = X(t) S_0 X(t)^T = \Theta_{X(t)}(S_0)$, where $X(t)$ is a fundamental solution and S_0 the covariance given at time $t = 0$. We say that the solutions of the covariance differential equation (2.5) are generated by the action by congruence of fundamental solutions on PD_n .

For the remainder of this thesis let the inner product on $\mathbb{R}^{n \times n}$ be given by the trace pairing of $A, B \in \mathbb{R}^{n \times n}$: $\langle A, B \rangle = \mathrm{tr}(A^T B)$. Denote the inverse transpose of $X \in \mathrm{GL}_n$ with the shorthand $X^{-T} := (X^{-1})^T$.

The following actions arise naturally in optimal control problems on Lie groups.

Definition 3.1.4: The *adjoint action* of \mathbf{G} on \mathfrak{g} is the left-action $\mathrm{Ad} : \mathbf{G} \times \mathfrak{g} \rightarrow \mathfrak{g}$ defined by $\mathrm{Ad}_X(A) := \mathrm{Ad}(X, A) = T_I(R_{X^{-1}} \circ L_X)(A)$.

Definition 3.1.5: The *coadjoint action* of \mathbf{G} on \mathfrak{g}^* is the left-action $\mathrm{Ad}^* : \mathbf{G} \times \mathfrak{g}^* \rightarrow \mathfrak{g}^*$ defined by $\mathrm{Ad}_X^*(P) := \mathrm{Ad}^*(X, P) = (\mathrm{Ad}_{X^{-1}})^*(P)$, where $(\mathrm{Ad}_{X^{-1}})^*$ is the dual of $\mathrm{Ad}_{X^{-1}}$, i.e., $(\mathrm{Ad}_{X^{-1}})^*(P) = (T_I(R_X \circ L_{X^{-1}}))^*(P)$.

For the case $\mathbf{G} = \mathrm{GL}_n$, the dual of $\mathrm{Ad}_{X^{-1}}(A) = X^{-1} A X$ is given by

$$\begin{aligned} \langle (\mathrm{Ad}_{X^{-1}})^*(P), A \rangle &= \langle P, \mathrm{Ad}_{X^{-1}}(A) \rangle = \langle P, X^{-1} A X \rangle \\ &= \langle X^{-T} P X^T, A \rangle, \end{aligned} \tag{3.3}$$

for all $X \in \mathrm{GL}_n$, $A \in \mathfrak{gl}_n$, $P \in \mathfrak{gl}_n^*$. That is, $\mathrm{Ad}^*(X, P) = X^{-T} P X^T$ is the coadjoint action. The adjoint action is, of course, $\mathrm{Ad}(X, A) = X A X^{-1}$.

A *group representation* of a Lie group \mathbf{G} is a homomorphism $\phi : \mathbf{G} \rightarrow \mathrm{Aut}(V)$, where $\mathrm{Aut}(V)$ is the group of automorphisms of a vector space V . The adjoint and

coadjoint actions define the adjoint and coadjoint representations $\text{Ad} : \mathbf{G} \rightarrow \text{Aut}(\mathfrak{g})$ and $\text{Ad}^* : \mathbf{G} \rightarrow \text{Aut}(\mathfrak{g}^*)$ by $X \mapsto \text{Ad}_X$ and $X \mapsto (\text{Ad}_{X^{-1}})^*$. See [33] for an application of concepts from representation theory to optimal control problems on Lie groups.

We now consider infinitesimal descriptions of Lie group actions. The infinitesimal characterization is useful for describing tangent spaces to orbits of actions [27] and necessary for defining so-called equations of Lie type [34]. Let $\mathcal{X}(M)$ denote the infinite-dimensional Lie algebra of smooth vector fields on M .

Definition 3.1.6: Given either a left- or right-action $\lambda : \mathbf{G} \times M \rightarrow M$, define a curve in M associated to some $A \in \mathfrak{g}$ by $\varphi_A^\lambda(t, S) := \lambda(\exp(tA), S)$. The *infinitesimal generator* of this curve is the vector field $\Phi_A^\lambda \in \mathcal{X}(M)$ defined by

$$\Phi_A^\lambda(S) = \left. \frac{d}{dt} \right|_{t=0} \lambda(\exp(tA), S).$$

The curve $\varphi_A^\lambda(t, S)$ is an integral curve of $\Phi_A^\lambda(S)$ for all $S \in M$ by construction.

The infinitesimal generator for congruence is

$$\Phi_A^\Theta(S) = \left. \frac{d}{dt} \right|_{t=0} \exp(tA) S \exp(tA)^T = A S + S A^T, \quad (3.4)$$

for all $A \in \text{GL}_n$ and $S \in \text{PD}_n$.

We reproduce the calculations for the infinitesimal generator of the adjoint and coadjoint actions from [27, Ch. 9.1, 9.3] for the sake of completeness. The infinitesimal generator of Ad is the *adjoint operator*, which we denote

$$\Phi_A^{\text{Ad}}(B) =: (\text{ad } A)(B) = [A, B],$$

for all $A, B \in \mathfrak{g}$. For the coadjoint action Ad^* , we compute

$$\begin{aligned} \langle \Phi_A^{\text{Ad}^*}(P), B \rangle &= \left\langle \frac{d}{dt} \Big|_{t=0} \text{Ad}_{\exp(tA)}^*(P), B \right\rangle = \frac{d}{dt} \Big|_{t=0} \langle P, \text{Ad}_{\exp(-tA)}(B) \rangle \\ &= \langle P, -(\text{ad } A)(B) \rangle = \langle -(\text{ad } A)^*(P), B \rangle \end{aligned}$$

for all $B \in \mathfrak{g}$, $P \in \mathfrak{g}^*$. Therefore, the infinitesimal generator of the coadjoint action is the negative of the dual of the adjoint operator: $\Phi_A^{\text{Ad}^*}(P) = -(\text{ad } A)^*(P)$. For the case $\mathbf{G} = \mathbf{GL}_n$ and $\mathfrak{g} = \mathfrak{gl}_n$, $\Phi_A^{\text{Ad}^*}(P) = [P, A^T]$.

Equations of Lie type: In preparation for our brief discussion of Lie's method of reduction in chapter 4 we introduce a special type of differential equation that is intimately related to Lie group actions. The infinitesimal generator of an action plays a key role.

It is well-known that for a left action $\lambda : \mathbf{G} \times M \rightarrow M$ the map $\Phi^\lambda : \mathfrak{g} \rightarrow \mathcal{X}(M)$, defined by $A \mapsto \Phi_A^\lambda$, is a Lie algebra *anti*-homomorphism, *i.e.*

$$\Phi_{[A,B]}^\lambda = -[\Phi_A^\lambda, \Phi_B^\lambda]$$

for all $A, B \in \mathfrak{g}$. See [27, 34] for proofs.

It is possible to make Φ^λ into a Lie algebra *homomorphism* by re-defining the Jacobi-Lie bracket. This is accomplished in [35] by reversing the sign of the Jacobi-Lie bracket (3.1) — giving the so-called negative Jacobi-Lie bracket — and defining the Lie algebra bracket using *right*- instead of left-invariant vector fields on the Lie group. Although, as noted in [34], this change of sign is not strictly necessary it does make the following discussion slightly more convenient. For the remainder of this section only we assume that the Jacobi-Lie bracket is the negative of (3.1) and that Φ^λ is a Lie algebra homomorphism.

Definition 3.1.7: Let $\Phi^\lambda : \mathfrak{g} \rightarrow \mathcal{X}(M)$ be a Lie algebra homomorphism, where M is a smooth manifold and \mathfrak{g} is the Lie algebra of the Lie group \mathbf{G} , and $A : \mathbb{R} \rightarrow \mathfrak{g}$. A differential equation for a curve $Y : \mathbb{R} \rightarrow \mathbf{G}$ of the form

$$\dot{Y}(t) = \Phi_{A(t)}^\lambda(Y(t)) \quad (3.5)$$

is called an *equation of Lie type*.

It is easy to see that the covariance differential equation is an equation of Lie type. With the drift and control matrices A and B fixed, the time-varying control gain $K : [0, T] \rightarrow \mathbb{R}^{m \times n}$ makes $(A + BK) : [0, T] \rightarrow \mathfrak{g}$ a curve in the Lie algebra and $S : \mathbb{R} \rightarrow \mathbf{PD}_n$ the evolution of the covariance. Then $\dot{S}(t) = \Phi_{A+BK}^\ominus(S(t)) = (A + BK)S(t) + S(t)(A + BK)^T$.

The solutions of equations of Lie type are determined by fundamental solutions. The following is Proposition 3 in [34].

Proposition 3.1.1 (Bryant): Let $A : \mathbb{R} \rightarrow \mathfrak{g}$ be a curve in the Lie algebra of the Lie group \mathbf{G} , and let $\lambda : \mathbf{G} \times M \rightarrow M$ be a left action on M . If $X : \mathbb{R} \rightarrow \mathbf{G}$ is the solution to $\dot{X} = A(t)X(t)$ with initial condition $X(0) = I$, then it is called a *right-invariant fundamental solution* and gives the solutions to equation (3.5) as $Y(t) = \lambda(X(t), Y_0)$ where $Y_0 = Y(0)$ is the initial condition of the curve $Y(t) \in M$.

This proposition rigorously generalizes what we already knew to be true from Proposition 2.2.1, that $S(t) = \Theta(X(t), S_0) = X(t) S_0 X(t)^T$. We will return, briefly, to equations of Lie type at the end of chapter 4.

3.2 Invariant Optimal Control Problems

The remainder of this chapter is concerned with solving the following right-invariant optimal control problem on GL_n :

$$\begin{aligned} \min_{K(t) \in \mathcal{U}, T \in [0, \infty)} \quad & \int_0^T \frac{1}{2} \langle K, K \rangle dt, \\ \text{s. t.} \quad & \dot{X} = (A + BK)X, \quad X(T) S_0 X(T)^T = S_f. \end{aligned} \tag{3.6}$$

where the final time T is free, S_f is the output of the semidefinite program (2.9), S_0 is given and $\mathcal{U} \subset \mathbb{R}^{m \times n}$ is a closed, convex subset.

The optimal solution trajectory $X(t)$ for (3.6) minimizes the matrix Frobenius norm of the control gain along fundamental solutions of the closed-loop linear system (2.4), while ensuring that the system reaches the covariance S_f in the unspecified time $T < \infty$. This problem belongs to a class of so-called invariant optimal control problems, in which the Hamiltonian can be represented by its value at the identity in \mathbf{G} ; this is accomplished by the identification of the cotangent bundle of the Lie group as $T^*\mathbf{G} = \mathfrak{g}^* \times \mathbf{G}$. This identification — called *trivialization of the cotangent bundle* [36] or simply *reduction* in the context of geometric mechanics [27] — has the potential to reduce the complexity of numerical algorithms for solving invariant optimal control problems. In [37] trivialization yielded closed-form solutions to optimal control problems on the special orthogonal group SO_3 and special Euclidean group SE_3 . Our problem is formulated on GL_n and so does not readily lend itself to generic closed-form solutions.

The Hamiltonian system on $\mathfrak{g}^* \times \mathbf{G}$ corresponding to an *invariant* optimal control problem always has the following properties:

(T1) The costate evolves on \mathfrak{g}^* independently of the state.

(T2) The costate differential equation is a *Lax equation* with isospectral flow.

These properties are highly desirable from a numerical perspective, since property (T2) introduces n additional invariants that must be preserved by the numerical integration and (T1) implies that the costate equation can be integrated independently of the state, or even possibly offline.

In this section I show that problem (3.6) benefits from these properties, but trivialization has the additional unforeseen benefit of reducing the 2-point boundary value problem to the determination of the initial or final value of the costate in \mathfrak{g}^* ; the transversality condition for the Hamiltonian system on the trivialized cotangent bundle does not depend on the final value of the state, so the problem reduces to an initial or final value problem on the costate.

Trivialization of the cotangent bundle: The identification $T^*\mathbf{G} = \mathfrak{g}^* \times \mathbf{G}$ can be made using either left or right translation, depending on whether the optimal control problem is left- or right-invariant. Optimal control problem (3.6) is right-invariant so we will need the correspondence due to right translation.

Definition 3.2.1: A *trivialization of the cotangent bundle* is a diffeomorphism $\Psi : E \times \mathbf{G} \rightarrow T^*\mathbf{G}$, where E is a vector space of the same dimension as \mathbf{G} , such that

1. for all $P \in E$ and $X \in \mathbf{G}$, $\pi \circ \Psi(P, X) = x$, where $\pi : T^*\mathbf{G} \rightarrow \mathbf{G}$ is the canonical projection;
2. the map $P \mapsto \Psi(P, X)$ is a linear isomorphism $E \rightarrow T_X^*\mathbf{G}$.

The trivialization identifies each fiber $T_X^*\mathbf{G}$ with the space E .

Let $T_Y^*L_X : T_{XY}^*\mathbf{G} \rightarrow T_Y^*\mathbf{G}$ and $T_Y^*R_X : T_{YX}^*\mathbf{G} \rightarrow T_Y^*\mathbf{G}$ denote the *cotangent lifts* of L_X and R_X at $Y \in \mathbf{G}$. The cotangent lift $T_Y^*R_X$ is defined by $\langle T_Y^*R_X P, Z \rangle = \langle P, T_Y R_X Z \rangle$ for all $Z \in T_Y\mathbf{G}$ and $P \in T_{YX}^*\mathbf{G}$. The cotangent lift $T_Y^*L_X$ is defined similarly.

Note that for all $X \in \mathbf{G}$, $T_X^*L_{X^{-1}}$ and $T_X^*R_{X^{-1}}$ are maps from $T_I^*\mathbf{G} = \mathfrak{g}^*$ to $T_X^*\mathbf{G}$. Following [38], we obtain a trivialization of the cotangent bundle using either the correspondence $(P, X) \leftrightarrow T_X^*L_{X^{-1}} P$ or $(P, X) \leftrightarrow T_X^*R_{X^{-1}} P$, where $(P, X) \in \mathfrak{g}^* \times \mathbf{G}$.

Define the *right trivialization* $\Psi : \mathfrak{g}^* \times \mathbf{G} \rightarrow T^*\mathbf{G}$ by the map $(P, X) \mapsto \bar{P}_X$, where $\bar{P}_X := T_X^* R_{X^{-1}} P$ is the right-invariant 1-form defined by its value at the identity,

$$\langle \bar{P}_X, AX \rangle = \langle P, T_X R_{X^{-1}} AX \rangle = \langle P, A \rangle,$$

for all $A \in \mathfrak{g}$. Following [36], for any $X \in \mathbf{G}$ consider the map $\Psi(\cdot, X) : \mathfrak{g}^* \rightarrow T_X^* \mathbf{G}$ and its dual, $\Psi^*(\cdot, X) : T_X \mathbf{G} \rightarrow \mathfrak{g}$. These are linear isomorphisms of the cotangent and tangent spaces, respectively. For the right trivialization we find that, for $P \in \mathfrak{g}^*$ and $Z \in \mathfrak{g}$,

$$\begin{aligned} \langle \Psi(\cdot, X), Z \rangle &= \langle \cdot, \Psi^*(Z, X) \rangle \\ &= \langle \cdot, T_X R_{X^{-1}} Z \rangle, \end{aligned}$$

that is, $\Psi^*(\cdot, X) = T_X R_{X^{-1}}(\cdot)$. It is seen that the inverse map $\Psi^{-1} : T^*\mathbf{G} \rightarrow \mathfrak{g}^* \times \mathbf{G}$ is given by $(\bar{P}_X) \mapsto (T_X^* R_X \bar{P}_X, X)$, since $P = T_X^* R_X(T_X^* R_{X^{-1}} P)$.

We want to derive the differential equations for Hamiltonian systems on $\mathfrak{g}^* \times \mathbf{G}$. These expressions are obtained in [36] as equations (18.14) for the *left trivialization* of the cotangent bundle, $(P, X) \mapsto \bar{P}_X = T_X^* L_{X^{-1}} P$. The derivation of the Hamiltonian differential equations for right trivialization are straightforward; the following is a corollary to the results in [36, Sec 18.3]. I only provide a sketch of the proof. We briefly recall some basic facts on differential forms.

A differential k -form ω on a smooth manifold M is a smooth assignment of a skew-symmetric, k -multilinear map to every point $x \in M$, that is, $\omega_x : T_x M \times \cdots \times T_x M \rightarrow \mathbb{R}$. The *interior product* of a k -form ω is the $k - 1$ form obtained by contraction with a vector field. It is the map $\omega \mapsto \omega(v, \cdot)$ denoted $\iota_v \omega$, for $v \in \mathcal{X}(M)$. The *differential* of a smooth 1-form ω on M is given by the formula

$$d\omega(X, Y) = X\langle \omega, Y \rangle - Y\langle \omega, X \rangle - \langle \omega, [X, Y] \rangle, \quad (3.7)$$

where $X, Y \in \mathcal{X}(M)$.

Corollary 3.2.1: The Hamiltonian system on $\mathfrak{g}^* \times \mathbf{G}$ corresponding to the smooth Hamiltonian function $H : \mathfrak{g}^* \times \mathbf{G} \rightarrow \mathbb{R}$ is given by the differential equations

$$\begin{aligned}\dot{X} &= T_X R_X \left(\frac{\partial H}{\partial P} \right) \\ \dot{P} &= -T_X^* R_X \left(\frac{\partial H}{\partial X} \right) - \left(\text{ad} \frac{\partial H}{\partial P} \right)^* (P)\end{aligned}\tag{3.8}$$

Proof (Sketch): For a given trivialization $\Psi : E \times \mathbf{G} \rightarrow T^*\mathbf{G}$, the system of differential equations for a Hamiltonian $H : E \times \mathbf{G} \rightarrow \mathbb{R}$ are provided as equations (18.13) in [36, Sec 18.3]:

$$\dot{X} = (\Psi^*)^{-1} \frac{\partial H}{\partial P}(P, X), \quad \dot{P} = -\Psi^{-1} \left(\frac{\partial H}{\partial X}(P, X) + \iota_{\dot{X}} d\Psi_{(P, X)} \right),$$

where the Ψ^{-1} is a slight abuse of notation, and for right trivialization evaluates to $\Psi^{-1} = T_X^* R_X$. From the discussion above we also note that $(\Psi^*)^{-1} = T_X R_X$ and $d\Psi_{(P, X)} = d\bar{P}_X$. To compute $d\bar{P}_X$ we use equation (3.7). Recall that $[AX, BX] = -[A, B]X$, for all $X \in \mathbf{G}$, $A, B \in \mathfrak{g}$. Then the action of $d\bar{P}_X$ on right-invariant vector fields is

$$\begin{aligned}d\bar{P}_X(AX, BX) &= (AX)\langle P, B \rangle - (BX)\langle P, A \rangle - \langle P, [AX, BX]X^{-1} \rangle \\ &= \langle P, [A, B] \rangle,\end{aligned}$$

where $\langle P, B \rangle$ and $\langle P, A \rangle$ are just constants. This yields

$$\begin{aligned}\iota_{\dot{X}} d\bar{P}_X &= d\bar{P}_X \left(T_X R_X \left(\frac{\partial H}{\partial P} \right), \cdot \right) = \left\langle P, \left[\frac{\partial H}{\partial P}, \cdot \right] \right\rangle \\ &= \left\langle \left(\text{ad} \frac{\partial H}{\partial P} \right)^* (P), \cdot \right\rangle,\end{aligned}$$

from which we obtain $\Psi^{-1} \iota_{\dot{X}} d\bar{P}_X = \left(\text{ad} \frac{\partial H}{\partial P} \right)^* (P)$ and the result follows. \square

Remark 3.2.2: In the engineering literature, the differential equation on \mathbf{G} in Corollary 3.2.1 is called the *state* equation, and the equation on $T^*\mathbf{G}$ (or in this case, on $\mathfrak{g}^* \times \mathbf{G}$) the *costate* equation.

Note that the infinitesimal generator for the coadjoint action appears in the expression for the costate. This means that the costate is generated by the coadjoint action when $\frac{\partial H}{\partial X} \equiv 0$. In [38] this very fact is obtained by instead considering the negative Jacobi-Lie bracket and defining the Lie algebra bracket on \mathfrak{g} using right-invariant vector fields. Arguably, this is because the optimal control problems of interest in that reference are all left-invariant.

The left trivialization with our definition of the (positive) Jacobi-Lie bracket and left-invariant Lie algebra bracket is, of course, covered in [36].

Transversality conditions: In preparation for our discussion of the necessary conditions for optimality of solutions to (3.6), we study the constraint set $\{X \in \mathbf{GL}_n : X S_0 X^T = S_f\}$, and determine its tangent and normal spaces.

Consider the terminal constraint function $G : \mathbf{PD}_n \times \mathbf{GL}_n \rightarrow \mathbf{Sym}_n$ defined by $G_{S_0}(X) := G(S_0, X) = \Theta_X(S_0) - S_f$, where $\mathbf{Sym}_n = \mathbb{R}^{n(n+1)/2}$ is the vector space of symmetric matrices. By the implicit function theorem, if G_{S_0} has constant rank on \mathbf{GL}_n then the level set $\mathcal{G} = G_{S_0}^{-1}(0)$ is a smooth embedded submanifold of \mathbf{GL}_n with tangent space given by

$$T_X \mathcal{G} = \{Z \in T_X \mathbf{GL}_n : DG(X) \cdot Z = 0\}.$$

The derivative of G_{S_0} is

$$DG_{S_0}(X) \cdot Z = \left. \frac{d}{dt} \right|_{t=0} (X + tZ) S_0 (X + tZ)^T = Z S_0 X^T + X S_0 Z^T,$$

and has full rank for all $X \in \mathbf{GL}_n$. It suffices to show that for every $X \in \mathbf{GL}_n$ and $\hat{Z} \in \mathbf{Sym}_n$ there exists a $Z \in T_X \mathbf{GL}_n$ such that $DG(X) \cdot Z = \hat{Z}$; see [39, 40] for

similar computations in other matrix Lie groups. Note that because $\hat{Z} \in \text{Sym}_n$ there exists $W \in \mathbb{R}^{n \times n}$ such that $\hat{Z} = W + W^T$. Then, by setting $Z = WX^{-T}S_0^{-1}$ we obtain $DG_{S_0}(X) \cdot Z = \hat{Z}$. Therefore, G_{S_0} has constant rank and \mathcal{G} is an embedded submanifold of dimension $n(n-1)/2$.

We now derive the conditions for a point $P \in T^*\text{GL}_n$ to be transverse to the tangent space $T_X\mathcal{G}$. Define the *normal space* to be $(T_X\mathcal{G})^\perp := \{P \in \mathbb{R}^{n \times n} : \langle P, Z \rangle = 0 \text{ for all } Z \in T_X\mathcal{G}\}$. This is a vector space of dimension $n(n+1)/2$ because $\dim(\mathcal{G}) = n(n-1)/2$.

Lemma 3.2.2: $(T_X\mathcal{G})^\perp = \{P = WX S_0 \in \mathbb{R}^{n \times n} : W = W^T \in \text{Sym}_n\}$.

Proof. Let $Z \in T_Z\mathcal{G}$. Then $Q^T(ZS_0X^T + XS_0Z^T) = 0$ for all $Q \in \mathbb{R}^{n \times n}$ so certainly $\langle Q, (ZS_0X^T + XS_0Z^T) \rangle = 0$. Simplifying,

$$\begin{aligned} 0 &= \langle Q, (ZS_0X^T + XS_0Z^T) \rangle \\ &= \langle QXS_0, Z \rangle + \langle Z, Q^TXS_0 \rangle \\ &= \langle (Q + Q^T)XS_0, Z \rangle \end{aligned}$$

This holds for all $Q \in \mathbb{R}^{n \times n}$, $Z \in T_X\mathcal{G}$ and $X \in \text{GL}_n$ and therefore $P \in (T_X\mathcal{G})^\perp \Leftrightarrow P = WX S_0$, for some $W = Q + Q^T \in \text{Sym}_n$. \square

Pontryagin's Maximum Principle: Summing up the previous results of this chapter and applying Pontryagin's Maximum Principle [36, Theorem 12.2, Theorem 12.4], we obtain the following characterization of optimal solutions to problem (3.6).

Proposition 3.2.3: The optimal trajectories of the right-invariant, free-time optimal control problem (3.6) are given by the following Hamiltonian system of differential equations on $\mathfrak{gl}_n^* \times \text{GL}_n$:

$$\begin{aligned} \dot{P} &= [P, (A + BK)^T] \\ \dot{X} &= T_X R_X (A + BK) = (A + BK)X, \end{aligned} \tag{3.9}$$

where the optimal controls are $K = B^T P$, the maximized Hamiltonian is

$$H(P, X) = \langle P, A + \frac{1}{2} B B^T P \rangle \equiv 0,$$

and the final costate satisfies $P(T) = W S_f$, where $W \in \text{Sym}_n$ and S_f is the terminal covariance.

Proof. The pre-Hamiltonian of the optimal control problem is a function $h : \mathfrak{gl}_n^* \times \text{GL}_n \times \mathbb{R}^{m \times n}$,

$$\begin{aligned} h(P, X, K) &= \langle T_X^* R_{X^{-1}} P, (A + BK)X \rangle - \frac{1}{2} \langle K, K \rangle \\ &= \langle P, A + BK \rangle - \frac{1}{2} \langle K, K \rangle \end{aligned}$$

The optimal control gain K maximizes $h(P, X, K)$, so that $0 = \partial h / \partial K = B^T P - K$. The maximized Hamiltonian $H : \mathfrak{gl}_n^* \times \text{GL}_n \rightarrow \mathbb{R}$ is therefore $H(P, X) = \max_{K(t) \in \mathcal{U}} h(P, X, K) = \langle P, A + \frac{1}{2} B B^T P \rangle$, and Theorem 12.2 of [36] establishes that the Hamiltonian equals 0 along optimal solutions of a free-time problem.

Note that $\partial H / \partial P = A + B B^T P$, $\partial H / \partial X = 0$, and recall that the infinitesimal generator of the coadjoint action is $(\text{ad } A)^*(P) = [A^T, P]$. The differential equations from Corollary 3.2.1 become

$$\begin{aligned} \dot{X} &= T_X R_X (A + B B^T P) = (A + B B^T P) X \\ \dot{P} &= -[(A + B B^T P)^T, P]. \end{aligned}$$

Finally, Theorem 12.4 of [36] establishes that a terminal constraint $X(T) \in \mathcal{G}$ implies a necessary transversality condition on the terminal costate $P \in T^* \text{GL}_n$: $P(T) \perp T_X \mathcal{G}$. This transversality condition can be translated to the identity to

obtain the equivalent condition on $P \in \mathfrak{g}^*$: $P(T) \perp T_X \mathcal{G} X^{-1}$. Applying Lemma 3.2.2, we find that at $t = T$,

$$0 = \langle P, T_X \mathcal{G} X^{-1} \rangle = \langle P X^{-T}, T_X \mathcal{G} \rangle,$$

which implies that $P(T)X(T)^{-T} = WX(T)S_0$ for some $W \in \text{Sym}_n$. If $X \in \mathcal{G}$ then $X(T)S_0X(T)^T = S_f$ and therefore $P(T) = WS_f$. \square

Remark 3.2.3: Properties (T2) and (T1) are apparent from the differential equations derived. The transversality condition reduces the number of unknowns in the boundary value problem only for the Hamiltonian system on $\mathfrak{gl}_n^* \times \text{GL}_n$. Otherwise there are $2n^2$ scalar unknowns corresponding to $X(T)$ and $P(T)$ or $X(T)$ and $P(0)$. It is not expected that other right- or left-invariant optimal control problems are always reducible in this manner to initial value problems.

As mentioned earlier, the same differential equations on \mathfrak{gl}_n^* are derived in [7] without formally trivializing $T^*\text{GL}_n$, and the transversality condition has the same effect of reducing the number of unknowns in the boundary value problem in that context, as well.

3.2.1 Application to a Planar Linear System

In this section I use a so-called indirect or shooting method to find the optimal solutions according to Proposition 3.2.3. Numerical experimentation shows that correctly specifying the transversality conditions for problem (3.6) is essential; omitting them can cause the solver to fail or converge to a local maximum of the Hamiltonian. We discuss some of these issues in this section, where we simulate a simple planar system.

Let us apply the results of the preceding section to the planar system (2.10) considered in section 2.3:

$$A = \begin{pmatrix} -1.417 & 1 \\ 2.86 & -1.183 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ -3.157 \end{pmatrix},$$

We construct our terminal constraint from (2.9),

$$S_f = \begin{pmatrix} 0.033246 & 0.047109 \\ 0.047109 & 0.16675 \end{pmatrix}.$$

Numerical solution: The numerical integration of the fundamental solutions of this simple planar system presents several challenges that are only partially surmounted by the standard *direct* and *indirect* methods for solving boundary value problems that one may find in [41], for example.

In the engineering literature there are two different numerical methods that one typically considers for solving an optimal control problem. In so-called *indirect methods* one guesses values for optimal solutions at the initial or final times and “shoots” the system forward or backward in time. The resulting trajectory is checked for optimality by an error function, which allows a root-finding algorithm, such as Matlab’s `fsolve`, to be used to refine the initial guess. Indirect methods will not converge to the optimal solution unless the initial guess is sufficiently close to the optimal.

For our optimal control problem (3.6), an indirect shooting method can either guess a symmetric matrix W and integrate the costate equation (3.9) backwards from $P(T) = WS_f$, or guess an initial matrix $P(0)$ and then verify that $P(T)S_f^{-1}$ is a symmetric matrix. Therefore, a shooting method that integrates backwards in time has $n(n+1)/2$ unknowns for W in the shooting method, while a forward-time integration has n^2 unknowns for $P(0)$. Numerical experimentation shows that backward-time integration in `ode113` is unstable in the shooting method, so it is

preferable to use a forward-time integration even though the number of unknowns is greater.

So-called *direct methods* approximate the trajectories of a dynamical system using piecewise polynomial splines. Such methods are often called collocation methods, because the splines are chosen to exactly satisfy the dynamics at specific points in time. Although it is preferable to construct splines that automatically lie in GL_n , it is possible to use direct methods on \mathbb{R}^{n^2} , just as in indirect methods. The issue is, however, that in collocation-based techniques, the interpolating spline is constructed to satisfy the dynamics exactly at specific points, and the number and placement of these points must be chosen carefully in order for the direct method to succeed. It is perhaps not clear how to choose the placement of these interpolating points in a matrix manifold; moreover, the direct method introduces n^2 additional unknowns for every set of spline parameters and interpolating points, so the dimensionality of a direct method can grow very quickly. It is expected that the extension of recent work on generating smooth splines on compact Lie groups [42] to the non-compact case should assist in the formulation of a computationally efficient direct method for this problem.

Because an indirect shooting method has fewer unknowns than a direct method, I solve the optimal control problem with a standard numerical integrator designed for problems on \mathbb{R}^{n^2} in a shooting method using Matlab's `ode113`, an adaptive step-size integrator designed to handle problems with stringent error tolerances. The adoption of `ode113` is due mainly to the lack of availability of Lie group integrators. The only publicly available package is called `DiffMan`¹ [43], but it only provides fixed step-size integrators. As noted in [44, pg. 129], and references therein, it is not possible to preserve both the Hamiltonian and the isospectrality of the costate equations using a fixed step-size numerical integrator; these are the very quantities that one must

¹During the writing of this thesis `DiffMan` was updated from version 2.0.0 to 2.0.1 for the first time in 12 years, on May 1, 2012.

use to confirm that the numerical solution of an invariant optimal control problem was accurate.

Using Matlab's root-finding package `fsolve` the optimal solution is found fairly quickly in a few iterations. `ode113` is set with relative error tolerance $1e-9$ and absolute error tolerance $1e-10$. With initial covariance $S_0 = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ the optimal final time is found to be $T = 1.3493$. The initial and final costates are

$$P(0) = \begin{pmatrix} 0.044591 & -0.15823 \\ -0.15823 & -0.23386 \end{pmatrix}, \quad P(T) = \begin{pmatrix} -0.34647 & -0.53583 \\ 0.035464 & 0.1572 \end{pmatrix} = W S_f$$

$$W = \begin{pmatrix} -9.7856 & -0.44878 \\ -0.44878 & 1.0695 \end{pmatrix}, \quad S_f = \begin{pmatrix} 0.033246 & 0.047109 \\ 0.047109 & 0.16675 \end{pmatrix}.$$

The optimal cost, maximized Hamiltonian, and costate eigenvalues are shown below in figures 3.1, 3.2 and 3.3. Note that the Hamiltonian is essentially zero and the costate eigenvalues are constant. This implies that our numerical solution is not inaccurate and we have found, at the very least, a local maximum of the Hamiltonian. It is always a concern that the obtained numerical solution corresponds not to a global maximum of the Hamiltonian, but to a local maximum. This can happen, for example, when a terminal constraint is enforced in the boundary value problem solver without its corresponding transversality condition. Verifying that a numerical solution is in fact a global maximum is difficult. Confirming the necessary condition that the Hamiltonian and costate eigenvalues are constant is much more straightforward.

The covariance and controls are shown in figures 3.4 and 3.5. Note that the controls do not settle at a constant value. In the next chapter I implement the optimal control problem presented here in receding-horizon fashion in order to stabilize the system near a desired covariance.

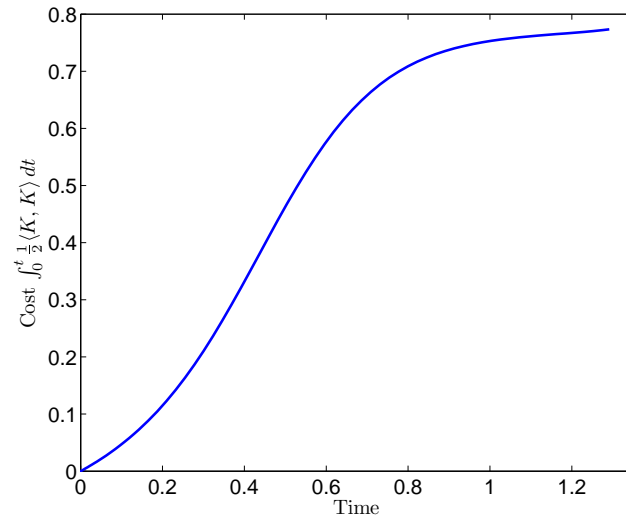


Fig. 3.1. The running cost along optimal trajectories. The optimal cost at the final time is $\int_0^{1.3493} \frac{1}{2} \langle K, K \rangle dt = 0.759$.

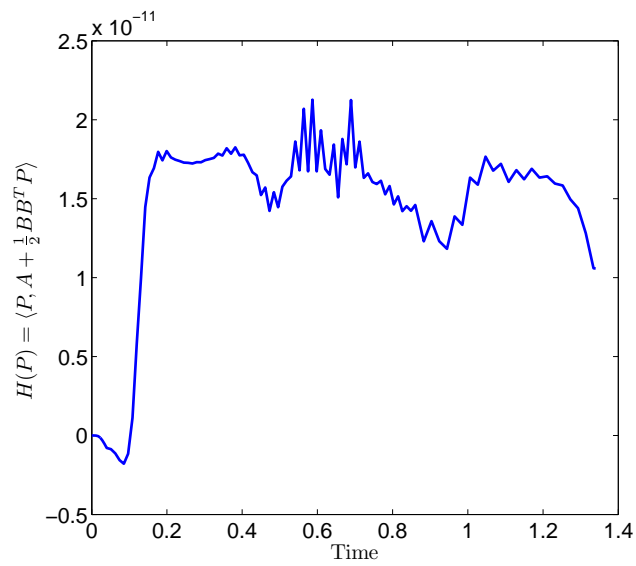


Fig. 3.2. The Hamiltonian of a free-time problem should be equal to 0. The error in the $H(P)$ is on the order of $1e-11$, an order of magnitude *smaller* than the integration tolerance.

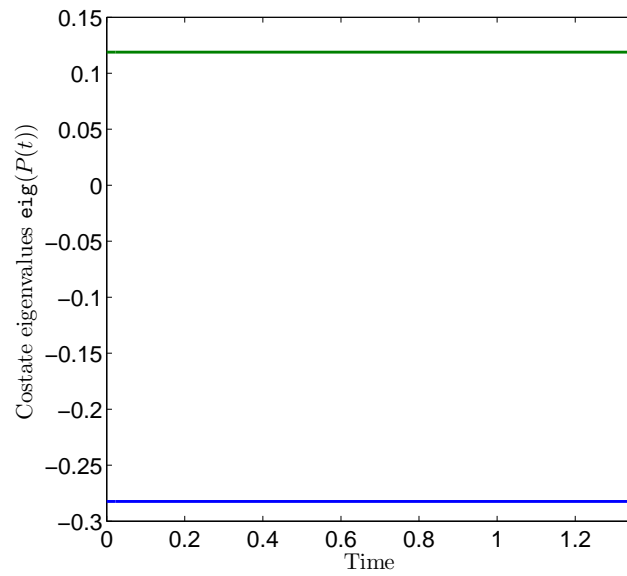


Fig. 3.3. The eigenvalues of the costate in \mathfrak{gl}_n^* over time. It is apparent that the trajectories are isospectral up to plotting accuracy.

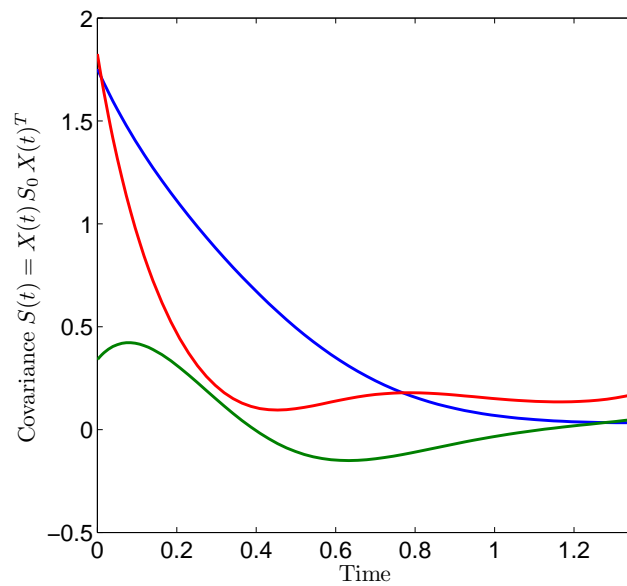


Fig. 3.4. The evolution of the covariance associated with the system. The desired final covariance is reached.

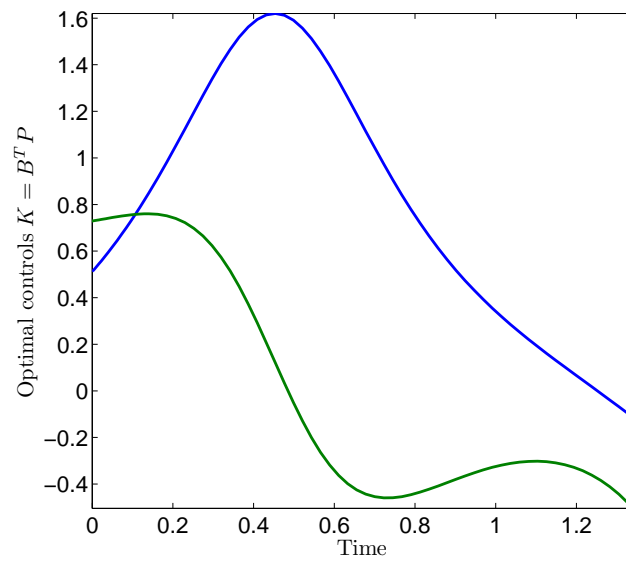


Fig. 3.5. The time-varying entries of the matrix $K(t) \in \mathbb{R}^{1 \times 2}$.

4. RECEDING HORIZON COVARIANCE CONTROL

The philosophy and motivation underlying receding horizon control methods is the desire to convert open-loop control laws into feedback control laws.

In the previous chapter we applied Pontryagin's Maximum Principle to obtain the time-varying controls that minimize a cost functional over a finite interval of time. Generally speaking, the solution of finite-horizon optimal control problems always yields time-dependent controls via solution of a two-point boundary value problem. It is by far preferable to instead obtain optimal controls as functions of the system state. Such feedback controls obtain from infinite-horizon control problems. The prototypical example of such problems is the Linear Quadratic Regulator problem [12], where the solution of a matrix Riccati equation yields optimal feedback controls provided the linear system is stabilizable. This prototypical example is, however, far from typical. In general, the length of the horizon effectively prevents the use of direct and indirect numerical methods and one solves the infinite horizon problem using the Hamilton-Jacobi-Bellman equation, a PDE which has the Hamiltonian differential equations as its characteristic equations [36, Ch. 17.2]. Depending on the optimal control problem, the PDE can be significantly more difficult to solve [45].

Receding horizon control represents a pragmatic — and successful — compromise to the dilemma presented by the fact that finite-horizon optimal control problems are easier to solve than infinite-horizon optimal control problems, by repeatedly solving a sequence of finite-horizon optimal control problems on the interval $[t, t + T]$ at discrete time intervals of length δt ; the corresponding receding horizon control law takes the solution at time t of the optimal control problem over the interval $[t, t + T]$ and applies it as a constant control over the interval $[t, t + \delta t]$. At time $t + \delta t$ another optimization problem is solved over $[t + \delta t, t + \delta t + T]$ and its optimal control at time $t + \delta t$ is fixed over the interval $[t + \delta t, t + 2\delta t]$. In this manner we obtain a receding

horizon control law that is inherently discrete because it samples the system state every δt seconds.

In the limit as $\delta t \rightarrow 0$ we obtain the *idealized receding horizon controller*, where a finite-horizon optimal control problem is solved at *every* instant of time. In this ideal case the receding horizon controller is said to be *inverse optimal* [45,46] because it minimizes a modified cost function over the interval $[0, \infty)$. It is due to this inverse optimal property of idealized receding horizon control strategies that we may claim that general receding horizon controllers approximate closed-loop feedback control laws. Note that this limiting case is very similar to Bellman's method of dynamic programming, as discussed in [36, Sec 17.3], the exception being that the Hamilton-Jacobi-Bellman equation resulting from inverse optimality is not exactly the Hamilton-Jacobi-Bellman equation for the optimal control problem solved at every instant in time.

Regardless, implementation of the idealized receding horizon controller is essentially equivalent to solving a Hamilton-Jacobi-Bellman equation. It is also impossible to numerically solve a boundary value problem at every instant of time. The challenge presented by receding horizon control, therefore, is to find the largest *update interval* δt and the smallest *horizon length* T such that the receding horizon control strategy stabilizes the control system with desirable steady-state properties. The update interval δt can be taken as an upper bound on the amount of time permitted for numerically solving the boundary value problem for each optimal control on the interval $[t, t + T]$. The horizon length T is regarded as a truncation of the infinite horizon problem. Intuitively speaking, an increase in the horizon length improves system stability at the cost of increasing the time needed to numerically solve the boundary value problem.

The situation is further complicated when one considers implementing receding horizon control in real-time. In this setting hardware limitations impose a hard lower upper bound on the update interval δt , which in turn implies a lower bound on the

horizon length T . Recently it was shown that there always exist a sufficiently large horizon T such that a receding horizon control strategy is stabilizing [47], but in general the horizon length must be chosen according to some heuristic process.

The remainder of this thesis is concerned with a receding horizon implementation of the results of the prior chapter, yielding, to my knowledge, the first implementation of a receding horizon control strategy to a control system on a matrix Lie group. In the first section I design a heuristic control strategy modeled after the dual-mode strategy of [48]. The first mode of the receding horizon controller solves a sequence of free-time optimal control problems with a terminal constraint on the terminal covariance. When the optimal time to reach the desired covariance falls below the update time of the receding horizon controller, the second mode triggers and Skelton's controller assigns the system to a covariance close to the desired covariance. I conclude the chapter with a discussion of the possible implications of solving families of boundary value problems for equations of Lie type on homogeneous spaces. As shown in chapter 3, geometric methods significantly reduced the complexity of the resulting boundary value problem; it is expected that Lie's method of reduction will reduce the complexity of receding horizon implementations.

4.1 Dual-Mode Receding Horizon Control Strategy

Our control system of interest is given by the differential equation $\dot{X} = (A + BK)X$, and our receding horizon control strategy will be formed by solving optimal control problem (3.6) repeatedly, but over different time intervals and with different initial covariances.

With the terminal covariance fixed at a pre-determined value S_f , the optimal control gain K and optimal final time T may be considered parameterized by an

initial covariance S . To emphasize the dependence of the optimal controls on the initial covariance, adopt the following notation for solutions of (3.6):

$$\begin{aligned} (K(\cdot, S), T(S)) &= \arg \min_{K, T} \int_0^T \frac{1}{2} \langle K, K \rangle dt, \\ \text{s. t. } X(T) S X(T)^T &= S_f, \\ \dot{X} &= (A + BK)X, \end{aligned} \tag{4.1}$$

where $T : \text{PD}_n \rightarrow [0, \infty)$ is the optimal time it takes the covariance to reach S_f from $S \in \text{PD}_n$ and $K : [0, T(S)] \times \text{PD}_n \rightarrow \mathbb{R}^{m \times n}$ is the control achieving that transfer.

Remark 4.1.1: In order to justify the claim that the optimal control and optimal terminal time are functions on $S \in \text{PD}_n$ it is necessary to show that optimal controls exist for every $S \in \text{PD}_n$. The notation in (4.1) is only for convenient exposition. If there exists one optimal trajectory for some particular initial covariance S , then there exists a sufficiently small neighborhood of that trajectory $X(t) S X(t)^T$ on which we may define $K(t, \cdot)$ and T as local functions. The question of existence and uniqueness of optimal controls is outside the scope of this thesis, but some discussion may be found in [36]. Questions of controllability are discussed in [7].

Using notation (4.1), we now implement our optimal control problem of interest in a receding-horizon fashion, as follows.

Definition 4.1.2: An *idealized receding horizon control strategy* for (3.6), denoted $\mathcal{RH}(0)$, is a state-feedback control law obtained from the repeated solution of (4.1) at every instant of time.

Let S_0 be the initial covariance given at the start of the receding horizon strategy. The state-feedback control is the map $K_{\mathcal{RH}(0)} : \text{PD}_n \rightarrow \mathbb{R}^{m \times n}$ defined by

$$K_{\mathcal{RH}(0)}(S) = K(0, S),$$

where $K(0, S) = B^T P(0)$ is the initial value of the optimal controls taking S to S_f in $T(S)$ units of time. The closed-loop dynamics are given by $\dot{X} = (A + BK(0, X S_0 X^T))X$.

First of all, the idealized receding horizon feedback control $K_{\mathcal{RH}(0)}$ requires solution of a boundary value problem at every instant of time; unless a feedback control law is already at hand — in which case a receding horizon implementation is unnecessary — the idealized receding horizon controller is impossible to implement. It is therefore necessary to introduce a delay greater than the computational time necessary to solve the boundary value problem when implementing (3.6) in a receding-horizon fashion in real-time.

Secondly, the limiting behavior of the idealized controller may not be stabilizing. The terminal constraint $X(T) S X(T)^T = S_f$ in problem (3.6) ensures that the receding horizon controller brings the system covariance arbitrarily close to S_f from the initial covariance S . When $S = S_f$, however, the optimal time is $T = 0$ and so the optimal controls cannot be found by applying Pontryagin's Maximum Principle to (3.6). This issue implies that a different control problem must be formulated in order to stabilize the system at S_f .

Both of these concerns motivate the following dual-mode receding horizon control strategy, modeled after the results of [48].

Definition 4.1.3: The *dual-mode receding horizon control strategy* for (3.6), denoted $\mathcal{RH}(\delta t)$, is a state-feedback control law obtained from the repeated solution of (4.1) every δt seconds.

Let S be the system covariance at time t_i . The optimal control $K(0, S) \in \mathbb{R}^{m \times n}$ is applied as a *constant* control over the interval $[t_i, t_i + \delta t]$ if the optimal time $T(S)$ to reach S_f from S is greater than δt . If $T(S)$ is less than the update time, the constant feedback gain (2.8),

$$K(S_f) := -B^+(AS_f + S_f A^T)(I - \frac{1}{2}BB^+)S_f^{-1} - B^+QS_f^{-1}$$

solving $0 = (A + BK(S_f))S_f + S_f(A + BK(S_f))$, is applied over the interval instead.

This dual-mode logic results in a state- and time-dependent, piecewise-constant feedback control $K_{\mathcal{RH}} : \text{PD}_n \rightarrow \mathbb{R}^{m \times n}$ defined by

$$K_{\mathcal{RH}}(S) = \begin{cases} K(0, S), & \text{if } T(S) > \delta t \\ K(S_f), & \text{if } T(S) \leq \delta t. \end{cases}$$

This control is fixed and constant over each interval $[t_i, t_i + \delta t]$. The closed-loop dynamics are given by $\dot{X} = (A + BK_{\mathcal{RH}})X$.

4.1.1 Application to a Planar Linear System

We illustrate the dual-mode strategy on our running example. The update time has a significant effect on the steady-state behavior of the system.

The figures below show the control effort and the evolution of the covariance for the dual-mode controller $\mathcal{RH}(\delta t)$ with $\delta t = 0.05$ seconds (20Hz). As we can see in Figure 4.1, the first-mode of the receding horizon controller terminates after $t = 1.3$ seconds, with the second mode kicking in at $t = 1.35$ seconds and applying the control $K(S_f) = (2.4947, -0.82357)$ computed in chapter 2. The dashed lines show the desired covariance S_f , and it is clear that the steady-state behavior of the covariance — which is stable due to the fact that $K(S_f)$ solves the steady-state covariance equation $0 = (A + BK(S_f))S_f + S_f(A + BK(S_f))$ — is to oscillate about the desired value S_f .

It is expected that with a smaller δt the first mode of the controller will terminate at a covariance that is closer to the desired, and the oscillations will be smaller. However, the update time cannot be made arbitrarily small, due to the practical concerns mentioned above and also due to the fact that an indirect or direct numerical method cannot solve an optimization problem of arbitrarily small duration. Numer-

ical experimentation indicates that our indirect shooting method cannot solve the boundary value problem when the optimal time is smaller than $\delta t = 0.05$ seconds.

It is worth noting the discontinuity in the controls caused by the switch in the controller from the first to the second mode; see Figure 4.2. The switch to the stabilizing mode of the controller results in a large discontinuity in the controls that might be impossible to implement. It is natural to approximate the discontinuity with an acceptably smooth curve using interpolating splines, for example. However, smoothing the control gain in this manner could cause the system to deviate from the desired covariance before $K(S_f)$ is reached.

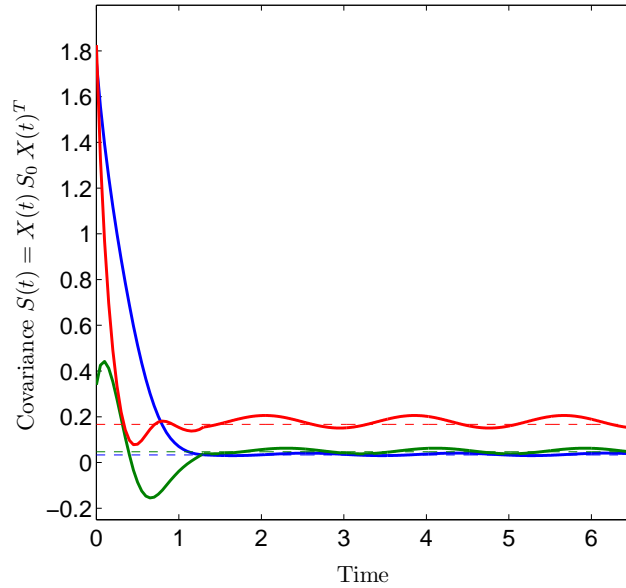


Fig. 4.1. The evolution of the covariance under the dual-mode receding horizon controller $\mathcal{RH}(0.05)$. The first mode of the controller brings the covariance as close as possible to the desired and the second applies ensures the covariance remains close to the desired covariance, shown dashed.

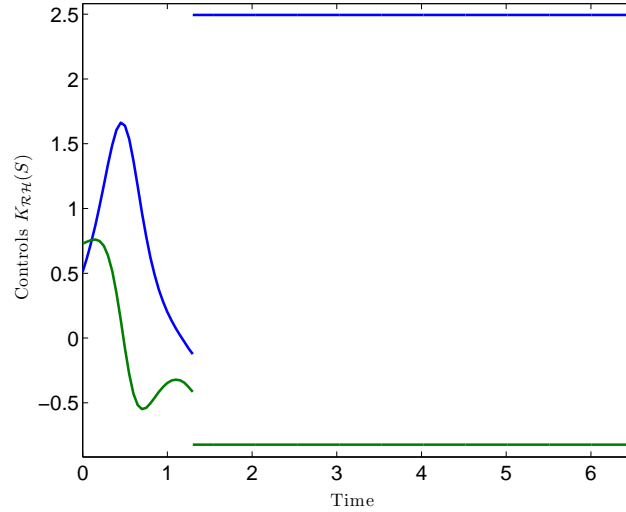


Fig. 4.2. The controls corresponding to the dual-mode receding horizon controller $\mathcal{RH}(0.05)$. Note the large discontinuity in the controls in the switch from the first to the second mode of the controller.

4.2 Lie's Method of Reduction

I conclude this chapter with a digression on a possible application of Lie's Method of Reduction, a classic approach to the integration of equations of Lie type, to the solution of the boundary value problem resulting from optimal control problem (3.6).

First, recall that the *isotropy subgroup* of the action by congruence is the subgroup $H_S^\Theta := \{X \in \mathbf{GL}_n : \Theta_X(S) = S\}$, and its Lie algebra is the *isotropy algebra* $\mathfrak{h}_S^\Theta := \{A \in \mathfrak{gl}_n : \Phi_A^\Theta(S) = 0\}$.

Lie's method of reduction begins with an arbitrary curve through a specified initial condition. Let $X_0 : [0, T] \rightarrow \mathbf{GL}_n$ be such a curve satisfying $\Theta(X_0(T), S_0) = S_f$. Let $X_1 : [0, T] \rightarrow \mathbf{GL}_n$ be the fundamental solution with differential equation $\dot{X}_1 = (A + BK_1)X_1$, satisfying $\Theta(X_1(T), S_1) = S_f$ where $S_1 = \Theta(Y, S_0)$, for some $Y \in \mathbf{GL}_n$ (S_1 and S_0 are congruent by Y). In this way we obtain two separate curves through S_1 , $\Theta(X_0(t)Y^{-1}, S_1)$ and $\Theta(X_1(t), S_1)$, terminating at the same point. Now, let $\tilde{X} :$

$[0, T] \rightarrow \mathbb{H}_{S_1}^\Theta$ be a curve in the isotropy subgroup of S_1 . Then, $\Theta(X_0(t) Y^{-1} \tilde{X}(t), S_1)$ is obviously also a curve through S_1 . Lie's Method of Reduction is to find the curve $\tilde{X}(t) \in \mathbb{H}_{S_1}^\Theta$ such that $X_0 Y^{-1} \tilde{X} = X_1$, that is, we seek the curve through the isotropy group of S_1 that transforms the arbitrary curve X_0 into a fundamental solution satisfying the differential equation $\dot{X}_1 = (A + BK_1)X_1$.

We proceed by deriving a differential equation for $\dot{\tilde{X}}$ for our particular application. Actually, this is a slight modification of Lie's method, see [34] for details. If $X_0 Y^{-1} \tilde{X} = X_1$ then

$$\begin{aligned} \dot{X}_1 &= (A + BK_1)X_1 = \dot{X}_0 Y^{-1} \tilde{X} + X_0 Y^{-1} \dot{\tilde{X}} \\ \dot{\tilde{X}} &= Y X_0^{-1} \left((A + BK_1)X_1 - \dot{X}_0 Y^{-1} \tilde{X} \right) \\ &= Y X_0^{-1} \left((A + BK_1)X_0 - \dot{X}_0 \right) Y^{-1} \tilde{X} \end{aligned}$$

If we assume that $\dot{X}_0 = (A + BK_0)X_0$ then this reduces to the right-invariant differential equation for $\tilde{X} \in \mathbb{H}_{S_1}^\Theta$,

$$\dot{\tilde{X}} = Y X_0^{-1} (B(K_1 - K_0)) X_0 Y^{-1} \tilde{X}.$$

That is, the difference between two different boundary value problems for the same differential equation can be expressed using a curve in the isotropy group whose evolution depends on the difference between controls. It remains to derive a numerical algorithm that utilizes this result in a direct or indirect method.

5. CONCLUSION

This thesis solved the problem of optimally controlling the covariance associated with a linear system from an arbitrary initial condition to a desired terminal condition. I showed that the formulation of the Hamiltonian system on the trivialization of the cotangent bundle resulted, with the transversality condition, in a boundary value problem with significantly fewer scalar unknowns. I implemented these results in a receding horizon control strategy suitable for implementation in a real-time setting. Judging from the dearth of literature on the subject, the resulting strategy appears to be the first example of a receding horizon controller formulated on a matrix Lie group, but suffers from discontinuities in the controls caused by switching between controller modes. I conclude with a number of recommendations for future work.

Geometric numerical integration and real-time optimal control: The conclusions of this thesis were illustrated on a planar linear system. The extension of these results, however, to linear systems of dimension $n \geq 3$, and even to nonlinear systems, necessitates numerical methods that are adapted to the geometric structure of the general linear group. In Chapter 3, a fast adaptive step-size Runge-Kutta integrator `ode113` designed for problems on \mathbb{R}^{n^2} was used to solve the boundary value problem. However, very high tolerances were necessary in order to preserve the isospectrality of the costate trajectories on \mathfrak{gl}_n^* and to ensure the Hamiltonian remained constant. Numerical experimentation also showed that the boundary value problem often could not be solved by relaxing these tolerances. It is possible that these difficulties were caused by the fact that `ode113` does not generate points that lie on the manifold of interest; it does not generate invertible matrices at every timestep. Noting also that fixed step-size Runge-Kutta schemes cannot preserve polynomial invariants of a dynamical system of degree $n \geq 3$ [49, Theorem 3.3, pg. 106], I provide the following conjecture.

- 1) The extension of this work to higher dimensional linear systems necessitates the use of fast algorithms that integrate *on* the manifold of interest.

It would be interesting to see how it might be possible to engineer trade-offs between accuracy and speed of numerical integration in a geometric integrator. Once a suitable set of direct and indirect methods for integration on Lie groups is identified, one might attempt to address the following question:

- 2) How can we achieve useful trade-offs between accuracy and computational speed in the real-time optimal control of control systems on matrix Lie groups?

Alternative receding horizon control strategies: The dual-mode receding horizon control strategy presented in the previous Chapter was based primarily on the approach presented in [48].

As a rule, optimal control problems with terminal constraints are more difficult to solve than those with terminal costs. According to the discussion in [47], it is possible to design a stabilizing receding horizon controller that has neither terminal costs nor terminal constraints, provided that the optimization horizon is sufficiently long. In our particular case, longer optimization horizons results in longer computational times and the inclusion of terminal costs usually resulted in an optimal control problem that was not right- or left-invariant. For this reason we designed the first mode of our receding horizon controller to solve a sequence of free-time optimal control problems with a terminal constraint. This was acceptable because the time to solve the boundary value problem decreased as the initial covariance neared the desired, and because the transversality condition for the trivialized Hamiltonian system reduced the number of scalar unknowns in the boundary value problem. As a consequence, however, it was necessary to design a second mode in the receding horizon control strategy to stabilize the system, introducing an unrealistic discontinuity in the controls.

- 3) Investigate alternative receding horizon implementations of optimal control problems on $\mathfrak{gl}_n^* \times \mathbf{GL}_n$ that do not rely on discontinuous controls.

Optimal control problems on quadratic matrix Lie groups: The objective of this thesis was to drive the covariance of a system as close as possible to a desired covariance. This goal was split into two separate objectives; the first was to control the system to a desired covariance and the second to fix the system at that covariance in steady-state. The latter objective is actually an example of an optimal control problem on a quadratic matrix Lie group. A special type of these optimal control problems was recently solved in [28].

- 4) Solve the problem of optimally controlling affine control systems in arbitrary quadratic matrix Lie groups.

REFERENCES

- [1] R. Bhattacharya, G. Balas, A. Kaya, and A. Packard, “Nonlinear receding horizon control of F-16 aircraft,” in *Proc. 2001 Amer. Control Conf.*, vol. 1, 2001, pp. 518–522 vol.1.
- [2] J. Rawlings and D. Mayne, *Model Predictive Control: Theory and Design*. Madison, WI: Nob Hill Publishing, 2008.
- [3] A. van der Schaft, “On a state space approach to nonlinear H_∞ control,” *Systems & Control Letters*, vol. 16, no. 1, pp. 1–8, 1991.
- [4] J. Primbs, “Nonlinear optimal control: a receding horizon approach,” Ph.D. dissertation, California Institute of Technology, 1999.
- [5] A. Hotz and R. E. Skelton, “Covariance control theory,” *International Journal of Control*, vol. 46, no. 1, pp. 13–32, 1987.
- [6] A. Ohara and T. Kitamori, “Geometric structures of stable state feedback systems,” *IEEE Trans. Autom. Control*, vol. 38, no. 10, pp. 1579–1583, Oct 1993.
- [7] R. Brockett, “Optimal control of the Liouville equation,” *AMS IP Studies in Advanced Mathematics*, vol. 39, p. 23, 2007.
- [8] J. Baillieul, “Geometric methods for nonlinear optimal control problems,” *Journal of optimization theory and applications*, vol. 25, no. 4, pp. 519–548, 1978.
- [9] A. Lasota and M. Mackey, *Probabilistic properties of deterministic systems*. Cambridge, UK: Cambridge University Press, 1985.
- [10] T. S. Parker and L. O. Chua, *Practical numerical algorithms for chaotic systems*. New York: Springer-Verlag, 1989.
- [11] M. W. Hirsch and S. Smale, *Differential equations, dynamical systems, and linear algebra*, ser. Pure and Applied Mathematics. New York-London: Academic Press [A subsidiary of Harcourt Brace Jovanovich, Publishers], 1974, vol. 60.
- [12] A. E. Bryson, Jr. and Y. C. Ho, *Applied optimal control*. Washington, D. C.: Hemisphere Publishing Corp., 1975, revised printing.
- [13] H. Schaub, J. Junkins, and J. Schetz, *Analytical Mechanics of Space Systems*, 2nd ed., ser. AIAA Education Series, J. A. Schetz, Ed. Reston, VA: American Institute of Aeronautics and Astronautics, 2009.
- [14] O. Junge, J. Marsden, and I. Mezic, “Uncertainty in the dynamics of conservative maps,” in *Proc. 43rd IEEE Conf. Decision Control*, vol. 2, Dec. 2004, pp. 2225–2230 Vol.2.
- [15] R. Brockett, *Finite dimensional linear systems*. New York: John Wiley & Sons, 1970, vol. 7.

- [16] R. Mañé, *Ergodic theory and differentiable dynamics*, ser. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Berlin: Springer-Verlag, 1987, vol. 8, translated from the Portuguese by Silvio Levy.
- [17] A. Halder and R. Bhattacharya, “Dispersion analysis in hypersonic flight during planetary entry using stochastic Liouville equation,” *Journal of Guidance, Control, and Dynamics*, vol. 34, no. 2, pp. 459–474, 2011.
- [18] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming,” Available from <http://stanford.edu/~boyd/cvx>, 2008.
- [19] R. Horn and C. Johnson, *Matrix Analysis*. Cambridge, UK: Cambridge University Press, 1990.
- [20] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities in system and control theory*, ser. SIAM Studies in Applied Mathematics. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 1994, vol. 15.
- [21] K. Grigoriadis and R. Skelton, “Alternating convex projection methods for covariance control design,” *International Journal of Control*, vol. 60, no. 6, pp. 1083–1106, 1994.
- [22] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [23] T. Iwasaki, R. Skelton, and M. Corless, “A recursive construction algorithm for covariance control,” *IEEE Trans. Autom. Control*, vol. 43, no. 2, pp. 268–272, Feb. 1998.
- [24] S.-i. Amari and H. Nagaoka, *Methods of information geometry*, ser. Translations of Mathematical Monographs. Providence, RI: American Mathematical Society, 2000, vol. 191, translated from the 1993 Japanese original by Daishi Harada.
- [25] L. Vandenberghe and S. Boyd, “Semidefinite programming,” *SIAM review*, pp. 49–95, 1996.
- [26] N. J. Higham, “Matrix nearness problems and applications,” in *Applications of matrix theory (Bradford, 1988)*, ser. Inst. Math. Appl. Conf. Ser. New Ser. New York: Oxford Univ. Press, 1989, vol. 22, pp. 1–27.
- [27] J. E. Marsden and T. S. Ratiu, *Introduction to mechanics and symmetry*, 2nd ed., ser. Texts in Applied Mathematics. New York: Springer-Verlag, 1999, vol. 17.
- [28] A. Bloch, P. Crouch, J. Marsden, and A. Sanyal, “Optimal control and geodesics on quadratic matrix Lie groups,” *Foundations of computational mathematics*, vol. 8, no. 4, pp. 469–500, 2008.
- [29] R. Brockett, “Lie theory and control systems defined on spheres,” *SIAM Journal on Applied Mathematics*, pp. 213–225, 1973.

- [30] M. Krstic and P. Tsiotras, “Inverse optimal stabilization of a rigid spacecraft,” *IEEE Trans. Autom. Control*, vol. 44, no. 5, pp. 1042–1049, May 1999.
- [31] M. Zefran and V. Kumar, “Planning of smooth motions on $se(3)$,” in *Proc. 1996 IEEE International Conference on Robotics and Automation*, vol. 1, Apr. 1996, pp. 121–126 vol.1.
- [32] A. Arvanitoyeorgos, *An introduction to Lie groups and the geometry of homogeneous spaces*, ser. Student Mathematical Library. Providence, RI: American Mathematical Society, 2003, vol. 22, translated from the 1999 Greek original and revised by the author.
- [33] V. Jurdjevic, “Optimal control on Lie groups and integrable Hamiltonian systems,” *Regular and Chaotic Dynamics*, vol. 16, no. 5, pp. 514–535, 2011.
- [34] R. L. Bryant, “An introduction to Lie groups and symplectic geometry,” in *Geometry and quantum field theory (Park City, UT, 1991)*, ser. IAS/Park City Math. Ser. Providence, RI: American Mathematical Society, 1995, vol. 1, pp. 5–181.
- [35] H. Munthe-Kaas, “High order Runge-Kutta methods on manifolds,” *Applied Numerical Mathematics*, vol. 29, no. 1, pp. 115–127, 1999.
- [36] A. A. Agrachev and Y. L. Sachkov, *Control theory from the geometric viewpoint*, ser. Encyclopaedia of Mathematical Sciences. Berlin: Springer-Verlag, 2004, vol. 87, control Theory and Optimization, II.
- [37] G. Walsh, R. Montgomery, and S. Sastry, “Optimal path planning on matrix lie groups,” in *Proc. 33rd IEEE Conf. Decision Control*, vol. 2, Dec. 1994, pp. 1258–1263 vol.2.
- [38] V. Jurdjevic, *Geometric control theory*, ser. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 1997, vol. 52.
- [39] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*. Princeton, NJ: Princeton University Press, 2008, with a foreword by Paul Van Dooren.
- [40] U. Helmke and J. B. Moore, *Optimization and dynamical systems*, ser. Communications and Control Engineering Series. London: Springer-Verlag, 1994, with a foreword by R. Brockett.
- [41] J. Betts, “Survey of numerical methods for trajectory optimization,” *Journal of Guidance, Control, and Dynamics*, vol. 21, no. 2, pp. 193–207, 1998.
- [42] J. Jakubiak, F. S. Leite, and R. C. Rodrigues, “A two-step algorithm of smooth spline generation on riemannian manifolds,” *Journal of Computational and Applied Mathematics*, vol. 194, no. 2, pp. 177–191, 2006.
- [43] K. Engø, A. Marthinsen, and H. Munthe-Kaas, “Diffman: An object-oriented matlab toolbox for solving differential equations on manifolds,” *Applied numerical mathematics*, vol. 39, no. 3, pp. 323–347, 2001.

- [44] K. Engø and S. Faltinsen, “Numerical integration of Lie-Poisson systems while preserving coadjoint orbits and energy,” *SIAM Journal on Numerical Analysis*, pp. 128–145, 2002.
- [45] R. Freeman and P. Kokotovic, “Inverse optimality in robust stabilization,” *SIAM Journal Control and Optimization*, vol. 34, no. 4, pp. 1365–1391, 1996.
- [46] L. Magni and R. Sepulchre, “Stability margins of nonlinear receding-horizon control via inverse optimality,” *Systems & Control Letters*, vol. 32, no. 4, pp. 241–245, 1997.
- [47] A. Jadbabaie and J. Hauser, “On the stability of receding horizon control with a general terminal cost,” *IEEE Trans. Autom. Control*, vol. 50, no. 5, pp. 674 – 678, May 2005.
- [48] H. Michalska and D. Mayne, “Robust receding horizon control of constrained nonlinear systems,” *IEEE Trans. Autom. Control*, vol. 38, no. 11, pp. 1623 –1633, Nov. 1993.
- [49] E. Hairer, C. Lubich, and G. Wanner, *Geometric numerical integration*, ser. Springer Series in Computational Mathematics. Heidelberg: Springer, 2010, vol. 31, reprint of the second (2006) edition.