

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Techniques for the optimization and control of large-scale systems
with application to jet noise

A dissertation submitted in partial satisfaction of the
requirements for the degree Doctor of Philosophy
in Engineering Sciences (Aerospace Engineering)

by

Laura Isabel Cerviño

Committee in charge:

Thomas R. Bewley, Chair
Miroslav Krstic
Sutanu Sarkar
Williams McEneaney
Philip Gill
Mohammed Ziane

2005

UMI Number: 3158468

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform 3158468

Copyright 2005 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

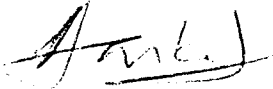
ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

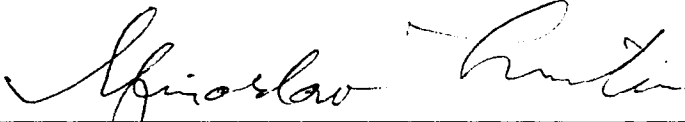
Copyright

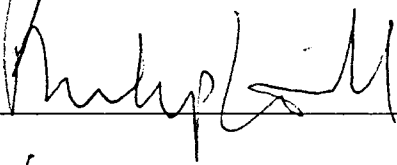
Laura Isabel Cerviño, 2005

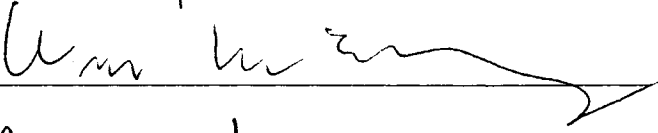
All rights reserved.

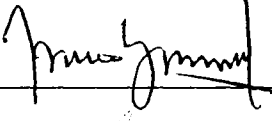
The dissertation of Laura Isabel Cerviño is approved, and it is acceptable in quality and form for publication on micro-film:

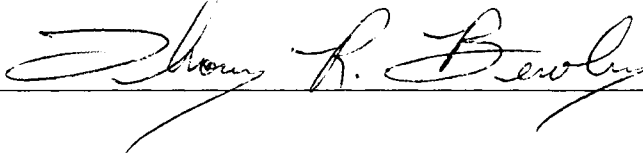












Chair

University of California, San Diego

2005

TABLE OF CONTENTS

Signature Page	iii
Table of Contents	iv
List of Figures	vi
Acknowledgements	viii
Vita	ix
Abstract	xi
1 Introduction	1
2 Adjoint analysis and optimization of flow acoustic interactions in a 2D jet . . .	7
1. Introduction	7
2. Perturbation and Adjoint analyses	9
3. Description of the system and simulation technique	12
1. Boundary conditions	15
4. Derivation of the adjoint system	16
5. Adjoint analysis	19
1. Calculation of an adjoint Green's function	19
2. An adjoint Green's function at temporal frequency f corresponding to far-field noise	20
3. Quantification of scattering of adjoint Green's functions	23
6. Optimization	27
1. Cost function	27
2. Gradient method	27
3. Controller	28
4. Definition of the adjoint field and derivation of the gradient	30
5. Control update	33
6. Control computation	33
7. Results	34
7. Concluding remarks	35
Appendix. Formulation of the equations in cylindrical coordinates	36
1. Introduction	36
2. Governing equations	36
3. Numerical problems	42
References	47

3	Extension of the complex-step derivative technique to pseudospectral algorithms	50
1.	Introduction	50
2.	Background	51
1.	The Finite-Difference (FD) approximation	51
2.	The Complex-Step Derivative (CSD) approximation	52
3.	Extension of the CSD approach to Fourier-based pseudospectral codes . .	53
4.	Discussion and extension to other pseudospectral methods	57
	References	58
4	Identification of time-periodic orbits in chaotic systems	59
1.	Introduction	59
2.	Obtaining several approximately time-periodic trajectories	62
3.	Refining the trajectories with time-evolving iterates	65
1.	Derivation of gradients	66
2.	Estimation of descent parameter	69
3.	Summary of numerical algorithm	71
4.	Refining the trajectories with time-periodic iterates	73
1.	Derivation of gradients	73
2.	Estimation of descent parameter	75
3.	Summary of numerical algorithm	75
5.	Numerical results	77
6.	Discussion	81
7.	Conclusions	82
	Appendix. Estimation of the descent parameter	83
1.	Introduction	83
2.	Description of the problem	83
3.	Derivation of gradients	84
4.	Estimation of the descent parameter	86
	References	93

LIST OF FIGURES

2.1	Perturbation analyses (left) characterize control $\rightarrow$ effect relationships. On the other hand, adjoint analyses (right) characterize effect $\rightarrow$ control relationships.	10
2.2	Perturbation analysis (top, representation of perturbation pressure) characterizes the effect on the entire flow resulting from a small change to a particular control quantity, taken here to be a sinusoidally-varying mass source at point x_c . Adjoint analysis (bottom, representation of adjoint density) characterizes the effect on a particular flow quantity, taken here to be high frequency noise at point x_e , due to small changes in the control applied anywhere in the flow. Vorticity contours are superimposed to the colored perturbation and adjoint fields.	11
2.3	Adjoint analysis (right) of sound waves (left) produced by a monopole sound source at the point marked by the X, in a stationary fluid. The adjoint field is forced by the acoustic energy measured in the rectangular area.	12
2.4	Visualization of the vortex breakdown in the two-dimensional strong jet.	14
2.5	Equivalence of a perturbation (left) and adjoint (right) problems when the boundary terms in the adjoint identity are 0.	18
2.6	Adjoint density (top) and adjoint pressure (bottom) reveals sensitivity of the pressure component of the perturbation field at point x_e at time t_e to additional control forcing in, respectively, the energy equation (top) and the continuity equation (bottom) everywhere in space x_c and for all times $t_c < t_e$. Note that, by causality, the adjoint field is zero for $t_c > t_e$; that is, the adjoint field marches backward in time from $t = t_e$	19
2.7	Evolution of adjoint pressure in time at the points $\{x, y\}$ of (solid) $\{5D, 0\}$, (dashed) $\{5D, 2.5D\}$, (dot-dashed) $\{5D, -2.5D\}$	20
2.8	Evolution of adjoint density in time at the points $\{x, y\}$ of (solid) $\{5D, 0\}$, (dashed) $\{5D, 2.5D\}$, (dot-dashed) $\{5D, -2.5D\}$	21
2.9	Evolution of the adjoint pressure at three different locations at the centerline: at (dot-dashed) $x = 8$, (dashed) $x = 9$, and (solid) $x = 10$	21
2.10	Evolution of the adjoint pressure at three different locations at the centerline: at (dot-dashed) $x = 8$, (dashed) $x = 9$, and (solid) $x = 10$ shifted the time corresponding to the convection velocity. There is an approximate superposition of the three lines, which indicates that these perturbations convect toward the nozzle at the convective speed of the jet.	22
2.11	Adjoint density field due to incoming waves from the far field.	23
2.12	Adjoint pressure wave corresponding to far-field noise at an angle of 60° off the jet axis and at a frequency of $St = 2.0$	24

2.13	Temporal spectra measured at the indicated points $\{x, y\}$ of (solid) $\{5D, 0\}$, (dashed) $\{5D, 2.5D\}$, (dot-dashed) $\{5D, -2.5D\}$ of (left) the adjoint pressure $\hat{p}_f^*$ and (right) the adjoint density $\hat{\rho}_f^*$ of incoming waves at the same angle and at a frequency of (top) $St = 0.8$, (center) $St = 2.0$, and (bottom) $St = 8.0$	25
2.14	Spectra of the vertical velocity of the flow at $\{x, y\} = \{1.5D, 0.5D\}$ (left) and $\{x, y\} = \{3D, 0.5D\}$ (right). The appearance of the peak at $S_t = 0.2$ at the downstream station (right) is a result of vortex pairing (recall that the jet is forced at $S_t = 0.4$).	26
2.15	Physics of the flow (top) and adjoint (bottom) fields.	29
2.16	Evolution of the adjoint density.	35
3.1	Relative error of the directional derivative of $f(x) = \frac{e^{-x}}{\sqrt{\tan x}}$ at $x = 1$ given by (solid) first-order, (dashed) second-order, and (dot-dashed) fourth-order FD approaches and (circles+dots) the CSD approach using single-precision arithmetic (left) and double-precision arithmetic (right). Note that both plots have essentially the same shape.	54
3.2	Relative error of (dashed) first-order and (dot-dashed) second-order FD approaches and the CSD approach using (solid) strategy A and (solid-circle) strategy B.	56
4.1	Two different views of the "skeleton" formed by the periodic orbits in the attractor (colored tubes), and a generic trajectory (thin black line). The generic trajectory passes close to the periodic orbits due to the ergodic nature of the system.	61
4.2	Periodic orbits with $T < 4$	78
4.3	Convergence history of the cost function defined in Eq. (4.4) in the time-evolving setting as a function of the optimization iteration.	80
4.4	Spectrum of the periodic orbit shown in Fig. 4.2(g).	81
4.5	Convergence history of the cost function defined in the time-periodic setting as a function of the optimization iteration (top) and same cost function nondimensionalized with the initial value (bottom).	90
4.6	Proximity of the attractor to a set of periodic orbits.	91
4.7	Estimated $\alpha_{est} = 0.1$ and optimal α_{opt}	91
4.8	Estimated $\alpha_{est} = 0.01$ and optimal α_{opt}	92
4.9	α estimated with the method described in Eq. (4.46).	92

ACKNOWLEDGEMENTS

I would like to thank Prof. Thomas R. Bewley for his continuous guidance and inspiration in these last years. It has been a pleasure to work with him and to get involved in his research projects.

I would also like to thank the people in my research group, Peter Blossey, Bartosz Protas, Patricia Cathalifaud, Catalin Trenchea, and Scott Miller, for their help and encouragement, and especially to Dr. Haoxiang Luo, with whom I shared office most of this time and with whom I started the program. Haoxiang and I have gone through the PhD together, and his help has been priceless.

I would like to thank Profs. Montañes and Lasheras for working on making possible for me to come here, and for their advice in these years.

I want to thank Dr. Rodriguez-Rodriguez and Dr. Mellado for proofreading parts of the thesis.

On a personal level, I want to mention my relatives in Spain, who gave me support during all this time and encouraged me throughout the good and bad moments, especially to Maribel, Valentin, Javier, and Angel.

I want to thank Profs. Jon Freund and Sanjiva Lele for helpful discussions and hosting our visit to the summer program of the Center for Turbulence Research (CTR), where much of the work in Chapter 2 was performed.

Finally I want to thank AFOSR (Dr. John Schmisser and Dr. Belinda King) for generous funding in support of this research.

The text of Chapter 3, in full, is a reprint of the material as it appears in Cerviño & Bewley, 2003, “On the extension of the complex-step derivative technique to pseudospectral algorithms”. *J. Comput. Phys.* 187, 544–549. The dissertation author was the primary author and the coauthor listed in this publication directed and supervised the research which forms the basis for this chapter.

VITA

Education:

- 1999 B.S., Aerospace Engineering, Universidad Politécnica de Madrid, Spain.
- 2001 M.S. Engineering Sciences (Aerospace Engineering), University of California, San Diego.
- 2005 Ph.D. Engineering Sciences (Aerospace Engineering), University of California, San Diego.
Techniques for the optimization and control of large-scale systems with application to jet noise.
GPA: 4.00

Professional Experience:

- Oct. 1999-Sep. 2004
Graduate research assistant, University of California, San Diego, Department of Mechanical and Aerospace Engineering.
- Oct. 2001-Dec. 2001
Teaching assistant for *Numerical Methods in Science and Engineering*, University of California, San Diego. Department of Mechanical and Aerospace Engineering.
- Oct. 2002-Dec. 2002
Teaching assistant for *Numerical Methods in Science and Engineering*, University of California, San Diego. Department of Mechanical and Aerospace Engineering.
- Summer 2003
Lecturer for *Introduction to mathematical physics*, University of California, San Diego. Department of Mechanical and Aerospace Engineering.

Awards and Honors:

- Fellowship at the Von Karman Institute for Fluid Dynamics, summer 1998.
- Fellowship at the Department of Aerospace Vehicles, 1998-99, conferred by Universidad Politécnica de Madrid.
- IBM Research Fellowship, 2000-01.
- Fellowship from Helmholtz Institute for Supercomputational Physics for the program *First summer school: Tools to Simulate Turbulence on Supercomputers*, Potsdam, Germany, Summer 2001.
- Academic award (Graduate Research), conferred by the Department of Mechanical and Aerospace Engineering at University of California, San Diego, June 2002.

Academic award (Graduate Research), conferred by the Department of Mechanical and Aerospace Engineering at University of California, San Diego, June 2003.

Final Dissertation Fellowship, Department of Mechanical and Aerospace Engineering, University of California, San Diego, Fall 2004.

Publications and Presentations:

- L. I. Cerviño and T. R. Bewley, *New adjoint-based tools for identification and control of flow-induced noise*, 54th Annual Meeting of the APS Division of Fluid Dynamics, San Diego, USA, November 2001.
- L. I. Cerviño, T. R. Bewley, J. B. Freund, and S. K. Lele, *Perturbation and adjoint analyses of flow-acoustic interactions in an unsteady 2D jet*, Center for Turbulence Research, Proceedings of the Summer Program 2002, 27-40 (2002).
- L. I. Cerviño and T. R. Bewley, *Adjoint analysis of flow/acoustic interactions in a 2D jet*, 55th Annual Meeting of the APS Division of Fluid Dynamics, Dallas, USA, November 2002.
- L. I. Cerviño and T. R. Bewley, *Adjoint-based analysis of noise control opportunities in a 2D jet*, AMS–IMS–SIAM Summer Research Conference on Hydrodynamics Stability and Flow Control, Snowbird, Utah, USA, July 2003.
- L. I. Cerviño and T. R. Bewley, *On the extension of the complex-step derivative technique to pseudospectral algorithms*, J. Comp. Phys., **187**, 544-549 (2003).
- L. I. Cerviño and T. R. Bewley, *Exploring controls for unsteady jet systems using adjoint analysis*, 7th Southern California Nonlinear Control Workshop, University of Southern California, USA, November 2003.
- L. I. Cerviño, S. Miller and T. R. Bewley, *Characterizing the stabilizability of complex flows using adjoint eigenfunctions*, 56th Annual Meeting of the APS Division of Fluid Dynamics, NYC/New Jersey, USA, November 2003.
- L. I. Cerviño and T. R. Bewley, *Adjoint Analysis and Control Opportunities in a 2D Jet*, Proceedings of the 42nd IEEE Conference on Decision and Control, 1642-1647, Maui, Hawaii, December 2003.
- L. I. Cerviño and T. R. Bewley, *Identification of time-periodic orbits in chaotic systems*, 57th Annual Meeting of the APS Division of Fluid Dynamics, Seattle Washington, USA, November 2004.

ABSTRACT OF THE DISSERTATION

Techniques for the optimization and control of large-scale systems
with application to jet noise

by

Laura Isabel Cerviño

Doctor of Philosophy in Engineering Sciences (Aerospace Engineering)

University of California, San Diego, 2005

Professor Thomas R. Bewley, Chair

The present work focuses on the analysis and optimization of jet noise and some of the related challenges that high-dimensional approximations of infinite-dimensional, chaotic, multiscale systems present to the gradient-based optimization framework.

We first present the optimization of flow/acoustic interactions in jets. This optimization is based on a gradient obtained from an adjoint field. The ability to modify favorably the high-frequency acoustic field via low-frequency modulation of the hydrodynamic field near the jet exit is confirmed with this method. Due to the complexity of the system under consideration, some simplifications have been applied in the derivation of equations governing the adjoint field. In order to evaluate the correctness of the adjoint code and the adequacy of these simplifications, complex-step derivative (CSD) method has been used to calculate the gradient directly from the flow solver. Until now, the CSD method has been applied only to physical space simulation codes. In the present work, a non-trivial extension of this technique to pseudospectral codes has been developed, as many of the numerical codes used for turbulent flows problems leverage pseudospectral techniques to calculate spatial derivatives in one or more directions.

Once the gradient information obtained with the adjoint field has been so validated, optimizations may be performed. Note that the control schedule that we have designed for the jet system is periodic in time, due both to the quasi-periodic nature of the jet system and reasons of practical implementation. We have encountered certain fundamental

challenges related to the fact that, when applied to time evolving systems, the adjoint field grows exponentially backward in time, resulting in a gradient dominated by effects in a narrow time window no matter what time horizon is used for the optimization. In such a situation, one can not leverage the ergodicity of the controlled chaotic system (by integrating over a sufficiently long time horizon) in order to achieve a general applicability of the control design. New methods to treat such control problems directly in the time-periodic framework have been proposed and are currently under investigation. A difficult step in such time-periodic control optimizations is the development of efficient techniques to obtain the time-periodic orbits of a chaotic system, as will be presented.

Chapter 1

Introduction

The noise radiated from the exhaust flow of jet engines is

- a well-known public nuisance both near airports and inside the cabin of modern commercial jet aircraft,
- a leading cause of structural damage to the tail surfaces of high-performance military aircraft,
- a persistent safety hazard for military personnel working on the flight line, and
- a significant financial concern for airlines and airplane/engine manufacturers due to strict governmental regulations on jet noise, both in the US and Europe.

Passive control techniques have already been applied to great extent to both subsonic and supersonic aircraft by decreasing the speed of the jet core, changing the shape of the jet, and adjusting the convective Mach number in the shear layer(s) of the jet. This can be accomplished by using a large diameter jet exit, a corrugated or non-axisymmetric nozzle, tabs near the nozzle lip that protrude into the jet, and adding an appropriate coflow. Such techniques have already been explored, perhaps to nearly the maximum extent possible, to achieve large reductions in jet noise. However, the projected growth in air traffic and increased sensitivity to noise near airports in the years to come will demand even quieter aircraft, keeping the problem of the mitigation of jet noise an important focus of research for the foreseeable future.

Active control techniques have the potential to reduce jet noise well below the levels already attained by passive strategies. Although active control strategies have not yet been implemented for the jet noise problem in commercial engines, actuators for active jet control (using high pressure air ducted from the compressor) have already been developed and tested on full size jet engines for other problems, such as jet mixing enhancement. The size of the actuators used in such tests is still much larger than ultimately desired in practical applications. However it is expected that the size of such actuators may be decreased significantly in the future, as control strategies are made more efficient. In order to coordinate an array of actuators to achieve the desired effect (that is, to reduce the noise in a specified direction from the jet), improved optimization strategies that are suitable for very large-scale chaotic systems must be developed.

Note that, at appreciable Reynolds numbers, the jet system under consideration is, for all practical purposes, both *uncontrollable* and *unstabilizable*. That is, using actuation only at the jet exit (as motivated by practical implementation issues), there is not sufficient control authority to stabilize the system with feedback control and thereby eliminate the transition to turbulence further downstream. This is typical in many flow systems (bluff body wakes, plane shear layers, etc.) dominated by a strong convective component, where the actuation is significantly far upstream of the flow instability leading to transition to turbulence. Effectively, such systems have a significant *time lag* between the application of a control input and the effect of this control on the regions of instability in the flow system. This being said, actuation near the jet nozzle, where the shear layers are thin, is capable of altering the nonlinear dynamics of the breakdown of the jet in a very significant fashion. The goal of this research is thus to determine efficient nonlinear *optimization* techniques (as opposed to feedback control techniques) to adjust the unsteady hydrodynamic field of a jet via an appropriate schedule of open-loop control forcing near the nozzle in order to modify the resulting acoustic field in an appropriate fashion. That is, we seek to develop a gradient-based strategy to tune a predetermined, time-periodic forcing schedule in order to tailor the statistics of the jet in a favorable manner.

Note also that simple “anti-noise” strategies that produce directly an acoustic field that is 180 degrees out of phase with the broad-band acoustic field of the jet are impractical in the present problem, primarily because such strategies require both far-field sensors and, for maximum effectiveness, actuators that are close to the region of the jet that the noise is created. Thus, “anti-noise” strategies are not the focus of the present research.

In a delicate large-scale system such as a turbulent jet, new optimization techniques must be developed that are both accurate and computationally tractable. In expensive high-dimensional optimization problems such as this, gradient-based strategies are numerically much more efficient than function-based (that is, derivative-free) approaches. In a dynamic system such as the present, it is commonly known that the required gradient may be determined via the calculation of an appropriately-defined *adjoint field*. However, high-dimensional (accurate) approximations of infinite-dimensional, chaotic, multiscale systems present a number of significant challenges to the gradient-based optimization framework. A major goal of the present research is to identify and address these challenges.

As laid out in greater detail below, this thesis first focuses on presenting the fundamental optimization problem of interest, that is, the optimization of flow/acoustic interactions in jets. As mentioned above, this optimization is based on a gradient obtained from an adjoint field. Due to the complexity of the system under consideration, some simplifications have been applied in the derivation of equations governing the adjoint field. In order to evaluate the correctness of the adjoint code and the adequacy of these simplifications, an expensive yet highly accurate method, the complex-step derivative (CSD) method, has also been used to calculate the gradient directly from the flow solver, for the purpose of comparison with the gradient obtained from the adjoint field. Until now, the CSD method has been applied only to physical space (e.g., finite difference or finite element) simulation codes. In the present work, a non-trivial extension of this technique to pseudospectral codes has been developed. This new technique is of broad applicability, as many of the numerical codes used for turbulent flow systems and weather forecasting problems leverage pseudospectral techniques to calculate spatial derivatives in one or

more directions.

Once the gradient information obtained with the adjoint field has been so validated, optimizations may be performed. Note finally that the control schedule that we have designed for the jet system is periodic in time, due both to the quasi-periodic nature of the jet system and reasons of practical implementation. We have encountered certain fundamental challenges related to the fact that, when applied to time evolving systems, the adjoint field grows exponentially backwards in time, resulting in a gradient dominated by effects in a narrow time window no matter what time horizon is used for the optimization. In such a situation, one can not leverage the ergodicity of the controlled chaotic system (by integrating over a sufficiently long time horizon) in order to achieve a general applicability of the control design to other realizations of the flow from different initial conditions. Thus, new methods to treat such control problems directly in the time-periodic framework have been proposed and are currently under investigation. A difficult yet important step in such time-periodic control optimizations is the development of efficient techniques to obtain the time-periodic orbits of a chaotic system, as laid out in the final chapter of this thesis.

Chapter 2 - Adjoint analysis and optimization of flow acoustic interactions in a 2D jet

The first chapter of the thesis presents the problem of jet noise reduction. The acoustic field radiated by the jet is computed accurately via Direct Numerical Simulation. The relevant adjoint field equations are derived, implemented numerically, and validated with the Complex-Step Derivative method (described further in chapter 3). Results for a 2D jet are illustrated, providing a foundation for 3D optimizations in the future. A detailed study of the information that can be gained from the adjoint analysis of the jet system is first performed. It is observed that low-frequency modulation of the hydrodynamic field near the jet exit can have a very significant effect on the high-frequency acoustic radiation to the far field (as mentioned previously, an “anti-noise” approach, in which the control forcing is at the the same frequency as the far field noise of concern in the cost function, is not the aim of the present research). Finally, an optimization is performed of low-frequency

forcing of the flow field near the jet exit in the 2D setting, using a gradient-based strategy leveraging iterative state and adjoint calculations.

Chapter 3 - Extension of the complex-step derivative technique to pseudospectral algorithms

When computing a directional derivative leveraging an existing (perhaps, extremely complicated) nonlinear simulation code using a simple Finite Difference (FD) strategy, it is critical to use a perturbation amplitude that will provide an accurate result. Unfortunately, such an appropriate perturbation amplitude is often quite difficult to determine. An alternative method to compute the directional derivative leveraging an existing nonlinear simulation code that circumvents this problem is called the Complex-Step Derivative (CSD) method. The CSD method makes use of complex variables and determines an accurate gradient for a very wide range of (sufficiently small) perturbation amplitudes. The CSD method has been implemented for checking the accuracy of the gradient of the acoustic energy of the jet with respect to the control parameters in Chapter 2. In Chapter 3, we extend the CSD method to pseudospectral simulation codes (which already incorporate complex arithmetic). This is a necessary development to extend this powerful method to compute directional derivatives in many turbulent simulation codes, such as those used for the present 3D turbulent jet as well as for many weather forecasting applications.

Chapter 4 - Identification of time-periodic orbits in chaotic systems

Many turbulent flow systems are dominated by large-scale coherent structures with quasi-periodic behavior, as is the case for the jet system under consideration here. In such systems, optimizations of time-periodic forcing schedules that are performed over finite time horizons might not necessarily generalize well to different realizations of the turbulent flow (e.g., from different initial conditions). A new strategy to favorably reconfigure (but not stabilize) an entire chaotic attractor, via optimization of the time-periodic orbits embedded within such an attractor, is under development by our group. An essential component of this control strategy is the development of efficient algorithms to

extract the first several periodic orbits of the chaotic system. In this chapter, methods to obtain such periodic orbits (up to a period T_{max}) are developed and evaluated for a model low-dimensional chaotic system.

Future work

An efficient, parallelized optimization code has been developed that is equipped to perform noise optimizations in a 3D round turbulent jet. In this thesis, this code has been effectively validated in the 2D setting. In future work, an adjoint analysis similar to that performed in Chapter 2 (on the 2D problem) will be performed on the full 3D problem in order to study the possibility of reducing the noise in the physical system of interest. Ultimately, once thoroughly validated on a variety of smaller problems, this code will be leveraged to perform numerically-expensive iterative state/adjoint optimizations of the full 3D system.

Chapter 2

Adjoint analysis and optimization of flow acoustic interactions in a 2D jet

2.1 Introduction

Jet exhaust is one of the dominant noise sources in modern turbofan engines, having significant engineering consequences, such as quality of life near airports (which has become an important environmental issue), financial incentives for airlines (regulations in US and Europe), safety concerns for carrier operations, and metal fatigue in tail surfaces of high-performance aircraft.

In order to reduce this noise radiated by turbulent jets, passive control strategies have already been pursued extensively, with great effect. Modern aircraft are about 20dB quieter than those of 50 years ago. Much of this reduction can be attributed to the increase of the bypass of turbofans in order to lower the nozzle exit velocity U (sound power level P is proportional to U^8). Tabs (protrusions into the flow) and chevrons (serration of the nozzle lip) disrupt the uniformity of vortex shedding by producing streamwise vortices (chevrons more gently than tabs), in order to mix the core flow with the fan flow, and the fan flow with the ambient flow, therefore reducing the noise (see Saiyed *et al.* (2000), where a 2.7dB reduction is achieved by using chevrons in both the core and the fan nozzles). These methods normally achieve noise reduction at the expense of cruise

thrust losses, being necessary to perform a trade-off between these two effects.

In supersonic jets, noise reduction has been achieved by modifying the shape of the nozzle (for example, making it non-axisymmetric). These modifications aim reduction of the convective Mach number, yielding a reduction in noise, and have been applied to both the supersonic regime (see, e.g., Papamoschou (1997)), and, more recently, subsonic turbofans (Papamoschou (2003)).

As an alternative to passive control, active control of jet noise appears to hold the potential for a significant additional impact on demand (near takeoff and landing) without disturbing the engine performance at cruise conditions, and forms the focus of the present work. To the best of our knowledge, no active control strategy has yet succeeded for the purpose of jet noise reduction due mainly to the complex nature of the physics involved and the high dimensional aspect of the control forcing schedule to be optimized, although effective, unsteady, individually-controllable actuators at the exit of the nozzle have already been developed and tested (at full scale) for the problem of jet mixing enhancement.

In order to obtain an optimal control schedule, we use adjoint-based nonlinear-system optimization, a popular technique in the flow control community. The adjoint system, when defined and calculated appropriately, gives very accurate gradient information with which controls may be tuned even in high dimensional systems with high dimensional controls. The performance of a control distribution optimized via this method cannot be guaranteed to be globally optimal, but it often far exceeds that possible with other control design techniques. Adjoint-based gradient optimization has proven to be effective in the analysis, control, optimization and forecasting of incompressible turbulence (see, e.g., Bewley *et al.* (2001)). Recently, there has been an increased interest in extending this approach to compressible flows (see, for example Wei & Freund (2002), where a 6.3dB reduction of noise in a 2D shear layer is obtained). The ultimate goal of the present project is to achieve 3D jet noise reduction via active control, obtaining an optimal azimuthal and temporal distribution of the actuator forcing with the adjoint-based method.

The high dimensionality of the jet when discretized makes the optimization procedure very expensive. Nonetheless, important information regarding control opportunities

can be gained from an adjoint analysis, and evaluation of whether the type of control which we intend to apply is appropriate for the problem is possible. The adjoint field reveals the sensitivity of the cost function to modification of the control actuation (furthermore, if an eigenvalue/eigenvector analysis were performed, it might be used to characterize the stabilizability of the system, as discussed in detail in Lauga & Bewley (2003)). Adjoint analysis thus provides quantification of the stabilizability and the suitability of proposed actuator configurations in a large system such as the turbulent jet, difficult to obtain by other methods.

In the first sections of the present chapter an adjoint analysis of the sensitivity of the far-field acoustic field to changes in the control is performed. It will be seen that it is possible to modify the hydrodynamic field at a low frequency and alter in that way the high-frequency far-field acoustics. This conclusion yields to the implementation of an optimization algorithm, considering periodic in time actuation (with low frequency). The performance of this actuation is also discussed.

2.2 Perturbation and Adjoint analyses

As summarized in Fig. 2.1 (left), perturbation analyses, which simulate directly the effect on the flow of a perturbation to the control distribution, characterize **control**→**effect** relationships (i.e., *if I change the control here, how and where will that effect the flow?*). A representative perturbation analysis of the present system is shown in the top row of Fig. 2.2. This analysis was obtained by the Complex Step Derivative method, which has been broadly used in the optimization literature (see, e.g., Lyness & Moler (1967), Martins *et al.* (2001), Cerviño & Bewley (2003)). Perturbation analyses characterize the propagation of disturbances in the system as it evolves forward in time. In order to perform an extensive investigation of the effect of different control possibilities, the perturbation field must be computed once for each control variable (whether it is a possible location of the actuators or a parameter), not giving at once a global view of the sensitivity of the cost function to changes in the control distribution.

In an adjoint analysis, an field system is defined and marched backward in time in order to identify the gradient of the cost function of interest to additional control forcing in the system. As depicted in Fig. 2.1 (right), such analyses characterize **effect**→**control** relationships (i.e., *if I want to achieve a desired effect here, how and where should I apply control to the flow?*). A representative adjoint analysis of the present system is depicted in the bottom row of Fig. 2.2. Once the cost function is defined, one adjoint simulation gives at once information about the sensitivity of this cost function to control actuation everywhere in the flow domain. Thus, for the purpose of control optimization, adjoint analyses provide much more valuable information than perturbation analyses, especially when working with high dimensional controllers.

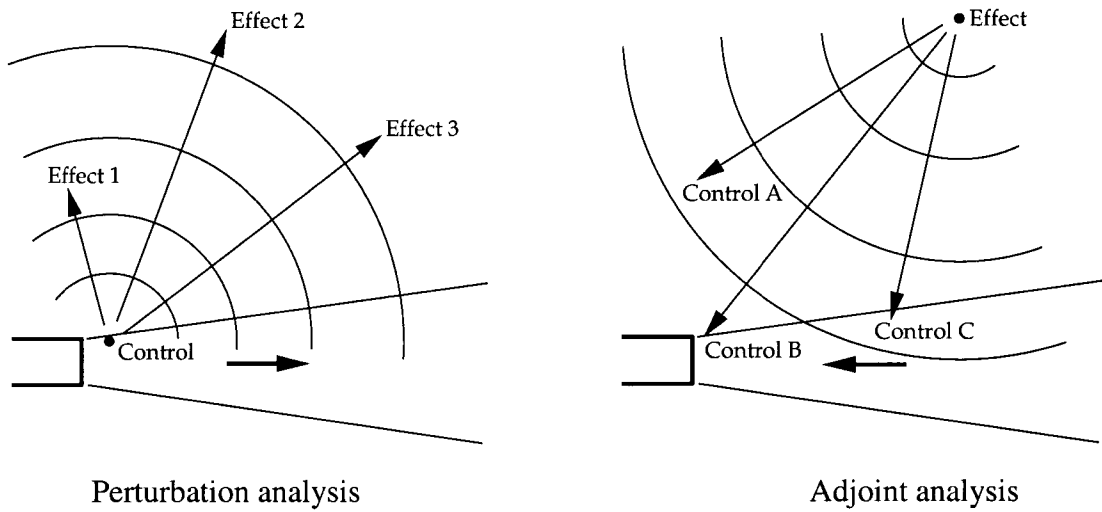


Figure 2.1: Perturbation analyses (left) characterize control → effect relationships. On the other hand, adjoint analyses (right) characterize effect → control relationships.

It is important to note that adjoint analyses do *not* identify the “origin” or “source” of the radiated sound in a system such as a turbulent jet. Rather, they identify how and where additional forcing may be applied to the existing system to modify in a desired manner the radiating noise already present. This point is readily evident by considering a simpler model system (without the jet present), as depicted in Fig. 2.3, which shows the acoustic field from a monopole, and the corresponding adjoint representing the sensitivity of the noise measured in the marked rectangular interrogation area with respect to control actu-

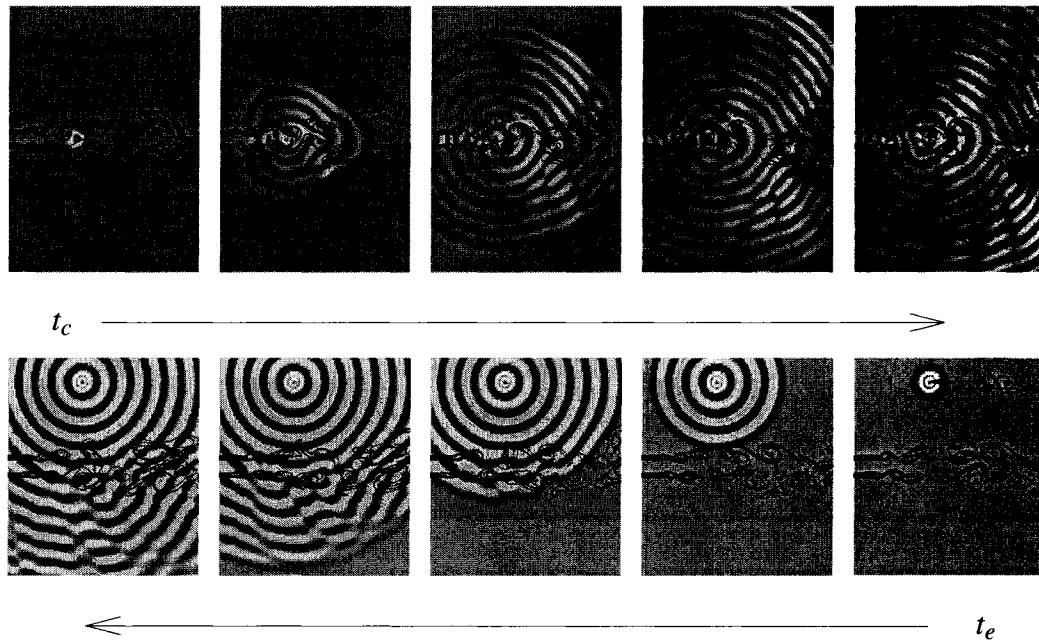


Figure 2.2: Perturbation analysis (top, representation of perturbation pressure) characterizes the effect on the entire flow resulting from a small change to a particular control quantity, taken here to be a sinusoidally-varying mass source at point x_c . Adjoint analysis (bottom, representation of adjoint density) characterizes the effect on a particular flow quantity, taken here to be high frequency noise at point x_e , due to small changes in the control applied anywhere in the flow. Vorticity contours are superimposed to the colored perturbation and adjoint fields.

ation at every point in space. The adjoint field is driven by the sound waves in the box and propagates away from it, illustrating possible locations for antinoise sources where additional forcing could be applied to achieve the desired effect (namely, to reduce the sound pressure level within the box). Even though the governing system represented here is a linear, constant-coefficient PDE and the cost function is quadratic in the state variables, the adjoint field identifies a range of effective “antinoise” forcing locations, and does not accurately identify the isolated sound source. Thus, identification of sound sources is not to be expected from adjoint analyses when applied to more complex systems, such as the unsteady jet considered in the present work.

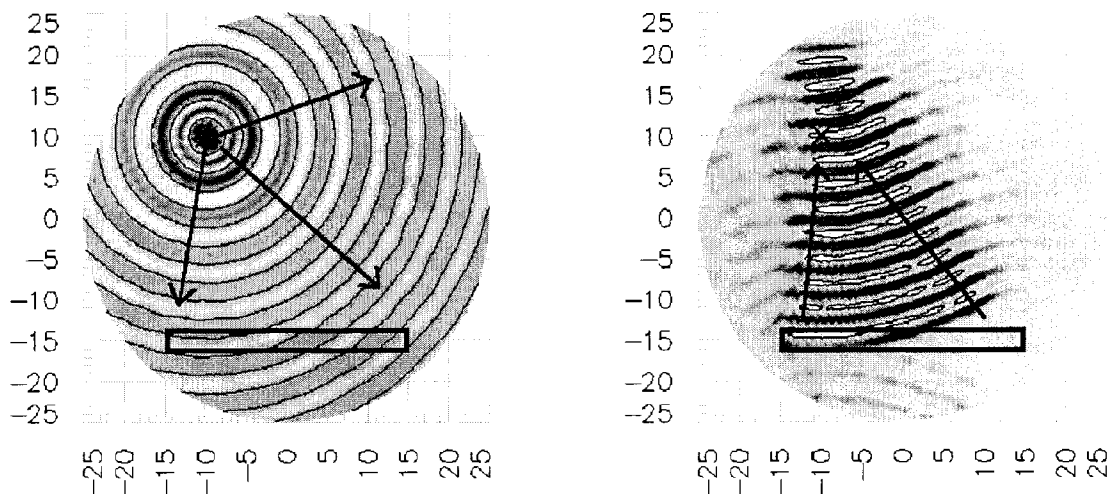


Figure 2.3: Adjoint analysis (right) of sound waves (left) produced by a monopole sound source at the point marked by the X, in a stationary fluid. The adjoint field is forced by the acoustic energy measured in the rectangular area.

Note that, in the remainder of the present analysis, the cost functions considered are essentially pointwise measurements of the sound field, and the adjoint field computations are therefore referred to as adjoint Green's functions.

2.3 Description of the system and simulation technique

The system under consideration is a Mach 0.5, 2D jet at a Reynolds number $Re_D = \rho_j D_j U_j / \mu_j = 5000$ (where the subscript j refers to the jet, U indicates velocity, D diameter, ρ density, and μ viscosity). The jet is sinusoidally excited near the jet exit at a Strouhal number $S_t = f_0 D_j / U_j = 0.4$ (f_0 is the frequency of excitation). The jet is not heated, being its temperature the same as the ambient temperature T_∞ . Refraction effects are expected to be significantly weaker in a cold jet than in a hot jet, as, in the former, the speed of sound is identical in the ambient fluid and the jet core. In fact, in sharp contrast with the perturbation and adjoint analyses of the mean of a heated jet, as considered by Tam & Auriault (1998), the corresponding analyses of the refraction due to the mean of the cold jet flow studied here exhibit very little refraction. Nevertheless, as shown in this paper, the acoustic scattering due to the unsteady vortex roll-up in the present flow

is quite pronounced even in this cold jet system, illustrating significant opportunities to control the hydrodynamic field (at low frequencies) in order to modify the high-frequency radiated noise.

This system is governed by the full nonlinear compressible Navier-Stokes equation

$$\mathcal{N}(\mathbf{q}) = \begin{pmatrix} \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{m} \\ \frac{\partial \mathbf{m}}{\partial t} + \nabla \cdot \frac{\mathbf{m} \otimes \mathbf{m}}{\rho} + \nabla p - \frac{1}{Re} \nabla \cdot \left(\mu \nabla \frac{\mathbf{m}}{\rho} \right) - \frac{1}{Re} \nabla \left[\mu \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \nabla \cdot \frac{\mathbf{m}}{\rho} \right] \\ \frac{\partial p}{\partial t} + \nabla \cdot \frac{p \mathbf{m}}{\rho} + (\gamma - 1) p \left(\nabla \cdot \frac{\mathbf{m}}{\rho} \right) - \frac{\gamma}{Re Pr} \nabla \cdot \left(\mu \nabla \frac{p}{\rho} \right) - \Phi \end{pmatrix} = \mathbf{g}. \quad (2.1)$$

where

$$\mathbf{q} = \begin{pmatrix} p \\ \rho \mathbf{u} \\ \rho \end{pmatrix} = \begin{pmatrix} p \\ \mathbf{m} \\ \rho \end{pmatrix} \quad (2.2)$$

is referred to as the state field. The operator $\mathcal{N}(\mathbf{q})$ represents the Navier-Stokes equation for an ideal gas with constant specific heats c_p and c_v , and constant Prandtl number Pr (appropriate expressions based on Sutherland's law have been given to the viscosities μ and μ_B). Φ is the irreversible viscous dissipation, and $\mathbf{g}$ is the control, here introduced as a right-hand-side forcing term in the governing equation.

It is known that the physics of a two-dimensional strong jet, such as the one under study, are different from a three-dimensional jet, as shown in Stanley & Sarkar (1997). In the two-dimensional strong jet there is a breakdown of the vortex street due to the interaction between vortices, product of the two-dimensionality of the flow, which does not occur in three-dimensional jet (neither in the two-dimensional weak jet, where the effects of vortex interactions are small compared with the strong convection downstream). This breakdown of vortices can be seen in Fig. 2.4. However, the results obtained in this paper concerning controllability of noise due to interaction between the hydrodynamic field and the acoustic field are expected to extend to three dimensions.

Even though there has been an extensive research on acoustics, there is still much work to be done in the field. In the last years, the most important advancements are

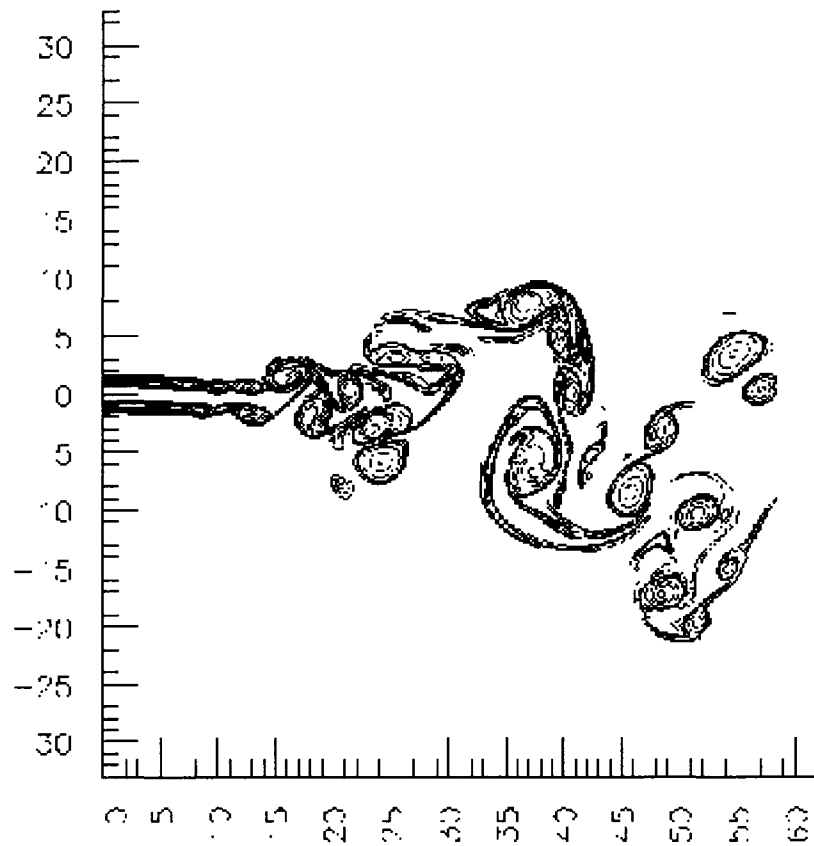


Figure 2.4: Visualization of the vortex breakdown in the two-dimensional strong jet.

related to the achievement of good numerical methods that allow us to simulate and obtain accurate far-field noise (DNS simulations are performed, for example, in Colonius *et al.* (1997), Mitchell *et al.* (1999), Suzuki & Lele (1999), Whitmire & Sarkar (2000), and Freund (2001)). The numerical method used in the present work for the direct numerical simulation of the flow follows closely that in Freund *et al.* (1997). The spatial derivatives are computed via sixth-order compact Padé approximations (which benefit from spectral-like resolution, as shown in Lele (1992)), while the integration in time is performed with a fourth-order Runge-Kutta scheme. A fourth-order compact numerical filter is utilized every few time-steps to damp any high wave-number oscillations.

2.3.1 Boundary conditions

The present simulations do not resolve any solid boundaries. Instead, artificial “buffer zones” have been placed around the domain of physical interest, coupled with characteristic based boundary conditions (in order to minimize reflections, as in Thompson (1987), Thompson (1990), once we obtain the characteristic equations of the system, the incoming waves are set to 0) on the computational boundaries. This type of *ad hoc* but effective numerical boundary conditions simulates the effect of quiescent far-field boundary conditions on the physical system, and has now become standard for this type of problem. It is further discussed in Colonius *et al.* (1993) and Freund (1997). With the non-reflecting boundary conditions, the viscous terms are considered as a regular forcing in the right hand side of the equation.

As mentioned buffer zones have been added around the physical domain, in order to drive the flow toward far-field conditions, or in order to impose the inflow profile (note that in the 2D case, entrainment has to be considered at the boundaries). In these regions, an extra forcing term is introduced on the right-hand side of the equations:

$$\mathcal{N}(\mathbf{q}) = -\sigma(\mathbf{q} - \mathbf{q}_{target}).$$

In order to obtain the inflow profile, the same kind of forcing is applied to drive the conditions to the desired values. The target axial velocity is a typical hyperbolic tangent profile:

$$v_{x,target} = U_j \frac{1}{2} \left(1 - \tanh \left[12.5 \left(\frac{r}{r_0} - \frac{r_0}{r} \right) \right] \right),$$

where r is the vertical direction.

One of the questions that arise when simulating the adjoint field is which boundary conditions are appropriate. For the flow field, the physical domain is unbounded, and the boundary conditions are only an artificial tool to simulate the unbounded domain. The information goes out the domain, and no information comes in. The adjoint field can be understood and treated in the same manner: an infinite physical domain, and non-reflective boundary conditions to simulate the far-field conditions, with a buffer zone to diminish the perturbations. If we take the adjoint of this right-hand-side term, we obtain,

in the right hand side of the adjoint field:

$$\mathcal{N}'(\mathbf{q})^* \mathbf{q}^* = -\sigma \mathbf{q}^*,$$

Note also that the adjoint of a buffer zone as the one given by Eq. (2.3.1) is a buffer zone as the one in Eq. (2.3.1).

2.4 Derivation of the adjoint system

As previously explained, the adjoint field will be used in order to obtain the sensitivity of the far-field noise to changes in the control actuation. In the development of the adjoint solver, certain additional approximations have been made, namely that the viscosity μ and the bulk viscosity of the flow μ_B are constant, and that the irreversible viscous dissipation in the energy equation is 0. These convenient simplifications are thought to be acceptable in the approximate adjoint analysis, as spatial and temporal variations of viscosity in the system and the irreversible viscous dissipation in the heat equation both effect the dynamics of the system only at the small length scales, and are thus thought to be relatively unimportant in terms of the mechanics of sound generation. Subject to these additional assumptions, and following the established procedure for performing an adjoint analysis (see, e.g., appendix B of Bewley *et al.* (2001) for the case of an unsteady compressible Euler system), we may take the Fréchet derivative of the governing equation (described in full in Cerviño *et al.* (2002)) to obtain a linearized equation of the form

$$\mathcal{N}'(\mathbf{q}) \mathbf{q}' = \mathbf{g}', \quad (2.3)$$

where $\mathbf{q}' = (p', \mathbf{m}', \rho')^T$ is referred to as the perturbation vector. Selecting an L_2 duality pairing of the form $\langle \mathbf{q}^*, \mathbf{q}' \rangle \triangleq \int_0^T \int_{\Omega} \mathbf{q}^* \cdot \mathbf{q}' d\mathbf{x} dt$, the linearized operator in Eq. (2.3) is then transformed according to the identity

$$\langle \mathbf{q}^*, \mathcal{N}'(\mathbf{q}) \mathbf{q}' \rangle = \langle \mathcal{N}'(\mathbf{q})^* \mathbf{q}^*, \mathbf{q}' \rangle + b, \quad (2.4)$$

where $\mathbf{q}^* = (p^*, \mathbf{m}^*, \rho^*)^T$ is referred to as the adjoint vector. Even though the L_2 norm has been selected here, the L_2 duality pairing is not necessarily the best choice for defining the adjoint operator in multiscale PDE systems such as the present, and incorporating

spatial or temporal derivatives into this pairing is recognized to have an important regularizing effect on the spectra of the resulting adjoint field that must be calculated (for further discussion of this important topic, see Protas *et al.* (2004)). After some algebra involving several integrations by parts, it is straightforward to obtain the adjoint operator corresponding to the approximate linearized form of the compressible Navier-Stokes equation in this framework:

$$\mathcal{N}'(\mathbf{q})^* \mathbf{q}^* = \begin{pmatrix} -\frac{\partial \rho^*}{\partial t} - \frac{\mathbf{m}}{\rho} \cdot \nabla \rho^* + (\gamma - 1) \rho^* \nabla \cdot \frac{\mathbf{m}}{\rho} - \nabla \cdot \mathbf{m}^* - \frac{\gamma \mu}{\rho Pr Re} \nabla^2 \rho^* \\ -\frac{\partial \mathbf{m}^*}{\partial t} - \frac{\gamma p}{\rho} \nabla \rho^* - \frac{(\gamma - 1) \rho^*}{\rho} \nabla p - \frac{\mathbf{m}}{\rho} \cdot (\nabla \otimes \mathbf{m}^* + (\nabla \otimes \mathbf{m}^*)^T) - \nabla p^* - \\ - \frac{\mu}{Re \rho} \left[\nabla^2 \mathbf{m}^* + \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \nabla (\nabla \cdot \mathbf{m}^*) \right] \\ -\frac{\partial p^*}{\partial t} + \frac{p \mathbf{m}}{\rho^2} \cdot \nabla \rho^* + \frac{(\gamma - 1) \mathbf{m}}{\rho^2} \cdot \nabla (\rho^* p) + \frac{\mathbf{m}}{\rho} \cdot \left(\frac{\mathbf{m}}{\rho} \cdot \nabla \right) \mathbf{m}^* + \\ + \frac{\mu}{Re \rho^2} \left[\mathbf{m} \cdot \nabla^2 \mathbf{m}^* + \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) (\mathbf{m} \cdot \nabla) (\nabla \cdot \mathbf{m}^*) \right] + \frac{\gamma \mu}{\rho Pr Re} \frac{p}{\rho} \nabla^2 \rho^* \end{pmatrix}. \quad (2.5)$$

It is important to note that, in the present derivation, we have associated the “adjoint pressure” p^* with additional forcing of the continuity equation $\partial \rho / \partial t$, and the “adjoint density” ρ^* with additional forcing of the selected form of the energy equation $\partial p / \partial t$ (note that this is in contrast with the nomenclature selected by Tam & Auriault (1998)). This is done, in part, so that the compressible adjoint equations reduce to the corresponding incompressible adjoint equations (see Bewley *et al.* (2001)) in the incompressible limit.

In a domain enclosed by solid boundaries, by selecting the appropriate adjoint boundary and initial conditions, we can make the boundary term b in Eq. (2.4), which results from the several integrations by parts, equal to zero. Alternatively, as in the present analysis, we may surround the physical part of the domain of interest in both the flow and adjoint problems with the numerical equivalent of quiescent far-field boundary conditions which propagate no information toward the physical domain of interest; this approach again effectively allows us to neglect the influence of b . By so doing, the adjoint identity Eq. (2.4) then reveals that the following two analyses are equivalent:

#1) analyzing the effect on $q_i'(\mathbf{x}_e, t_e)$ (that is, the effect on the i 'th component of the

perturbation field at point $\mathbf{x} = \mathbf{x}_e$ and time $t = t_e$) created by applying a localized force $g'_j = \delta(\mathbf{x} - \mathbf{x}_c)\delta(t - t_c)$ to the j 'th component of the perturbation equation, and

#2) analyzing the effect on $q'_i(\mathbf{x}_e, t_e)$ created by applying a localized force $g_i^* = \delta(\mathbf{x} - \mathbf{x}_e)\delta(t - t_e)$ to the i 'th component of the adjoint equation.

By the identity Eq. (2.4), we may relate the perturbation and adjoint fields in these two analyses, as indicated in Fig. 2.5, by

$$q'_i(\mathbf{x}_e, t_e) = q_j^*(\mathbf{x}_c, t_c). \quad (2.6)$$

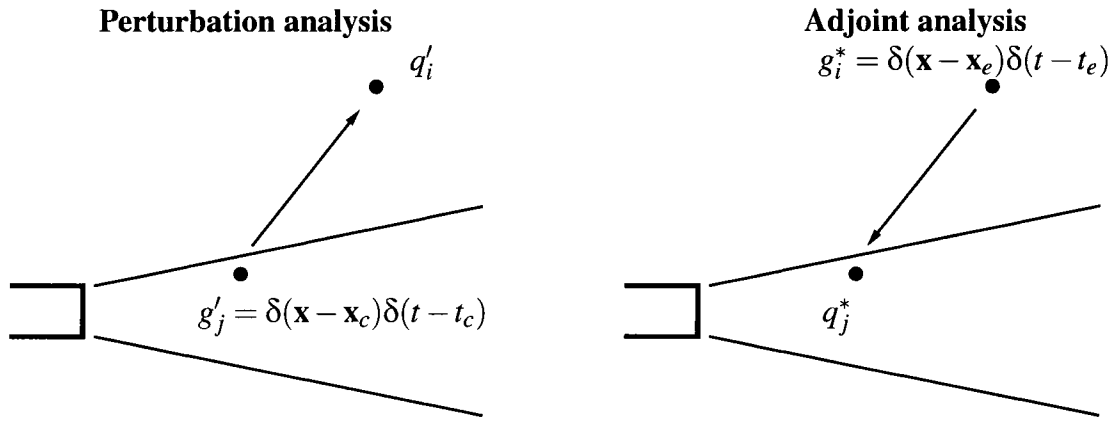


Figure 2.5: Equivalence of a perturbation (left) and adjoint (right) problems when the boundary terms in the adjoint identity are 0.

Note that the point $\mathbf{x}_c$ and time t_c do not appear in the formulation of the adjoint system in problem #2, but arise only in the subsequent analysis of the resulting adjoint field. Thus, a *single* adjoint calculation allows us to quantify the effect of forcing *anywhere* in the flow system (for any $\mathbf{x}_c$, t_c , and j) on the particular flow quantity $q'_i(\mathbf{x}_e, t_e)$. This relation between the perturbation and adjoint Green's functions provides an alternative but equivalent explanation of the significance of adjoint analyses to the one provided in Fig. 2.1.

2.5 Adjoint analysis

2.5.1 Calculation of an adjoint Green's function

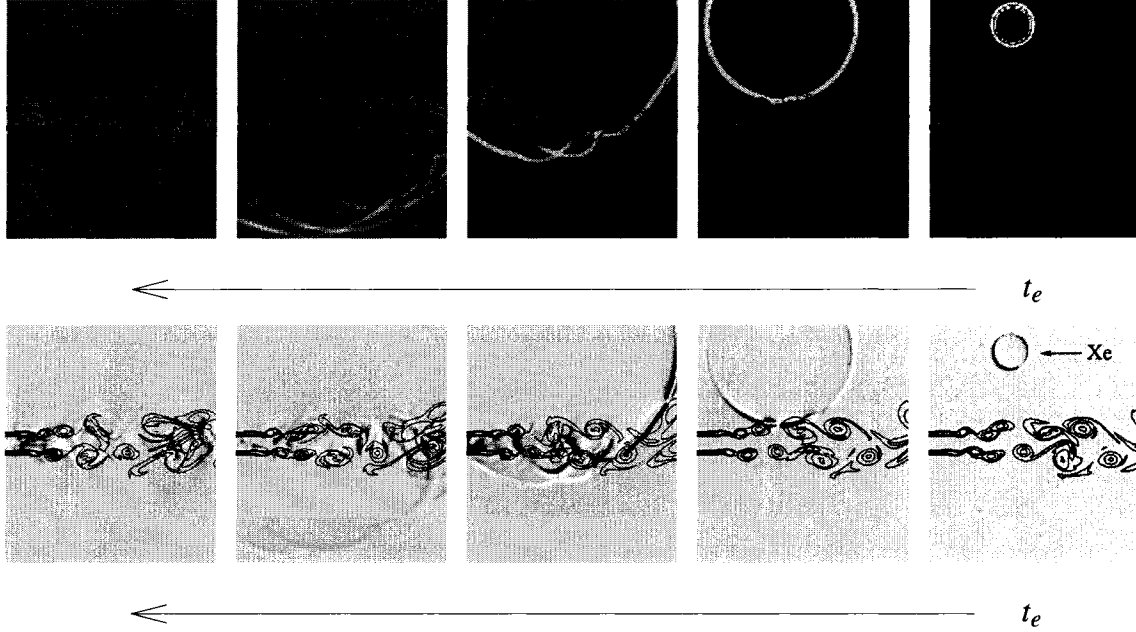


Figure 2.6: Adjoint density (top) and adjoint pressure (bottom) reveals sensitivity of the pressure component of the perturbation field at point x_e at time t_e to additional control forcing in, respectively, the energy equation (top) and the continuity equation (bottom) everywhere in space x_c and for all times $t_c < t_e$. Note that, by causality, the adjoint field is zero for $t_c > t_e$; that is, the adjoint field marches backward in time from $t = t_e$.

Fig. 2.6 illustrates a computation of the adjoint Green's function, as formulated at the end of the previous section, obtained by forcing the adjoint system $\mathcal{N}'(\mathbf{q})^* \mathbf{q}^* = \mathbf{g}^*$ with an isolated force at a particular point in space and time, that is, $g_i^* = \delta(\mathbf{x} - \mathbf{x}_e)\delta(t - t_e)$. As discussed above, each component j of the resulting adjoint Green's function, at each point in space $\mathbf{x}_c$ and each instant in time t_c , may be interpreted as the i 'th component of the perturbation to the flow at point $\mathbf{x}_e$ and time t_e that would arise due to localized forcing of the corresponding component j of the flow system at the corresponding point in space $\mathbf{x}_c$ and time t_c . The calculation reported in Fig. 2.6 takes $i = 1$, that is, the adjoint

field shown characterizes the effect on the perturbation pressure $p'(x_e, t_e)$.

It is interesting to note that the disturbance in the adjoint pressure grows rapidly as it propagates within the jet toward the nozzle at the convective velocity when the adjoint field evolves (backward in time). In contrast, the disturbance in the adjoint density essentially propagates right through the jet, experiencing significant refraction. This behavior is further quantified in Figures 2.7, 2.8, 2.9, and 2.10. The component of the adjoint density that propagates at the convective speed of the jet within the jet shear layers is found to be quite small. This result indicates, as one might expect, that mass sources are more efficient than energy sources in modifying the hydrodynamic field in a way which changes the radiated noise.

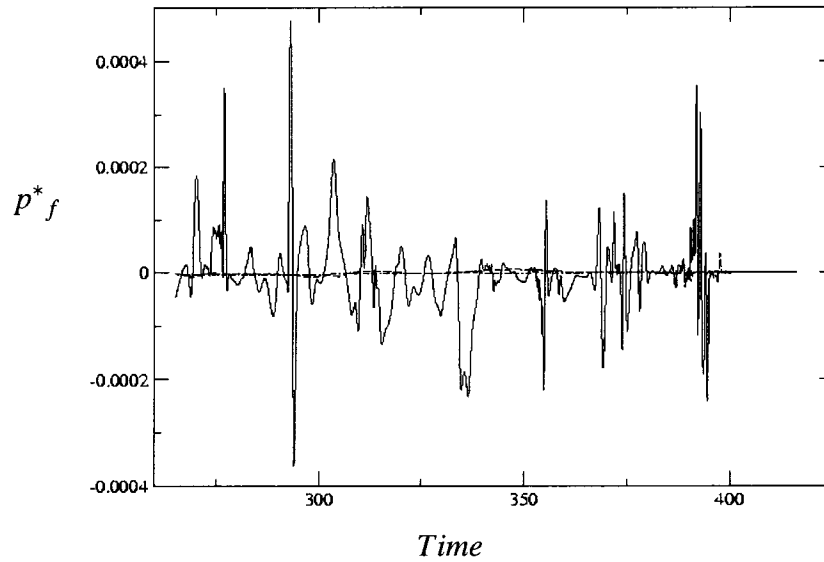


Figure 2.7: Evolution of adjoint pressure in time at the points $\{x, y\}$ of (solid) $\{5D, 0\}$, (dashed) $\{5D, 2.5D\}$, (dot-dashed) $\{5D, -2.5D\}$.

2.5.2 An adjoint Green's function at temporal frequency f corresponding to far-field noise

An alternative to forcing the adjoint problem at an isolated time t_e is to force it at a specific temporal frequency f . This corresponds roughly to looking at the sensitivity

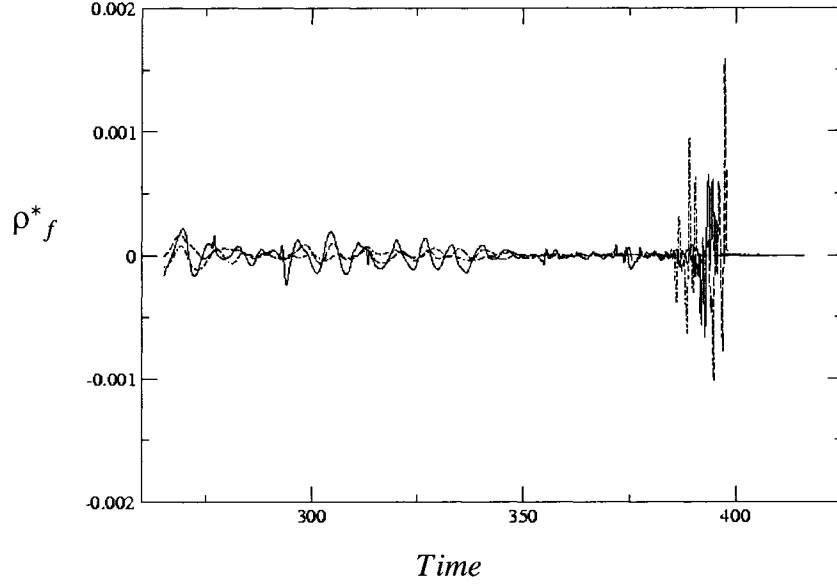


Figure 2.8: Evolution of adjoint density in time at the points $\{x, y\}$ of (solid) $\{5D, 0\}$, (dashed) $\{5D, 2.5D\}$, (dot-dashed) $\{5D, -2.5D\}$.

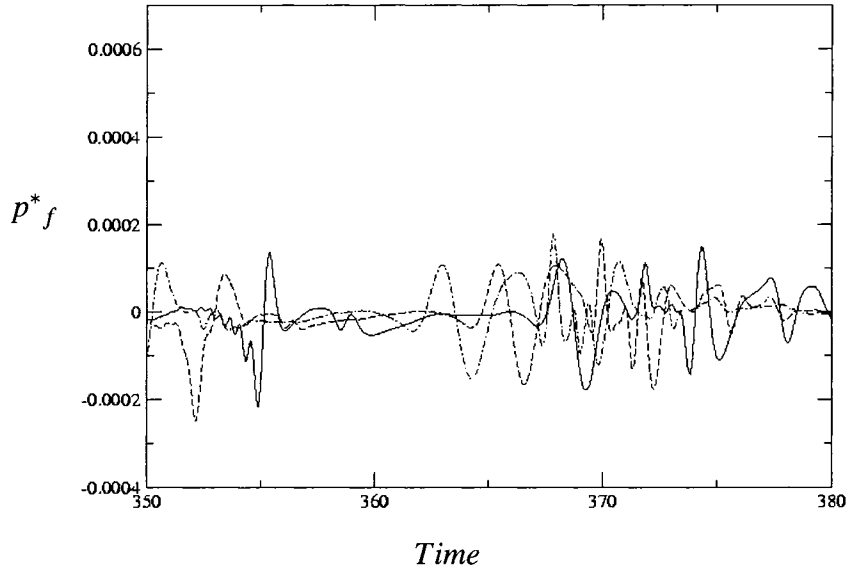


Figure 2.9: Evolution of the adjoint pressure at three different locations at the centerline: at (dot-dashed) $x = 8$, (dashed) $x = 9$, and (solid) $x = 10$.

of the sound field at point x_e (at the frequency and phase selected) to additional control forcing in the governing equations. This correspondence is only approximate, however, as the system under consideration has time-varying coefficients, and therefore frequency-

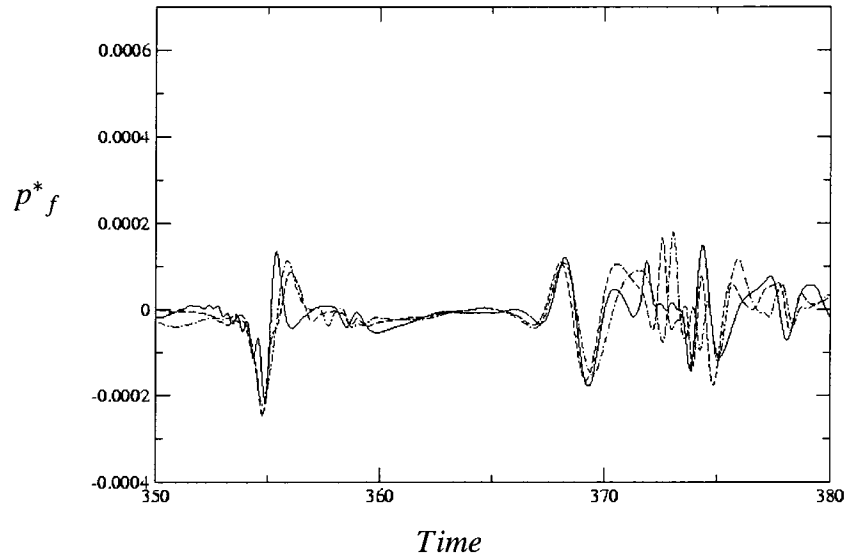


Figure 2.10: Evolution of the adjoint pressure at three different locations at the centerline: at (dot-dashed) $x = 8$, (dashed) $x = 9$, and (solid) $x = 10$ shifted the time corresponding to the convection velocity. There is an approximate superposition of the three lines, which indicates that these perturbations convect toward the nozzle at the convective speed of the jet.

based characterizations of the system's response are of limited usefulness. Note that, in systems with constant coefficients, a Bode plot completely characterizes the frequency response of the system. Such a frequency-domain analysis may only be applied to the case where only the mean flow is considered. Nonetheless, an approximate characterization of this sort may still be developed for the present system (in the time domain) simply by forcing the adjoint system sinusoidally at the frequency of interest during the backward march for the adjoint field. The computation corresponding to this kind of forcing is shown in the bottom row of Fig. 2.2.

Instead of forcing the adjoint problem at an isolated point in the computational domain x_e we can force it along a line near the boundary of the computational domain (that is, in the “buffer zone” used to approximate the far-field boundary conditions). By so doing, one may set up a propagating wave in the adjoint field which is the same as if the computational domain extended deep into the far field and the adjoint problem was forced

at an isolated point a very long distance away. By varying the forcing along this line sinusoidally, one may simulate the arrival of a wave in the adjoint field corresponding to the effect on the far-field noise in any direction of interest. A representative example is given in Fig. 2.11. Note that both reflection and refraction of the adjoint field are observed in this computation.

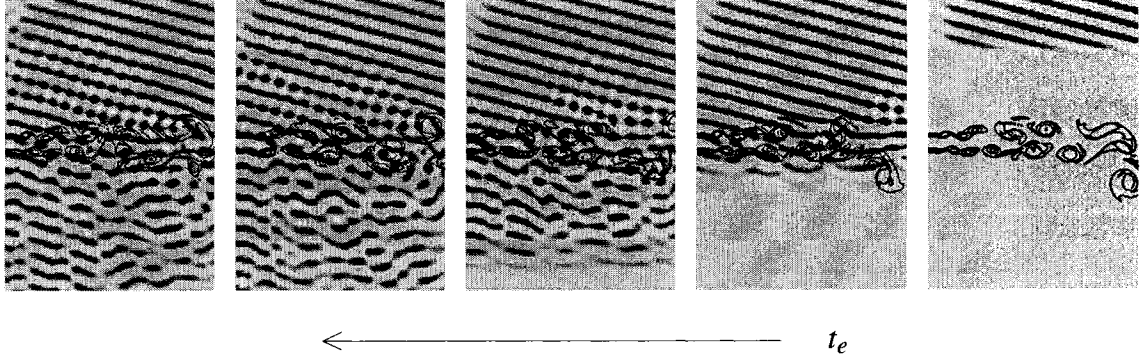


Figure 2.11: Adjoint density field due to incoming waves from the far field.

2.5.3 Quantification of scattering of adjoint Green's functions

In an attempt to quantify the scattering of a wave in the adjoint field due to the unsteady vortex roll-up, the values of the adjoint density and adjoint pressure have been measured at three different points in the representative adjoint Green's function analysis illustrated in Fig. 2.12. The points where the adjoint density and adjoint pressure were measured are above the jet (where the scattering will be referred to as reflection), at the centerline, and below the jet (where the scattering will be referred to as refraction). The time series of these measurements were Fourier-transformed in time, and the results are plotted in Fig. 2.13. The analysis was performed for adjoint forcing at two different Strouhal numbers: $St = 0.8$ ($2\times$ the vortex roll-up frequency), $St = 2.0$ ($5\times$ the vortex roll-up frequency), and $St = 8.0$ ($20\times$ the vortex roll-up frequency).

Perhaps the most important observation to make in Fig. 2.13 is that there is very significant frequency broadening in all of the measured adjoint spectra. The adjoint systems are excited by forcing at the single frequency indicated ($St = 0.8, 2.0$, or 8.0) but, due to the time-varying coefficients (from the unsteady flow field $\mathbf{q}$) in the adjoint operator, the

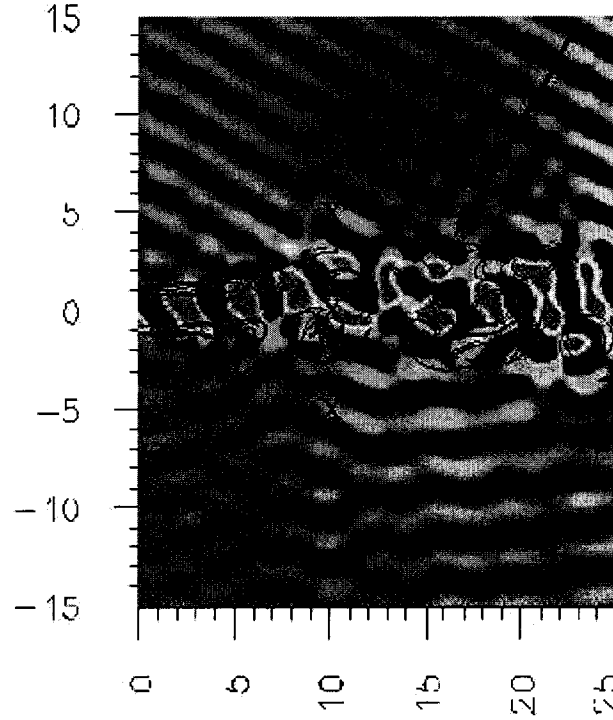


Figure 2.12: Adjoint pressure wave corresponding to far-field noise at an angle of 60° off the jet axis and at a frequency of $St = 2.0$.

measurements of the adjoint field at the indicated points exhibit energy over a broad range of temporal frequencies. For comparison, the spectra of the hydrodynamic fluctuations of the base flow is shown in Fig. 2.14. Note that the frequency broadening of the adjoint field cannot be captured by a steady-flow analysis.

The frequency broadening present when the adjoint field is forced at a higher frequency is much larger than when it is forced at a lower frequency. This fact was noticed by Suzuki (2001) for the direct problem, and was described as “multiple scattering”. In the present adjoint analysis, this suggests that high-frequency noise may be modified by a broad range of possible forcing frequencies.

Note in particular that the frequency spectrum is generally narrower at the point above the jet (dashed lines) than below the jet (dot-dashed lines), apparently because the refraction of the traveling wave in the adjoint field is stronger than the reflection of this wave for the incidence angle tested. Within the jet (solid line), it is observed that the

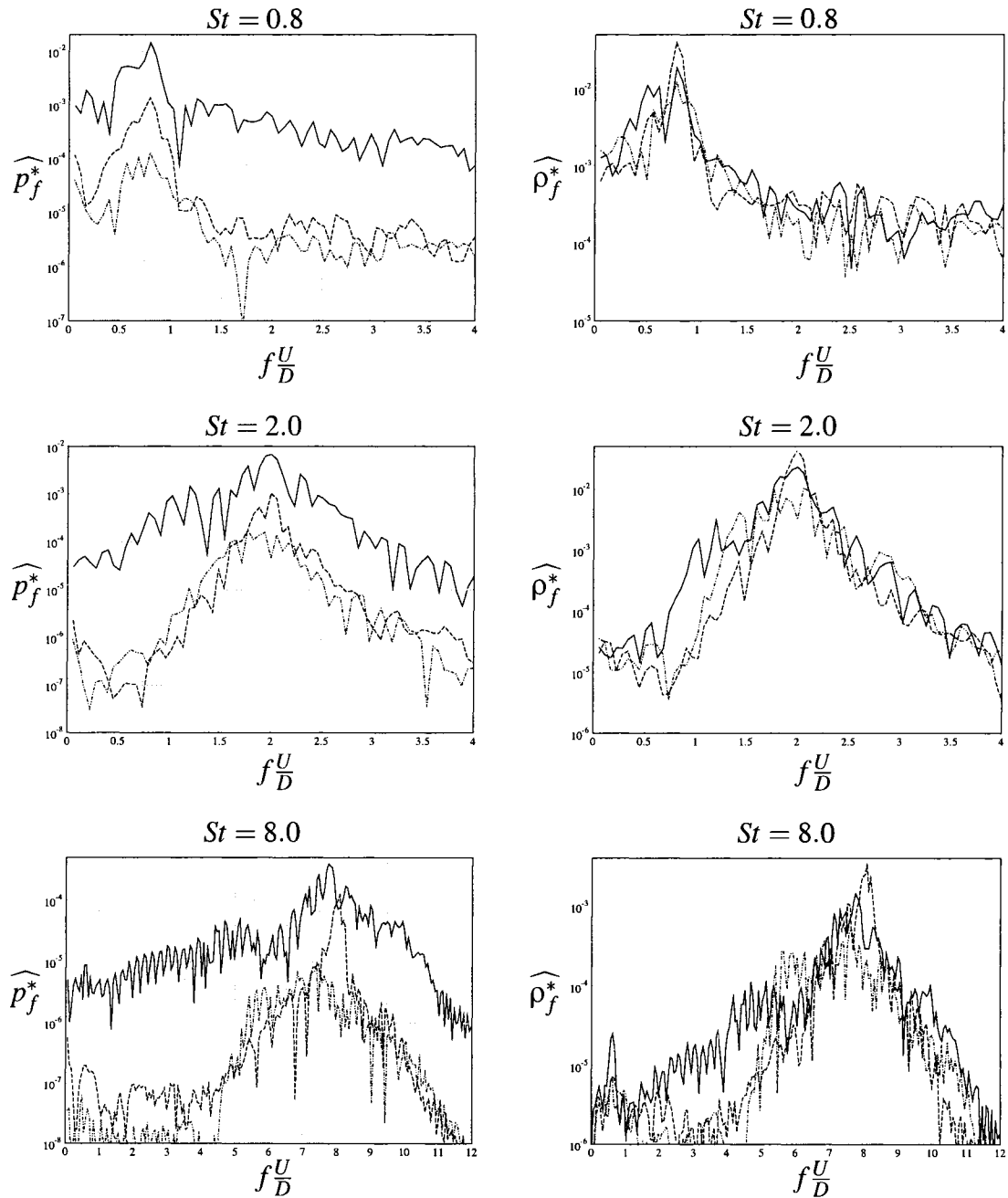


Figure 2.13: Temporal spectra measured at the indicated points $\{x, y\}$ of (solid) $\{5D, 0\}$, (dashed) $\{5D, 2.5D\}$, (dot-dashed) $\{5D, -2.5D\}$ of (left) the adjoint pressure $\widehat{p}_f^*$ and (right) the adjoint density $\widehat{\rho}_f^*$ of incoming waves at the same angle and at a frequency of (top) $St = 0.8$, (center) $St = 2.0$, and (bottom) $St = 8.0$.

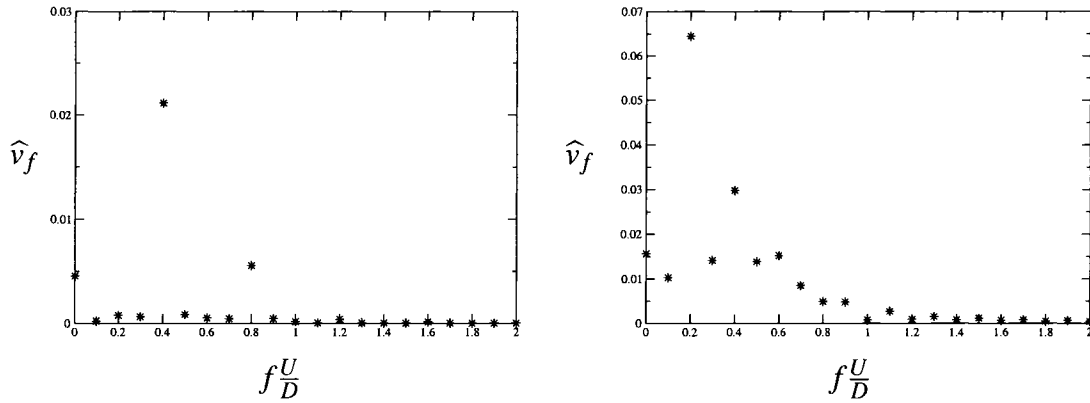


Figure 2.14: Spectra of the vertical velocity of the flow at $\{x,y\} = \{1.5D, 0.5D\}$ (left) and $\{x,y\} = \{3D, 0.5D\}$ (right). The appearance of the peak at $S_t = 0.2$ at the downstream station (right) is a result of vortex pairing (recall that the jet is forced at $S_t = 0.4$).

frequency broadening is strongest.

The low-wavenumber components of the spectra of the adjoint pressure at the centerline are especially strong for all three forcing frequencies tested. This indicates that low-frequency modulation of the hydrodynamic field via mass sources within the jet can have a significant impact on the high-frequency noise in the far field, and provides impetus for further studies in jet-noise control based on such characterizations to exploit this sensitivity.

Note also that all of the spectra are somewhat jagged, and the distance between each small peak in this jaggedness is $\Delta f = 0.2D/U$, which is exactly half of the vortex roll-up frequency. This appears to indicate (as one might expect) that the scattering of the wave in the adjoint field is closely related to its interactions with the large-scale vortex roll-up.

A second set of cases was also run in which the wave in the adjoint field approaches the jet at a 90° angle off the jet axis (cf. Fig. 2.12). The results showed very similar trends, and are thus not included here.

2.6 Optimization

It has been shown in the previous section that a low-frequency modulation of the hydrodynamic field will be effective in order to modify the high-frequency acoustic field at the far field. In the present section, we develop a low-frequency control and apply it to the jet in order to gain some reduction in the noise radiated by the jet in a given direction.

As mentioned in the previous sections, a common method to optimize a high dimensional system is using a gradient-based optimization technique. These approaches consist of iterative processes which reduce gradually the cost function of interest by doing successive line minimizations in the direction of the gradient.

2.6.1 Cost function

The first step in any optimization is to select the quantity we want to optimize, that is, the cost function. In the adjoint analyses performed in the previous sections, the cost function which defined the adjoint field was a pointwise measurement at the far field of the perturbation pressure at a certain frequency and a certain phase. When interest is on performing an optimization of the noise radiated by a jet, one of the possibilities for the cost function is a measurement of pressure fluctuations within an area of interest. These pressure fluctuations are the base of the cost function chosen here

$$\mathcal{J}(\phi) = \frac{1}{2} \int_{T_0}^T \int_{\Omega} H(\mathbf{x})(p - p_{target})^2 dx dt, \quad (2.7)$$

where $H(\mathbf{x})$ defines the interrogation area (it is valued 1 inside the area of interest Ω_1 and 0 outside). A penalty on the control effort may be added to the cost function, as in

$$\mathcal{J}(\phi) = \frac{1}{2} \int_{T_0}^T \int_{\Omega} H(\mathbf{x})(p - p_{target})^2 dx dt + \frac{d^2}{2} \int_{T_0}^T \int_{\Omega} \phi^* \phi dx dt, \quad (2.8)$$

although in the present 2D optimization, it has not been included.

2.6.2 Gradient method

Once an appropriate cost function has been chosen, a method to compute the gradient of the cost function with respect to the control parameters has to be selected. Methods

such as Finite Differences or Complex Step Derivative (described in more detail in Chapter 3) perform one computation per control parameter, being suitable in problems where the dimensionality of the controller is low, but resulting in an expensive method for high dimensional systems. Adjoint methods, on the other hand, require only one flow calculation and one adjoint computation regardless of the number of control parameters (the adjoint operator has a dependency on the flow variables, making the simulation of the flow field necessary). Flow control problems such as blowing/suction in a channel flow, Bewley *et al.* (2001), or aerodynamic shape optimization, Jameson *et al.* (1998), have already successfully applied adjoint methods.

In the problem of interest, changes in the hydrodynamic field near the exit of the nozzle will be convected downstream by the jet, and will perturb the far-field acoustics. It is the aim of this research to change this flow field in order to reduce the noise in an interrogation area Ω_1 (see Fig. 2.15) via the adjoint based optimization procedure. The evolution (backward in time) of the corresponding adjoint field, which is forced in the same interrogation area, needs to be simulated. When the adjoint perturbations reach the jet, they are convected upstream, and give the required gradient information for the optimization problem.

The adjoint based method requires a good database of the flow field and its evolution in time. The analysis will be limited to low Reynolds numbers due to the inherent memory storage and time requirements of Direct Numerical Simulations (DNS).

2.6.3 Controller

Adjoint methods have been applied in high dimensional problems where the number of control parameters is such that other methods like finite differences or complex step derivative are very expensive. Bewley *et al.* (2001) apply it in a channel flow, where the control parameters were blowing/suction at the walls, being the actuation at every point on the wall independent. Wei & Freund (2002) apply a generalized forcing in a small region over a shear layer, the forcing term taking a different and independent value at every point in the control region.

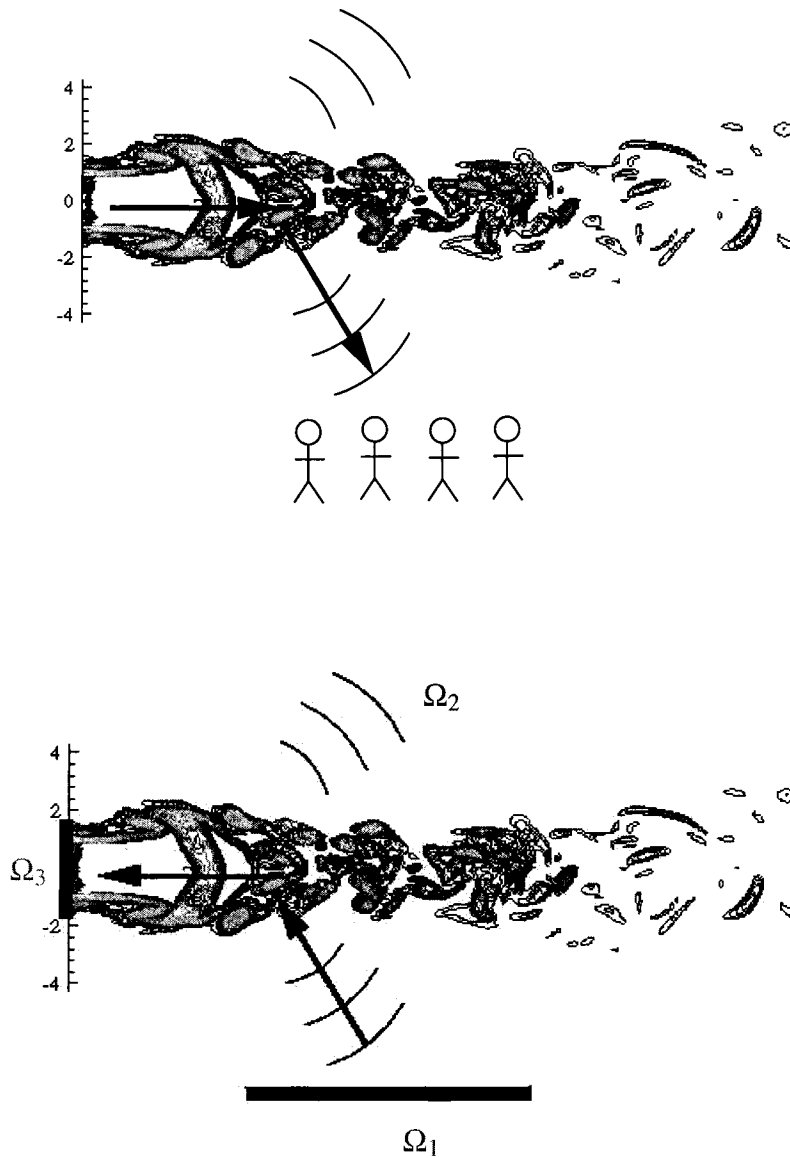


Figure 2.15: Physics of the flow (top) and adjoint (bottom) fields.

The effect of applying periodic control actuation in chaotic systems has been studied in previous works. Mirus & Sprott (1999) claims that the optimum actuation frequencies correspond to low-order rational multiples of the frequencies of the unstable periodic orbits embedded in the attractor, as opposed to other researches who claim that the most successful actuation frequencies are rational multiples of the periodic frequencies that initiated the chaos, as in Lima & Pettini (1990), the natural frequencies in a period-doubling

route in chaos, Azevedo & Rezende (1991), or frequencies corresponding to peaks in the power spectrum, S  therblom (1997). A periodic-in-time actuation has been preferred in the present work over some other approaches for applicability reasons.

As seen in previous sections, mass sources are more effective than heat sources in order to modify the acoustic field with low-amplitude modulation of the hydrodynamic field. For this reason, the control consists on a forcing term on the right-hand side of the continuity equation (first subequation of Eq. (2.1))

$$\mathcal{N}(\mathbf{q}) = S(\mathbf{x}) \cdot f_p(\theta, t), \quad (2.9)$$

where $f(\theta, t)$ is a periodic function in $t \in [0, T]$. It has been generalized to the case where there is an azimuthal component for future application to the round jet. Note that the formulation in the present section considers θ as a variable, although this dependence is ignored in the 2D setting.

The general form given to f is

$$f_p(\theta, t) = \sum_{k=1}^{M_t} \sum_{j=1}^{M_\theta} A_{kj} \sin(k\omega t + z_k) \cos((j-1)\theta + y_j), \quad (2.10)$$

where ω is a fixed frequency. M_t is the maximum number of modes in time considered, and M_θ is the maximum azimuthal modes considered in the actuation. The control parameters are the amplitudes A_{kj} , and the phases z_k and y_j .

2.6.4 Definition of the adjoint field and derivation of the gradient

The derivation of the adjoint operator is exactly the same as the derivation in Section 2.4. The simplifications assumed there are again used here for the optimization. The forcing term $\mathbf{g}$ on the right-hand side of Eq. (2.1) is substituted by the control above described. The adjoint operator is the same as in Section 2.4. However, the forcing and the cost function will define a different adjoint field, as described below.

Applying again the adjoint identity Eq. (2.4), the boundary term b is given by

$$\begin{aligned}
 b = & \int_{\Omega} (p^* \rho' + m_i^* m'_i + \rho^* p') \big|_0^T dx + \int_0^T \int_{\partial\Omega} n_j [p^* m'_j + m_j^* p' + \\
 & + m_i^* \left(\frac{m_i m'_j}{\rho} + \frac{m'_i m_j}{\rho} - \frac{\rho' m_i m_j}{\rho^2} \right) - v \left(m_i^* \frac{\partial m'_i}{\partial x_j} + m'_i \frac{\partial m_i^*}{\partial x_j} \right) + \\
 & + \rho^* \left(\frac{p' m_j}{\rho} + \frac{p m'_j}{\rho} - \frac{\rho' p m_j}{\rho^2} \right) + (\gamma - 1) \frac{p \rho^*}{\rho} (m'_j - \frac{\rho' m_j}{\rho})] dx dt.
 \end{aligned} \tag{2.11}$$

Taking into account that the initial conditions of the perturbation are known

$$\mathbf{q}'(t = 0) = 0, \tag{2.12}$$

and defining initial and far-field conditions for the adjoint field

$$\mathbf{q}^*(t = T) = 0 \tag{2.13}$$

$$\mathbf{q}^*(\mathbf{x} \rightarrow \infty) = 0, \tag{2.14}$$

the boundary term b becomes zero, as it was the case in the adjoint analysis performed in the previous sections.

The only step left in the complete definition of the adjoint field (once the operator and the initial and boundary conditions are defined) is the definition of the right-hand-side forcing of the adjoint equation. The main consideration to take into account when defining the forcing is to obtain a closed form of the gradient by substituting terms in the adjoint identity Eq. (2.4) and in the perturbation cost function

$$\mathcal{J}'(\phi; \phi') = \int_{T_0}^T \int_{\Omega} H(\mathbf{x}) p'(p - p_{target}) dx dt. \tag{2.15}$$

Defining the right-hand-side forcing of the adjoint equation as

$$\begin{aligned}
 \mathcal{J}'(\phi; \phi') &= \int_{T_0}^T \int_{\Omega} H(\mathbf{x}) p'(p - p_{target}) dx dt. \\
 \mathcal{N}'(\mathbf{q})^* \mathbf{q}^* &= \begin{pmatrix} H(\mathbf{x}) (p - p_{target}) \\ 0 \\ 0 \end{pmatrix},
 \end{aligned} \tag{2.16}$$

and substituting in the identity Eq. (2.4) in Section 2.4, a different form of the perturbation cost function given in Eq. (2.15) is obtained

$$\mathcal{J}'(\phi; \phi') = \int_{T_0}^T \int_{\Omega} S(\mathbf{x}) p^* f'_\rho(\theta, t) dx dt \tag{2.17}$$

Substituting the form of $f'_0(\theta, t)$ given in Eq. (2.10), the final expression for the perturbed cost function is

$$\begin{aligned} \mathcal{J}'(\phi; \phi') = & \sum_{k=1}^{M_t} \sum_{j=1}^{M_\theta} A'_{kj} \int_{T_0}^T \int_{\Omega} [S(\mathbf{x}) p^* \sin(k\omega t + z_k) \cos((j-1)\theta + y_j)] dx dt + \\ & + \sum_{k=1}^{M_t} \sum_{j=1}^{M_\theta} z'_k \int_{T_0}^T \int_{\Omega} [S(\mathbf{x}) p^* A_{kj} \cos(k\omega t + z_k) \cos((j-1)\theta + y_j)] dx dt - \\ & - \sum_{k=1}^{M_t} \sum_{j=1}^{M_\theta} p'_j \int_{T_0}^T \int_{\Omega} [S(\mathbf{x}) p^* A_{kj} \sin(k\omega t + z_k) \sin((j-1)\theta + y_j)] dx dt \end{aligned} \quad (2.18)$$

A closed expression for the gradients are now directly available from Eq. (2.18):

$$\begin{aligned} \frac{D\mathcal{J}}{DA_{kj}} &= \int_{T_0}^T \int_{\Omega} [S(\mathbf{x}) p^* \sin(k\omega t + z_k) \cos((j-1)\theta + y_j)] dx dt \\ \frac{D\mathcal{J}}{Dz_k} &= \int_{T_0}^T \int_{\Omega} [S(\mathbf{x}) p^* A_{kj} \cos(k\omega t + z_k) \cos((j-1)\theta + y_j)] dx dt \\ \frac{D\mathcal{J}}{Dy_j} &= \int_{T_0}^T \int_{\Omega} [S(\mathbf{x}) p^* A_{kj} \sin(k\omega t + z_k) \sin((j-1)\theta + y_j)] dx dt \end{aligned} \quad (2.19)$$

Accuracy of the gradient

Both the accuracy of the gradient and the approximations made and explained in Section 2.4 have been checked with the Complex-Step Derivative method (the result given by the adjoint method, finite differences, or Complex-Step derivative, should be the same). This method provides an accurate gradient of the system by transforming all the variables of the flow into complex variables, applying a perturbation in the complex part of the controller, and looking at the complex part of the resulting cost function. It is further described in Chapter 3. The error found in the gradient of the cost function with respect to the amplitudes A_{kj} is about 10%, which is an acceptable error in such a high-dimensional system where a long-time simulation is performed. The error in the gradient of the cost function with respect to the phase z_k is about 20%, higher than what would be desirable. One of the reasons of the high error is that the adjoint field grows exponentially backward in time, largely amplifying errors toward the end of the adjoint computation.

2.6.5 Control update

As mentioned before, the optimization procedure is an iterative process where numerous line minimizations in direction of the gradient are performed (Steepest Descent Method). An alternative which converges fastest (exactly in N iterations when the cost function is quadratic, being N the number of parameters to optimize) is the Conjugate Gradient Method. For this reason, Conjugate Gradient Method, which adds a momentum term to the gradient direction, has been used in the present optimization. Basically, the conjugate gradient method updates the control in the following manner

$$\phi^{k+1} = \phi^k + \alpha^k h^k, \quad (2.20)$$

where h^k is the update direction, and has the expression

$$h^k = - \left(\frac{Dj}{D\phi} \right)^k + \beta^k h^{k-1}, \quad (2.21)$$

where β^k is the momentum term, and $Dj/D\phi$ is the gradient. In the steepest descent method, $\beta^k = 0$. In the conjugate gradient method, the update direction is initialized as $h^1 = - \left(\frac{Dj}{D\phi} \right)^1$. Considering the Polak-Ribiere variant of the method, the expression for the momentum term is

$$\beta^k = \frac{\left[\left(\frac{Dj}{D\phi} \right)^k - \left(\frac{Dj}{D\phi} \right)^{k-1} \right] \cdot \left(\frac{Dj}{D\phi} \right)^k}{\left(\frac{Dj}{D\phi} \right)^{k-1} \cdot \left(\frac{Dj}{D\phi} \right)^{k-1}} \quad (2.22)$$

For further discussion of the conjugate gradient, see Chapter 5 of Bewley (Preprint).

2.6.6 Control computation

The adjoint simulation is coded in the same manner as the flow simulation. The flow field is stored every 25 iterations, and linearly interpolated in the adjoint simulation. Conjugate gradient (Polak-Ribiere variant) is used in order to perform the control update, and the parameter α^k is computed via Brent's method, described in Press *et al.* (1992).

2.6.7 Results

Before proceeding with the optimization, the flow field is first evolved for 600 numerical seconds in order to reach a statistically steady state and flush out of the domain all the transient structures. After this time, the control is switched on, as indicated in Eq. (2.9). However the cost function is not computed until $t > 640$, so that there is enough time for the actuation to reach the interrogation area Ω_1 (Fig. 2.15), which has been selected to be the rectangle $x \in [4, 16]$, $y \in [9.5, 10.5]$. The function $f(\theta, t)$ has been built as the summation of two modes in time. Therefore there are 4 control variables: two amplitudes (A_{10} and A_{20}), and two phases (z_1 and z_2). The frequency of the controller has been chosen to be the same as the one perturbing the jet (corresponding to a Strouhal number of 0.4). The Reynolds number based on the jet diameter is 5000. The system has been discretized using 800×1600 points (axial and vertical directions respectively), which have been uniformly distributed throughout the domain.

The convergence criterion in the optimization is that the change in cost function in two consecutive iterations is smaller than 1%. This criterion has been satisfied after 4 iterations, resulting in a reduction of 6% in the cost function.

Fig. 2.16 shows the evolution of the adjoint density in order to observe the growth of the adjoint field as it approaches the jet exit, as it has been mentioned in previous sections. The smaller value assigned to the contourlines is the same in all the frames. However, as the adjoint evolves backward in time, the maximum value increases, and more contourlines are represented. It can be seen that there is a larger concentration of contourlines at the exit of the nozzle (jet evolves from left to right in each frame), and that the number of contourlines increases as time decreases. This growth of the adjoint field will cause a stronger dependence of the gradient on the flow state within a narrow window of time, which will prevent the optimization to achieve optimal time-average statistics.

The reduction in noise that has been achieved with the present approach is 6%, which translates to 0.5dB. This reduction is lower than desired. Two main reasons are attributed to the low reduction. First, the low number of control parameters we are using. In the round jet, more parameters are going to be used (time modes and azimuthal modes).

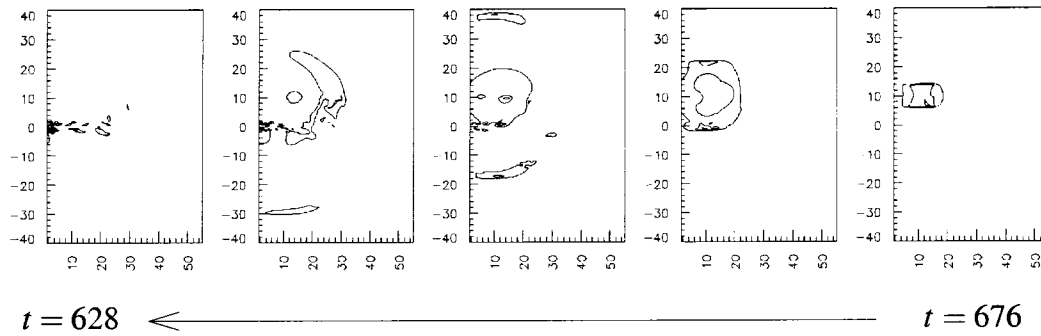


Figure 2.16: Evolution of the adjoint density.

Second, the system has been forced at a Strouhal number of 0.4. This value has been taken also in the control actuation. However, as seen by Mirus & Sprott (1999), the optimum actuation frequencies correspond to the frequency of the unstable periodic orbits embedded in the chaotic attractor. A dominant frequency in the jet is $St = 0.2$, caused by the pairing of the vortices. This frequency is not represented in the controller. Further reduction of the noise at the far-field is expected when including this frequency in the controller, or when optimizing the frequency.

2.7 Concluding remarks

An adjoint analysis in an unsteady compressible 2D jet has been performed in order to obtain insight on control opportunities in this system. Attention has been focused on the scattering of adjoint Green's functions corresponding to far-field high-frequency noise. Significant scattering of the adjoint field is detected both above and below the jet, as quantified by a spectral analysis of the adjoint field. This scattering is a direct result of system unsteadiness (vortex roll-up), and cannot be captured by mean flow analyses.

The degree to which frequency broadening extends into the low frequencies within the jet in the adjoint analyses indicates the degree to which low-frequency alteration of the hydrodynamic field can be used to effect the high-frequency radiated acoustic field. This distinguishes promising low-frequency “hydrodynamic” control strategies from simple (but perhaps impractical) “antinoise” control strategies, which must be applied at the frequency of the radiated noise.

An optimization has been performed in a $Re = 5000$, $M = 0.5$ jet. A periodic mass forcing has been used as the control, with two modes in time. A reduction in noise of 6% has been achieved on a specified region. The reduction is expected to be higher as more modes are considered, and probably when the pairing frequency of the vortices is included in the control actuation.

Appendix

Formulation of the equations in cylindrical coordinates

2.A.1 Introduction

When performing Direct Numerical Simulation of a round turbulent jet, cylindrical coordinates are preferred over cartesian coordinates in order to better resolve the shear layers. The new equations in cylindrical coordinates are laid out in the present Appendix, as well as the discussion of some numerical issues that arise when using these coordinates.

2.A.2 Governing equations

Navier-Stokes equation in cylindrical coordinates is

$$\begin{aligned}
 \frac{\partial}{\partial t}(\rho u_x) + \frac{\partial}{\partial x}(\rho u_x u_x) + \frac{1}{r} \frac{\partial}{\partial r}(r \rho u_x v_r) + \frac{1}{r} \frac{\partial}{\partial \theta}(r \rho u_x v_\theta) &= -\frac{\partial p}{\partial x} + V_x \\
 \frac{\partial}{\partial t}(\rho v_r) + \frac{\partial}{\partial x}(\rho v_r u_x) + \frac{1}{r} \frac{\partial}{\partial r}(r \rho v_r v_r) + \frac{1}{r} \frac{\partial}{\partial \theta}(r \rho v_r v_\theta) - \frac{\rho v_\theta^2}{r} &= -\frac{\partial p}{\partial r} + V_r \\
 \frac{\partial}{\partial t}(\rho v_\theta) + \frac{\partial}{\partial x}(\rho v_\theta u_x) + \frac{1}{r} \frac{\partial}{\partial r}(r \rho v_\theta v_r) + \frac{1}{r} \frac{\partial}{\partial \theta}(r \rho v_\theta v_\theta) + \frac{\rho v_r v_\theta}{r} &= -\frac{1}{r} \frac{\partial p}{\partial \theta} + V_\theta \\
 \frac{\partial e}{\partial t} + \frac{\partial}{\partial x}[u_x(e+p)] + \frac{1}{r} \frac{\partial}{\partial r}[r v_r(e+p)] + \frac{1}{r} \frac{\partial}{\partial \theta}[r v_\theta(e+p)] &= -\frac{\partial q_x}{\partial x} - \frac{1}{r} \frac{\partial r q_r}{\partial r} - \frac{1}{r} \frac{\partial q_\theta}{\partial \theta} + \\
 &\quad + u_x V_x + v_r V_r + v_\theta V_\theta + \Phi
 \end{aligned}$$

where V_x , V_r and V_θ are the viscous terms in the momentum equations, e is the total energy per unit volume, Φ is the irreversible viscous dissipation and q_x , q_r , and q_θ are the axial, radial and azimuthal heat fluxes respectively.

The viscous terms appearing in the momentum equations are:

$$\begin{aligned} V_x &= \frac{\partial \tau_{xx}}{\partial x} + \frac{1}{r} \frac{\partial r \tau_{xr}}{\partial r} + \frac{1}{r} \frac{\partial \tau_{x\theta}}{\partial \theta} \\ V_r &= \frac{\partial \tau_{xr}}{\partial x} + \frac{1}{r} \frac{\partial r \tau_{rr}}{\partial r} + \frac{1}{r} \frac{\partial \tau_{r\theta}}{\partial \theta} - \frac{1}{r} \tau_{\theta\theta} \\ V_\theta &= \frac{\partial \tau_{x\theta}}{\partial x} + \frac{1}{r} \frac{\partial r \tau_{r\theta}}{\partial r} + \frac{1}{r} \frac{\partial \tau_{\theta\theta}}{\partial \theta} + \frac{1}{r} \tau_{r\theta} \end{aligned} \quad (2.23)$$

where τ is the symmetric viscous stress tensor. The irreversible viscous dissipation is

$$\begin{aligned} \Phi &= \tau_{xx} \frac{\partial u_x}{\partial x} + \tau_{xr} \frac{\partial v_r}{\partial x} + \tau_{x\theta} \frac{\partial v_\theta}{\partial x} + \\ &+ \tau_{xr} \frac{\partial u_x}{\partial r} + \tau_{rr} \frac{\partial v_r}{\partial r} + \tau_{r\theta} \frac{\partial v_\theta}{\partial r} + \\ &+ \frac{\tau_{x\theta}}{r} \frac{\partial u_x}{\partial \theta} + \frac{\tau_{r\theta}}{r} \frac{\partial v_r}{\partial \theta} + \frac{\tau_{\theta\theta}}{r} \frac{\partial v_\theta}{\partial \theta} + \\ &+ \frac{v_r \tau_{\theta\theta}}{r} - \frac{v_\theta \tau_{r\theta}}{r}. \end{aligned} \quad (2.24)$$

The components of the stress tensor are

$$\begin{aligned} \tau_{xx} &= 2 \frac{\mu}{Re_a} \frac{\partial u_x}{\partial x} - \frac{2}{3} \frac{\mu}{Re_a} \Theta + \frac{\mu_B}{Re_a} \Theta \\ \tau_{xr} &= \frac{\mu}{Re_a} \left(\frac{\partial v_r}{\partial x} + \frac{\partial u_x}{\partial r} \right) \\ \tau_{x\theta} &= \frac{\mu}{Re_a} \left(\frac{\partial v_\theta}{\partial x} + \frac{1}{r} \frac{\partial u_x}{\partial \theta} \right) \\ \tau_{rr} &= 2 \frac{\mu}{Re_a} \frac{\partial v_r}{\partial r} - \frac{2}{3} \frac{\mu}{Re_a} \Theta + \frac{\mu_B}{Re_a} \Theta \\ \tau_{r\theta} &= \frac{\mu}{Re_a} \left(\frac{\partial v_\theta}{\partial r} + \frac{1}{r} \frac{\partial u_r}{\partial \theta} - \frac{v_\theta}{r} \right) \end{aligned} \quad (2.25)$$

where Θ is the dilatation

$$\Theta = \frac{\partial u_x}{\partial x} + \frac{1}{r} \frac{\partial r v_r}{\partial r} + \frac{1}{r} \frac{\partial v_\theta}{\partial \theta} \quad (2.26)$$

and Re_a is the Reynolds number based on the speed of sound

$$Re_a = \frac{\rho_\infty a_\infty r_0^*}{\mu_\infty} \quad (2.27)$$

This Reynolds number has been selected for its convenience to nondimensionalize the equations. It has little physical relevance, but it is related to a more pertinent Reynolds number $Re_\infty = \rho_\infty U r_0^* / \mu_\infty$ by the Mach number $M = U / a_\infty$. It is assumed that the bulk viscosity is related to the shear viscosity by a constant ratio.

The heat fluxes are given by Fourier's Law

$$\begin{aligned} q_x &= \frac{-\mu}{RePr} \frac{\partial T}{\partial x} \\ q_r &= \frac{-\mu}{RePr} \frac{\partial T}{\partial r} \\ q_\theta &= \frac{-\mu}{RePr} \frac{1}{r} \frac{\partial T}{\partial \theta} \end{aligned} \quad (2.28)$$

where the Prandtl number

$$Pr = \frac{c_p^* \mu^*}{k^*}$$

is assumed to be constant and equal to 0.7 (k^* is the thermal conductivity of the fluid).

Energy and viscosity

The total energy is the sum of the internal and kinetic energy, which, for a perfect gas (where the internal energy is related by the temperature by $e_i^* = \rho c_v^* T^*$, c_v^* being the specific heat at constant volume), can be written as follows

$$e = \frac{\rho T}{\gamma} + \frac{1}{2} \rho (u_x^2 + v_r^2 + v_\theta^2), \quad (2.29)$$

where we have used the nondimensionalized expression of the internal energy $e_i = \rho T / \gamma$. The perfect gas assumption also relates the thermodynamic pressure, density and temperature by the ideal gas equations of state $p^* = \rho^* R T^*$, where R is the dimensional gas constant. When nondimensionalized, the state equation is

$$p = \rho \frac{\gamma - 1}{\gamma} T. \quad (2.30)$$

With these relationships, we can relate the pressure to the total energy by

$$p = (\gamma - 1) \left(e - \frac{1}{2} \rho (u_x^2 + v_r^2 + v_\theta^2) \right). \quad (2.31)$$

To specify the viscous stress tensor τ we have assumed a Newtonian fluid, and so its components are linearly related to the fluid strain rates by the viscosity $\mu^* = \mu^*(T)$. We assume Sutherland's law gives the functional dependence on the temperature:

$$\frac{\mu^*}{\mu_\infty} = \left(\frac{T^*}{T_0^*} \right)^{\frac{3}{2}} \frac{T_0^* + S^*}{T^* + S^*}. \quad (2.32)$$

When computing the viscous terms, some care has to be taken in order to not forget the derivatives of the viscosity, since it is not constant.

Passive scalar transport equation

The passive scalar, a variable from which the mixing can be studied, is solved simultaneously with the flow equations.

$$\frac{\partial}{\partial t}(\rho\xi) + \frac{\partial}{\partial x}(\rho\xi u_x) + \frac{1}{r} \frac{\partial}{\partial r}(r\rho\xi v_r) + \frac{1}{r} \frac{\partial}{\partial \theta}(\rho\xi v_\theta) = -\frac{\partial G_x}{\partial x} - \frac{1}{r} \frac{\partial r G_r}{\partial r} - \frac{1}{r} \frac{\partial G_\theta}{\partial \theta}. \quad (2.33)$$

The molecular diffusion fluxes are specified by Fick's law

$$\begin{aligned} G_x &= \frac{-\mu}{ReSc} \frac{\partial \xi}{\partial x} \\ G_r &= \frac{-\mu}{ReSc} \frac{\partial \xi}{\partial r} \\ G_\theta &= \frac{-\mu}{ReSc} \frac{1}{r} \frac{\partial \xi}{\partial \theta} \end{aligned} \quad (2.34)$$

and the Schmidt number, Sc , is defined as

$$Sc = \frac{\mu^*}{\rho^* D^*},$$

where D^* is the dimensional mass diffusion coefficient. The Schmidt number is assumed to be unity, which is appropriate for gas phase flow.

Adjoint field

With the same assumptions as mentioned in Section 2.4, the resulting adjoint operator is

$$\mathcal{N}'(\mathbf{q})^* \mathbf{q}^* = \left(\begin{array}{l} -\frac{\partial \rho^*}{\partial t} - \frac{\mathbf{m}}{\rho} \cdot \nabla \rho^* + (\gamma - 1) \rho^* \nabla \cdot \frac{\mathbf{m}}{\rho} - \nabla \cdot \mathbf{m}^* - \frac{\gamma \mu}{\rho Pr Re} \nabla^2 \rho^* \\ -\frac{\partial \mathbf{m}^*}{\partial t} - \frac{\gamma p}{\rho} \nabla \rho^* - \frac{(\gamma - 1) \rho^*}{\rho} \nabla p - \frac{\mathbf{m}}{\rho} \cdot \left(\nabla \otimes \mathbf{m}^* + (\nabla \otimes \mathbf{m}^*)^T \right) - \nabla p^* - \\ - \frac{\mu}{Re \rho} \left[\nabla^2 \mathbf{m}^* + \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \nabla (\nabla \cdot \mathbf{m}^*) \right] - \frac{\phi}{\rho} \nabla \phi^* \\ -\frac{\partial p^*}{\partial t} + \frac{p \mathbf{m}}{\rho^2} \cdot \nabla \rho^* + \frac{(\gamma - 1) \mathbf{m}}{\rho^2} \cdot \nabla (\rho^* p) + \frac{\mathbf{m}}{\rho} \cdot \left(\frac{\mathbf{m}}{\rho} \cdot \nabla \right) \mathbf{m}^* + \\ + \frac{\mu}{Re \rho^2} \left[\mathbf{m} \cdot \nabla^2 \mathbf{m}^* + \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) (\mathbf{m} \cdot \nabla) (\nabla \cdot \mathbf{m}^*) \right] + \\ + \frac{\phi \mathbf{m}}{\rho^2} \cdot \nabla \phi^* + \frac{\mu}{Re Sc \rho^2} \nabla^2 \phi^* + \frac{\gamma \mu}{\rho Pr Re \rho} p \nabla^2 \rho^* \\ -\frac{\partial \phi^*}{\partial t} - \frac{\mathbf{m}}{\rho} \cdot \nabla \phi^* - \frac{\mu}{Re Sc \rho} \nabla^2 \phi^* \end{array} \right). \quad (2.35)$$

where $\phi = \rho \xi$. For the jet problem the equation Eq. (2.35) must be expressed in cylindrical coordinates. The equations then acquire a complicated form. They are given in below.

The first equation of the adjoint field yields (*continuity equation*)

$$\begin{aligned} -\frac{\partial \rho^*}{\partial t} - \frac{1}{r} \frac{\partial m_r^*}{\partial r} - \frac{1}{r} \frac{\partial m_\theta^*}{\partial \theta} - \frac{\partial m_x^*}{\partial x} + (\gamma - 1) \rho^* \left[\frac{1}{r} \frac{\partial r u_r}{\partial r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} + \frac{\partial u_x}{\partial x} \right] - \\ - \frac{m_r}{\rho} \frac{\partial \rho^*}{\partial r} - \frac{m_\theta}{\rho r} \frac{\partial \rho^*}{\partial \theta} - \frac{m_x}{\rho} \frac{\partial \rho^*}{\partial x} - \frac{\gamma \mu}{\rho Pr Re} \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial \rho^*}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 \rho^*}{\partial \theta^2} + \frac{\partial^2 \rho^*}{\partial x^2} \right] = rhs \end{aligned} \quad (2.36a)$$

The *momentum* equations are

$$\begin{aligned}
& -\frac{\partial m_x^*}{\partial t} - \frac{\partial p^*}{\partial x} - \frac{\gamma p}{\rho} \frac{\partial \rho^*}{\partial x} - \frac{(\gamma-1)\rho^*}{\rho} \frac{\partial p}{\partial x} - \\
& -\frac{m_r}{\rho} \left[\frac{\partial m_r^*}{\partial x} + \frac{\partial m_x^*}{\partial r} \right] - \frac{m_\theta}{\rho} \left[\frac{\partial m_\theta^*}{\partial x} + \frac{1}{r} \frac{\partial m_x^*}{\partial \theta} \right] - 2 \frac{m_x}{\rho} \frac{\partial m_x^*}{\partial x} - \\
& -\frac{\mu}{Re\rho} \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial m_x^*}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 m_x^*}{\partial \theta^2} + \frac{\partial^2 m_x^*}{\partial x^2} + \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \frac{\partial \Theta^*}{\partial x} \right] - \\
& -\frac{\phi}{\rho} \frac{\partial \phi^*}{\partial x} = rhs \\
& -\frac{\partial m_r^*}{\partial t} - \frac{\partial p^*}{\partial r} - \frac{\gamma p}{\rho} \frac{\partial \rho^*}{\partial r} - \frac{(\gamma-1)\rho^*}{\rho} \frac{\partial p}{\partial r} - \\
& -2 \frac{m_r}{\rho} \frac{\partial m_r^*}{\partial r} - \frac{m_\theta}{\rho} \left[\frac{1}{r} \frac{\partial m_r^*}{\partial \theta} + \frac{\partial m_\theta^*}{\partial r} - \frac{m_\theta^*}{r} \right] - \frac{m_x}{\rho} \left[\frac{\partial m_r^*}{\partial x} + \frac{\partial m_x^*}{\partial r} \right] - \\
& -\frac{\mu}{Re\rho} \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial m_r^*}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 m_r^*}{\partial \theta^2} + \frac{\partial^2 m_r^*}{\partial x^2} - \frac{m_r^*}{r^2} - \frac{2}{r^2} \frac{\partial m_\theta^*}{\partial \theta} + \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \frac{\partial \Theta^*}{\partial r} \right] - \\
& -\frac{\phi}{\rho} \frac{\partial \phi^*}{\partial r} = rhs \\
& -\frac{\partial m_\theta^*}{\partial t} - \frac{1}{r} \frac{\partial p^*}{\partial \theta} - \frac{\gamma p}{\rho r} \frac{\partial \rho^*}{\partial \theta} - \frac{(\gamma-1)\rho^*}{\rho r} \frac{\partial p}{\partial \theta} - \\
& -\frac{m_r}{\rho} \left[\frac{1}{r} \frac{\partial m_r^*}{\partial \theta} + \frac{\partial m_\theta^*}{\partial r} - \frac{m_\theta^*}{r} \right] - 2 \frac{m_\theta}{\rho} \left[\frac{1}{r} \frac{\partial m_\theta^*}{\partial \theta} + \frac{m_r^*}{r} \right] - \frac{m_x}{\rho} \left[\frac{\partial m_\theta^*}{\partial x} + \frac{1}{r} \frac{\partial m_x^*}{\partial \theta} \right] - \\
& -\frac{\mu}{Re\rho} \left[\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial m_\theta^*}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 m_\theta^*}{\partial \theta^2} + \frac{\partial^2 m_\theta^*}{\partial x^2} - \frac{m_\theta^*}{r^2} + \frac{2}{r^2} \frac{\partial m_r^*}{\partial \theta} + \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \frac{1}{r} \frac{\partial \Theta^*}{\partial \theta} \right] - \\
& -\frac{\phi}{\rho} \frac{1}{r} \frac{\partial \phi^*}{\partial \theta} = rhs
\end{aligned} \tag{2.36b}$$

The *energy* equation yields

$$\begin{aligned}
& -\frac{\partial p^*}{\partial t} + \frac{\gamma p}{\rho^2} \left[m_r \frac{\partial \rho^*}{\partial r} + \frac{m_\theta}{r} \frac{\partial \rho^*}{\partial \theta} + m_x \frac{\partial \rho^*}{\partial x} \right] + \\
& + \frac{(\gamma-1)\rho^*}{\rho^2} \left[m_r \frac{\partial p}{\partial r} + \frac{m_\theta}{r} \frac{\partial p}{\partial \theta} + m_x \frac{\partial p}{\partial x} \right] + \frac{m_r}{\rho^2} \left[m_r \frac{\partial m_r^*}{\partial r} + \frac{m_\theta}{r} \frac{\partial m_r^*}{\partial \theta} - \frac{m_\theta m_\theta^*}{r} + m_x \frac{\partial m_r^*}{\partial x} \right] + \\
& + \frac{m_\theta}{\rho^2} \left[m_r \frac{\partial m_\theta^*}{\partial r} + \frac{m_\theta}{r} \frac{\partial m_\theta^*}{\partial \theta} + \frac{m_\theta m_r^*}{r} + m_x \frac{\partial m_\theta^*}{\partial x} \right] + \frac{m_x}{\rho^2} \left[m_r \frac{\partial m_x^*}{\partial r} + \frac{m_\theta}{r} \frac{\partial m_x^*}{\partial \theta} + m_x \frac{\partial m_x^*}{\partial x} \right] + \\
& + \frac{\mu}{Re \rho^2} \left[m_r \left(\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial m_r^*}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 m_r^*}{\partial \theta^2} + \frac{\partial^2 m_r^*}{\partial x^2} - \frac{m_r^*}{r^2} - \frac{2}{r^2} \frac{\partial m_\theta^*}{\partial \theta} \right) + \right. \\
& \quad + m_\theta \left(\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial m_\theta^*}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 m_\theta^*}{\partial \theta^2} + \frac{\partial^2 m_\theta^*}{\partial x^2} - \frac{m_\theta^*}{r^2} + \frac{2}{r^2} \frac{\partial m_r^*}{\partial \theta} \right) + \\
& \quad \left. + m_x \left(\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial m_x^*}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 m_x^*}{\partial \theta^2} + \frac{\partial^2 m_x^*}{\partial x^2} \right) \right] + \\
& + \frac{\mu}{Re \rho^2} \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \left(m_r \frac{\partial \Theta^*}{\partial r} + \frac{m_\theta}{r} \frac{\partial \Theta^*}{\partial \theta} + m_x \frac{\partial \Theta^*}{\partial x} \right) + \\
& + \frac{\phi}{\rho^2} \left[m_r \frac{\partial \phi^*}{\partial r} + \frac{m_\theta}{r} \frac{\partial \phi^*}{\partial \theta} + m_x \frac{\partial \phi^*}{\partial x} \right] + \frac{\mu \phi}{Re Sc \rho^2} \left[\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial \phi^*}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 \phi^*}{\partial \theta^2} + \frac{\partial^2 \phi^*}{\partial x^2} \right] + \\
& + \frac{\gamma \mu}{\rho Pr Re \rho} \left[\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial \rho^*}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 \rho^*}{\partial \theta^2} + \frac{\partial^2 \rho^*}{\partial x^2} \right] = rhs
\end{aligned} \tag{2.36c}$$

Finally, the adjoint *scalar* equation is:

$$\begin{aligned}
& -\frac{\partial \phi^*}{\partial t} - \frac{1}{\rho} \left[m_r \frac{\partial \phi^*}{\partial r} + \frac{m_\theta}{r} \frac{\partial \phi^*}{\partial \theta} + m_x \frac{\partial \phi^*}{\partial x} \right] - \frac{\mu}{Re Sc \rho} \left[\frac{1}{r} \frac{\partial}{\partial r} (r \frac{\partial \phi^*}{\partial r}) + \frac{1}{r^2} \frac{\partial^2 \phi^*}{\partial \theta^2} + \frac{\partial^2 \phi^*}{\partial x^2} \right] = rhs
\end{aligned} \tag{2.36d}$$

In the previous equations we have defined

$$\Theta^* = \frac{1}{r} \frac{\partial}{\partial r} (r m_r^*) + \frac{1}{r} \frac{\partial m_\theta^*}{\partial \theta} + \frac{\partial m_x^*}{\partial x} . \tag{2.37}$$

The right-hand side of the equations (forcing term) has been obtained in the previous sections considering a time-periodic control actuation.

2.A.3 Numerical problems

CFL restriction

Due to the non-preferred azimuthal direction of the jet, the discretization in the azimuthal direction θ is uniform. For this reason, the CFL number is very restricted near the

centerline. To avoid it, only a few modes are kept in that region. As indicated in Mohseni & Colonius (2000), if we consider a grid defined by $r_n = n\Delta r$, with $n = 0, 1, 2, \dots$, the effective spacing in the azimuthal direction is $\Delta x_\theta = 2\pi\Delta r/N_\theta$, where N_θ is the number of modes retained. The CFL constraint is dictated only by Δr when the number of retained modes is

$$N_\theta = 2\pi n. \quad (2.38)$$

Singularity at the origin

The equations have a singularity at the centerline ($r=0$). Therefore this point needs special treatment. There are different ways to solve this issue. One the methods to avoid the problem is the use of cartesian coordinates at the origin, as for example in Freund *et al.* (1997). Mohseni & Colonius (2000) use a mesh which does not contain the node at the origin. Other people use l'Hôpital rule. The method that has given the best results to our problem is the one proposed in Constantinescu & Lele (2001). The variables are expanded (series expansion), and a set of new exact equations is obtained at the centerline.

The general expansions of a single valued quantity S (like p) and a multi-valued quantity M (like u_r and u_θ) are:

$$S(r, \theta) = \sum_{m=0}^{\infty} r^m \left(\sum_{n=0}^{\infty} \alpha_{mn}^{(S)} r^{2n} \right) \cos(m\theta) + \sum_{m=0}^{\infty} r^m \left(\sum_{n=0}^{\infty} \beta_{mn}^{(S)} r^{2n} \right) \sin(m\theta)$$

$$M(r, \theta) = \frac{1}{r} \sum_{n=1}^{\infty} A_{0n}^{(M)} r^{2n} + \sum_{m=1}^{\infty} r^{m-1} \sum_{n=0}^{\infty} \left(A_{mn}^{(M)} \cos(m\theta) + B_{mn}^{(M)} \sin(m\theta) \right) r^{2n}$$

Substituting all the variables with these expansions, and taking the limit when $r \rightarrow 0$ yields new equations at the origin for the coefficients. The equations at the origin corresponding to the compressible flow are given in Constantinescu & Lele (2001). The new equations for the adjoint field are given below.

In the case of the *continuity* equation, we now have

$$-\frac{\partial \rho^*}{\partial t} - \left[2A_{01}^{(m_r^*)} + \frac{\partial \alpha_{00}^{(m_x^*)}}{\partial x} \right] + (\gamma - 1)\alpha_{00}^{(\rho^*)} \left[2A_{01}^{(u_r)} + \frac{\partial \alpha_{00}^{(u_x)}}{\partial x} \right] -$$

$$-\frac{\alpha_{00}^{(m_x)}}{\alpha_{00}^{(\rho)}} \frac{\partial \alpha_{00}^{(\rho^*)}}{\partial x} - \frac{A_{10}^{(m_r)} \alpha_{10}^{(\rho^*)} + B_{10}^{(m_r)} \beta_{10}^{(\rho^*)}}{\alpha_{00}^{(\rho)}} - \frac{\gamma}{Pr} \alpha_{00}^{(\frac{\mu}{\rho k_e})} \left[\frac{\partial^2 \alpha_{00}^{(\rho^*)}}{\partial x^2} + 4\alpha_{01}^{(\rho^*)} \right] = rhs$$

The equation for the x -momentum equation at the centerline is

$$\begin{aligned}
 & -\frac{\partial(m_x^*)}{\partial t} - \frac{\partial\alpha_{00}^{(p^*)}}{\partial x} - \frac{\gamma\alpha_{00}^{(p)}}{\alpha_{00}^{(\rho)}} \frac{\partial\alpha_{00}^{(\rho^*)}}{\partial x} - \frac{(\gamma-1)\alpha_{00}^{(\rho^*)}}{\alpha_{00}^{(\rho)}} \frac{\partial\alpha_{00}^{(p)}}{\partial x} - \\
 & -2\alpha_{00}^{(u_x)} \frac{\partial\alpha_{00}^{(m_x^*)}}{\partial x} - A_{10}^{(u_r)} \left(\frac{\partial A_{10}^{(m_r^*)}}{\partial x} + \alpha_{10}^{(m_x^*)} \right) - B_{10}^{(u_r)} \left(\frac{\partial B_{10}^{(m_r^*)}}{\partial x} + \beta_{10}^{(m_x^*)} \right) - \\
 & -v \left[\frac{\partial^2 \alpha_{00}^{(m_x^*)}}{\partial x^2} + 4\alpha_{01}^{(m_x^*)} \right] - \alpha_{00}^{(\frac{\mu}{\rho Re})} \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \left[2 \frac{\partial A_{01}^{(m_r^*)}}{\partial x} + \frac{\partial^2 \alpha_{00}^{(m_x^*)}}{\partial x^2} \right] - \alpha_{00}^{(\xi)} \frac{\partial\alpha_{00}^{(\phi^*)}}{\partial x} = rhs
 \end{aligned}$$

The equation for the r -momentum yields:

$$\begin{aligned}
 & -\frac{\partial m_r^*}{\partial t} - (\alpha_{10}^{(p^*)} + \frac{\gamma\alpha_{00}^{(p)}}{\alpha_{00}^{(\rho)}} \alpha_{10}^{(\rho^*)} + \frac{(\gamma-1)\alpha_{00}^{(\rho^*)}}{\alpha_{00}^{(\rho)}} \alpha_{10}^{(p)}) \cos \theta - (\beta_{10}^{(p^*)} + \frac{\gamma\alpha_{00}^{(p)}}{\alpha_{00}^{(\rho)}} \beta_{10}^{(\rho^*)} + \\
 & + \frac{(\gamma-1)\alpha_{00}^{(\rho^*)}}{\alpha_{00}^{(\rho)}} \beta_{10}^{(p)}) \sin \theta - 2(A_{01}^{(m_r^*)} A_{10}^{(u_r)} + A_{10}^{(u_r)} A_{20}^{(m_r^*)} + B_{10}^{(u_r)} B_{20}^{(m_r^*)}) \cos \theta - \\
 & - \alpha_{00}^{(u_x)} \left(\frac{\partial A_{10}^{(m_r^*)}}{\partial x} + \alpha_{10}^{(m_x^*)} \right) \cos \theta - 2(A_{01}^{(m_r^*)} B_{10}^{(u_r)} + A_{10}^{(u_r)} B_{20}^{(m_r^*)} - B_{10}^{(u_r)} A_{20}^{(m_r^*)}) \sin \theta - \\
 & - \alpha_{00}^{(u_x)} \left(\frac{\partial B_{10}^{(m_r^*)}}{\partial x} + \beta_{10}^{(m_x^*)} \right) \sin \theta - v \left[2(A_{11}^{(m_r^*)} - B_{11}^{(m_\theta^*)}) + \frac{\partial^2 A_{10}^{(m_r^*)}}{\partial x^2} \right] \cos \theta - \\
 & - v \left[2(B_{11}^{(m_r^*)} + A_{11}^{(m_\theta^*)}) + \frac{\partial^2 B_{10}^{(m_r^*)}}{\partial x^2} \right] \sin \theta - \\
 & - \alpha_{00}^{(\frac{\mu}{\rho Re})} \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \left[\left(3A_{11}^{(m_r^*)} + B_{11}^{(m_\theta^*)} + \frac{\partial\alpha_{10}^{(m_x^*)}}{\partial x} \right) \cos \theta + \right. \\
 & \quad \left. + \left(3B_{11}^{(m_r^*)} - A_{11}^{(m_\theta^*)} + \frac{\partial\beta_{10}^{(m_x^*)}}{\partial x} \right) \sin \theta \right] - \\
 & - \alpha_{00}^{(\xi)} (\alpha_{10}^{(\phi^*)} \cos \theta + \beta_{10}^{(\phi^*)} \sin \theta) = rhs
 \end{aligned}$$

For the θ -momentum equation we have to take into account that m_r can be obtained from m_θ by a counterclockwise rotation of $\pi/2$. Then, if

$$m_r^* = A_{10}^{(m_r^*)} \cos \theta + B_{10}^{(m_r^*)} \sin \theta$$

the azimuthal component can be written as:

$$m_\theta^* = A_{10}^{(m_\theta^*)} \cos \theta + B_{10}^{(m_\theta^*)} \sin \theta = -A_{10}^{(m_r^*)} \sin \theta + B_{10}^{(m_r^*)} \cos \theta$$

Energy equation:

$$\begin{aligned}
& -\frac{\partial p}{\partial t} + \frac{\gamma \alpha_{00}^{(p)}}{(\alpha_{00}^{(\rho)})^2} \left[A_{10}^{(m_r)} \alpha_{10}^{(\rho^*)} + B_{10}^{(m_r)} \beta_{10}^{(\rho^*)} + \alpha_{00}^{(m_x)} \frac{\partial \alpha_{00}^{(\rho^*)}}{\partial x} \right] + \\
& + \frac{(\gamma - 1) \alpha_{00}^{(\rho^*)}}{(\alpha_{00}^{(\rho)})^2} \left[A_{10}^{(m_r)} \alpha_{10}^{(p)} + B_{10}^{(m_r)} \beta_{10}^{(p)} + \alpha_{00}^{(m_x)} \frac{\partial \alpha_{00}^{(p)}}{\partial x} \right] + \\
& + \frac{1}{(\alpha_{00}^{(\rho)})^2} \left[\left((A_{10}^{(m_r)})^2 + (B_{10}^{(m_r)})^2 \right) A_{01}^{(m_r^*)} + \left((A_{10}^{(m_r)})^2 - (B_{10}^{(m_r)})^2 \right) A_{20}^{(m_r^*)} + 2A_{10}^{(m_r)} B_{10}^{(m_r)} B_{20}^{(m_r^*)} \right] + \\
& + \frac{\alpha_{00}^{(m_x)}}{(\alpha_{00}^{(\rho)})^2} \left[A_{10}^{(m_r)} \frac{\partial A_{10}^{(m_r^*)}}{\partial x} + B_{10}^{(m_r)} \frac{\partial B_{10}^{(m_r^*)}}{\partial x} \right] + \\
& + \frac{\alpha_{00}^{(m_x)}}{(\alpha_{00}^{(\rho)})^2} \left[A_{10}^{(m_r)} \alpha_{10}^{(m_x^*)} + B_{10}^{(m_r)} \beta_{10}^{(m_x^*)} + \alpha_{00}^{(m_x)} \frac{\partial \alpha_{00}^{(m_x^*)}}{\partial x} \right] + \\
& + \alpha_{00}^{(\frac{\mu}{\rho Re})} \left[A_{10}^{(u_r)} \left(2(A_{11}^{(m_r^*)} - B_{11}^{(m_\theta^*)}) + \frac{\partial^2 A_{10}^{(m_r^*)}}{\partial x^2} \right) + B_{10}^{(u_r)} \left(2(B_{11}^{(m_r^*)} + A_{11}^{(m_\theta^*)}) + \frac{\partial^2 B_{10}^{(m_r^*)}}{\partial x^2} \right) + \right. \\
& + \alpha_{00}^{(u_x)} \left(\frac{\partial^2 \alpha_{00}^{(m_x^*)}}{\partial x^2} + 4\alpha_{01}^{(m_x^*)} \right) \left. + \alpha_{00}^{(\frac{\mu}{\rho Re})} \left(\frac{\mu_B}{\mu} + \frac{1}{3} \right) \left[A_{10}^{(u_r)} \left(3A_{11}^{(m_r^*)} + B_{11}^{(m_\theta^*)} + \frac{\partial \alpha_{10}^{(m_x^*)}}{\partial x} \right) + \right. \right. \\
& + B_{10}^{(u_r)} \left(3B_{11}^{(m_r^*)} - A_{11}^{(m_\theta^*)} + \frac{\partial \beta_{10}^{(m_x^*)}}{\partial x} \right) + \alpha_{00}^{(u_x)} \left(2 \frac{\partial A_{01}^{(m_r^*)}}{\partial x} + \frac{\partial^2 \alpha_{00}^{(m_x^*)}}{\partial x^2} \right) \left. \right] + \\
& + \alpha_{00}^{(\xi)} \left[\alpha_{00}^{(u_x)} \frac{\partial \alpha_{00}^{(\phi^*)}}{\partial x} + A_{10}^{(u_r)} \alpha_{10}^{(\phi^*)} + B_{10}^{(u_r)} \beta_{10}^{(\phi^*)} \right] + \\
& + \frac{\alpha_{00}^{(\frac{\mu}{\rho Re})}}{Sc} \left[\frac{\partial^2 \alpha_{00}^{(\phi^*)}}{\partial x^2} + 4\alpha_{01}^{(\phi^*)} \right] + \frac{\gamma}{Pr} \alpha_{00}^{(\frac{\mu}{\rho Re})} \frac{\alpha_{00}^{(p)}}{\alpha_{00}^{(\rho)}} \left[\frac{\partial^2 \alpha_{00}^{(\rho^*)}}{\partial x^2} + 4\alpha_{01}^{(\rho^*)} \right] = rhs
\end{aligned}$$

And finally, the *scalar* equation:

$$-\frac{\partial \phi^*}{\partial t} - \alpha_{00}^{(u_x)} \frac{\partial \alpha_{00}^{(\phi^*)}}{\partial x} - A_{10}^{(u_r)} \alpha_{10}^{(\phi^*)} - B_{10}^{(u_r)} \beta_{10}^{(\phi^*)} - \frac{\alpha_{00}^{(\frac{\mu}{\rho Re})}}{Sc} \left[\frac{\partial^2 \alpha_{00}^{(\phi^*)}}{\partial x^2} + 4\alpha_{01}^{(\phi^*)} \right] = rhs$$

Accuracy at the centerline

In order to obtain a better accuracy close to the centerline, it would be desirable to use our sixth-order Padé. In Freund *et al.* (1997), where cartesian coordinates were used at $r = 0$, the accuracy is the same as at the boundaries, since the origin is treated as an artificial boundary when computing the radial derivatives. Then, the accuracy is less than 6th order.

Transforming the domain from $(0, R) \times (0, 2\pi)$ to $(-R, R) \times (0, \pi)$ when computing the radial derivatives achieves sixth order, since the point $r = 0$ is treated as any other point in the grid.

References

- AZEVEDO, A. & REZENDE, S. 1991 Controlling chaos in spin-wave instabilities. *Phys. Rev. Lett.* **66**, 1342.
- BEWLEY, T. Preprint Numerical methods in science and engineering. Available at http://turbulence.ucsd.edu/numerical_methods/notes.pdf.
- BEWLEY, T., MOIN, P. & TEMAM, R. 2001 DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms. *J. Fluid Mech.* **447**, 179–225.
- CERVINO, L. & BEWLEY, T. 2003 On the extension of the complex-step derivative technique to pseudospectral algorithms. *J. Comput. Phys.* **187**, 544–549.
- CERVINO, L., BEWLEY, T., FREUND, J. & LELE, S. 2002 Perturbation and adjoint analyses of flow-acoustic interactions in an unsteady 2d jet. In *Center for Turbulence Research, Proceedings of the Summer Program 2002*, pp. 27–40.
- COLONIUS, T., LELE, S. & MOIN, P. 1993 Boundary conditions for direct computation of aerodynamic sound generation. *AIAA J.* **31**, 1574–1582.
- COLONIUS, T., LELE, S. & MOIN, P. 1997 Sound generation in a mixing layer. *J. Fluid Mech.* **330**, 375–409.
- CONSTANTINESCU, G. S. & LELE, S. K. 2001 A Highly accurate technique for the treatment of flow equations at the polar axis in cylindrical coordinates using series expansion. *Submitted J. Comp. Phys.* .
- FREUND, J. 1997 Proposed inflow/outflow boundary condition for direct computation of aerodynamic sound. *AIAA J.* **35**, 740–742.
- FREUND, J. 2001 Noise sources in a low-reynolds-number turbulent jet at mach 0.9. *J. Fluid Mech.* **438**, 277–305.
- FREUND, J., P. MOIN, P. & LELE, S. 1997 Compressibility effects in a turbulent annular mixing layer. *Tech. Rep. TF-72*. Mech. Eng. Dept., Stanford University.
- JAMESON, A., ALONSO, J. J., REUTHER, J. J., MARTINELLI, L. & VASSBERG, J. 1998 Aerodynamic shape optimization techniques based on control theory. In *AIAA Paper 98-2538*.
- LAUGA, E. & BEWLEY, T. 2003 The decay of stabilizability with reynolds number in a linear model of spatially developing flows. *Proc. R. Soc. Lond. A* **459**, 2077–2095.
- LELE, S. 1992 Compact finite difference schemes with spectral-like resolution. *J. Comput. Phys.* **103**, 16–42.
- LIMA, R. & PETTINI, M. 1990 Suppression of chaos by resonant parametric perturbations. *Phys. Rev. A* **41**, 726–733.

- LYNESS, J. & MOLER, C. 1967 Numerical differentiation of analytic functions. *SIAM J. Numer. Anal.* **4**, 202–210.
- MARTINS, J., STURDZA, P. & ALONSO, J. 2001 The connection between the complex step derivative approximation and algorithmic differentiation. In *AIAA Paper 01-0921*.
- MIRUS, K. & SPROTT, J. 1999 Controlling chaos in a high dimensional system with periodic parametric perturbations. *Phys. Lett. A* **254**, 275–278.
- MITCHELL, B., LELE, S. & MOIN, P. 1999 Direct computation of the sound generated by vortex pairing in an axisymmetric jet. *J. Fluid Mech.* **383**, 113–142.
- MOHSENI, K. & COLONIUS, T. 2000 Numerical Treatment of Polar Coordinate Singularities. *J. Comp. Phys.* **157**, 787–795.
- PAPAMOSCHOU, D. 1997 Mach wave elimination in supersonic jets. In *AIAA Paper 97-0147*.
- PAPAMOSCHOU, D. 2003 A new method for jet noise reduction in turbofan engines. In *AIAA Paper 03-1059*.
- PRESS, W., FLANNERY, B., TEUKOLSKY, S. & VETTERLING, W. 1992 *Numerical Recipes..* Cambridge University Press.
- PROTAS, B., BEWLEY, T. & HAGEN, G. 2004 A comprehensive framework for the regularization of adjoint analysis in multiscale pde systems. *J. Comput. Phys.* **195** (1), 49–89.
- SAIYED, N., MIKKELSEN, K. & BRIDGES, J. 2000 Acoustics and thrust of separate flow exhaust nozzles with mixing devices investigated for high bypass ratio engines. In *AIAA Paper 00-1961*.
- SÄTHERBLOM, H.-E. 1997 Dynamic control of the reversed-field pinch in numerical magnetohydrodynamic simulations. *Phys. Rev. Lett.* **4** (1), 174–178.
- STANLEY, S. & SARKAR, S. 1997 Simulations of spatially developing two dimensional shear layers and jets. *Theoret. Comput. Fluid Dynamics* **9**, 121–147.
- SUZUKI, T. 2001 Acoustic wave propagation in transversely sheared flows. PhD thesis, Stanford University, Department of Aeronautics and Astronautic, also available as SUDAAR 739.
- SUZUKI, T. & LELE, S. K. 1999 Acoustic scattering from a mixing layer: role of instability waves. In *AIAA Paper 99-0228*.
- TAM, C. & AURIAULT, L. 1998 Mean flow refraction effects on sound radiated from localized sources in a jet. *J. Fluid Mech.* **370**, 149–174.

THOMPSON, K. W. 1987 Time Dependent Boundary Conditions for Hyperbolic Systems. *J. Comp. Phys.* **68**, 1–24.

THOMPSON, K. W. 1990 Time Dependent Boundary Conditions for Hyperbolic Systems, II. *J. Comp. Phys.* **89**, 439–461.

WEI, M. & FREUND, J. 2002 Optimal control of free shear flow. In *AIAA Paper 02-0665*.

WHITMIRE, J. & SARKAR, S. 2000 Validation of acoustic-analogy predictions for sound radiated by turbulence. *Phys. Fluids* **12** (2), 381–391.

Chapter 3

Extension of the complex-step derivative technique to pseudospectral algorithms

3.1 Introduction

The complex-step derivative (CSD) technique is a convenient and highly-accurate strategy to perform a linearized “perturbation” analysis to determine a “directional derivative” via a minor modification of an existing nonlinear simulation code. The technique has previously been applied to nonlinear simulation codes (such as finite-element codes) which employ real arithmetic only. The present note examines the suitability of this technique for extension to efficient pseudospectral simulation codes which nominally use the fast Fourier transform (FFT) to convert back and forth between the physical and transformed representations of the system. It is found that, if used carefully, this extension retains the remarkable accuracy of the CSD approach. However, to perform this extension without sacrificing this accuracy, particular care must be exercised; specifically, the state (real) and perturbation (imaginary) components of the complexified system must be transformed separately and arranged in such a manner that they are kept distinct during the process of differentiation in the transformed space in order to avoid the linear combi-

nation of the large and small quantities in the analysis. It is shown that this is relatively straightforward to implement even in complicated nonlinear simulation codes, thus positioning the CSD approach as an attractive and relatively simple alternative to hand coding a perturbation (a.k.a. “tangent linear”) code for determining the directional derivative even when pseudospectral algorithms are employed.

3.2 Background

The idea of using complex variables in order to differentiate a function was (apparently) first mentioned in Lyness & Moler (1967) and Lyness (1967). The applicability of this old technique to differentiate a complicated function, such as that computed by an involved finite-element code, was recognized by Squire & Trapp (1998). Recently, the CSD technique has become popular in aerodynamic optimization Newman *et al.* (1998); Anderson *et al.* (1999); Martins *et al.* (2000, 2001). This approach has proven to be very accurate and easy to apply to finite-element and finite-difference simulation codes but, to the best of our knowledge, has not yet been extended to pseudospectral simulation codes. The present note explores the suitability of this extension.

3.2.1 The Finite-Difference (FD) approximation

The directional derivative d of a continuous function $J(\phi)$ in the direction ϕ_1 is defined by

$$d \triangleq \frac{\mathcal{D}J(\phi)}{\mathcal{D}\phi} \cdot \phi_1 = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{\varepsilon} [J(\phi + \varepsilon \phi_1) - J(\phi)], \quad (3.1)$$

which is simply the amount $J(\phi)$ changes when ϕ is updated in the direction ϕ_1 , scaled by the size of the update, in the limit that the size of the update approaches zero. There are a variety of ways to calculate numerically this seemingly simple quantity. The most straightforward method to compute d is to consider a Taylor series expansion of the function $J(\phi + \varepsilon \phi_1)$ near ϕ , from which a first-order finite-difference (FD) formula for the

directional derivative is easily obtained:

$$\mathcal{J}(\phi + \varepsilon \phi_1) = \mathcal{J}(\phi) + \varepsilon \frac{\mathcal{D}\mathcal{J}(\phi)}{\mathcal{D}\phi} \cdot \phi_1 + O(\varepsilon^2) \quad \Rightarrow \quad d = \frac{\mathcal{J}(\phi + \varepsilon \phi_1) - \mathcal{J}(\phi)}{\varepsilon} + O(\varepsilon).$$

Likewise, second- and higher-order finite-difference formulae for the directional derivative can be easily obtained. The drawback with these methods when using a finite-precision arithmetic computer is the difficulty in finding the most suitable value for the step size ε , since the accuracy of the numerical approximation of d is very sensitive to this value. When ε is large, the Taylor-series truncation is not valid, and when it is small, subtractive cancellation errors dominate.

To illustrate, the derivative of a scalar nonlinear function $f(x) = \frac{e^{-x}}{\sqrt{\tan x}}$ at $x = 1$ has been computed using FD approximations with both single- and double-precision arithmetic. Figure 3.1 shows that, for large ε , the error of these FD approximations scales with ε^n , where n is the order of truncation of the higher-order terms of the corresponding FD formulae. For small ε , the error of all three FD formula in Figure 3.1 is $O(1/\varepsilon)$ due to subtractive cancellation errors. In other words, when comparing two numbers which are almost the same using finite-precision arithmetic, the relative round-off error is proportional to the inverse of the difference between the two numbers. If the difference between the two numbers is decreased by an order of magnitude, the relative accuracy with which this difference may be calculated using finite precision arithmetic is also decreased by an order of magnitude.

3.2.2 The Complex-Step Derivative (CSD) approximation

The CSD approximation makes use of complex variables in order to compute the directional derivative. If the complex extension $\mathcal{J}(z)$ of a real-valued function $\mathcal{J}(\phi)$ is analytic, it can be expanded with a complex Taylor series. In particular, the expansion of $\mathcal{J}(\phi + i\varepsilon\phi_1)$, where $i = \sqrt{-1}$, may be written:

$$\mathcal{J}(\phi + i\varepsilon\phi_1) = \mathcal{J}(\phi) + i\varepsilon \frac{\mathcal{D}\mathcal{J}(\phi)}{\mathcal{D}\phi} \cdot \phi_1 - \varepsilon^2 E_1 - i\varepsilon^3 E_2 + O(\varepsilon^4), \quad (3.2)$$

where E_1 and E_2 are real and are related to the higher-order derivatives of $\mathcal{J}$. Taking the imaginary part and dividing by ε gives a formula for the directional derivative:

$$d = \Im\left\{\frac{\mathcal{J}(\phi + i\varepsilon\phi_1)}{\varepsilon}\right\} + O(\varepsilon^2); \quad (3.3)$$

note that the error in this formula is $O(\varepsilon^2)$ because the leading-order error of Eq. (3.2) is real. This method has the advantage that it does not face the problem of the subtractive cancellation error; that is, d is computed simply by taking the imaginary part of the complex number $\mathcal{J}(\phi + i\varepsilon\phi_1)$. As seen in Figure 3.1, for small ε , the relative error of the CSD approximation reaches an asymptotic value given by the machine precision. Thus, when the CSD method is applied to this test problem, *any* choice of the step-size (providing it is sufficiently small) gives a very accurate result. Similar results are also seen when the CSD approach is applied to much more complicated functions, such as the calculation of the drag of a wing using an involved finite-element code Martins *et al.* (2000, 2001). As the CSD technique is both highly accurate and quite easy to apply to an existing nonlinear simulation code (simply convert all real variables in the code to complex and perturb the imaginary component of the control variable), it has become a popular technique for computing the directional derivative.

3.3 Extension of the CSD approach to Fourier-based pseudospectral codes

Many codes in both fundamental turbulence research and numerical weather prediction are pseudospectral. In such codes, products are computed in physical space, spatial derivatives are computed in some “transformed” space, and the conversion between these two representations is made efficiently with a fast transform technique, usually some variant of the fast Fourier transform (FFT). Pseudospectral techniques are essential for the accurate computation of spatial derivatives in multiscale fluid systems when marginal resolution is employed.

The most common and, perhaps, simplest pseudospectral approach is that in which

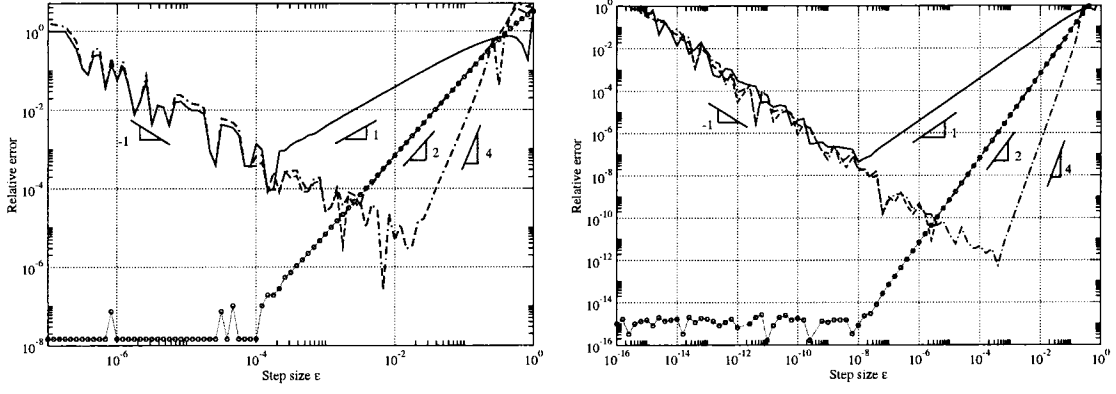


Figure 3.1: Relative error of the directional derivative of $f(x) = \frac{e^{-x}}{\sqrt{\tan x}}$ at $x = 1$ given by (solid) first-order, (dashed) second-order, and (dot-dashed) fourth-order FD approaches and (circles+dots) the CSD approach using single-precision arithmetic (left) and double-precision arithmetic (right). Note that both plots have essentially the same shape.

the transformed system is represented as a Fourier series. Unfortunately, as Fourier-based pseudospectral codes already employ complex arithmetic, direct application of the standard complex-step derivative method, as discussed above, is not possible. However, as the FFT is simply a linear manipulation of an otherwise real problem, the idea behind the CSD method is still valid. In order to implement the CSD method in an existing Fourier-based pseudospectral code, it is necessary to convert all formerly real arrays into complex arrays, double the size of all formerly complex arrays¹, and modify the FFT calls appropriately; we consider here two possible strategies:

- A. Simply replace the real $\rightarrow$ complex FFTs with complex $\rightarrow$ complex FFTs.
- B. Transform the state (real) and perturbation (imaginary) components with separate real $\rightarrow$ complex transforms, storing the results of these two transforms in the original (complex) data array (now doubled in size) in such a manner that performing spatial differentiation in the transformed space does not combine the state and perturbation components of the analysis. For example, in the Fourier-based case with transforms

¹Note that, when computing the FFT of a real function f (in one-dimensional, two-dimensional and three-dimensional problems), it is only necessary to store half of the Fourier coefficients, since the other half may be recovered by the identity $\hat{f}(\mathbf{k}) = \hat{f}^*(-\mathbf{k})$. This identity no longer holds if f is complex.

applied in the x direction only, defining $f = f_r + f_i$ where f_r is real and f_i is pure imaginary, it follows that $\hat{f}_r(k_x) = \hat{f}_r^*(-k_x)$ and $\hat{f}_i(k_x) = -\hat{f}_i^*(-k_x)$. Thus, for example, one may define the result of the physical $\rightarrow$ Fourier transform of the quantity f as:

$$\hat{f}(k_x) \triangleq \begin{cases} \hat{f}_r(k_x) & \text{for } k_x \geq 0, \\ \hat{f}_i(k_x) & \text{for } k_x < 0, \\ \hat{f}_i(0) & \text{for } k_x = NX/2 \text{ (the so-called "oddball" wavenumber).} \end{cases}$$

Spatial derivatives may then be computed using the original simulation code (however complicated) without further modification. Further, this representation provides enough information to recover both f_r and f_i upon separate complex $\rightarrow$ real inverse transforms while insuring that the state and perturbation components of the analysis are not combined during either the (forward and inverse) transforms or when applying the differentiation operations at each wavenumber.

To illustrate the difference in accuracy between the FD method and the two variants of the CSD method described above when applied to a pseudospectral simulation code, both methods have been implemented on an artificial optimization problem in which the dynamical system is governed by the nonlinear viscous Burgers equation

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} = \nu \frac{\partial^2 u}{\partial x^2}$$

in the periodic domain $x \in [0, 2\pi]$. A pseudospectral technique (FFT) is used to compute the spatial derivatives. The system is advanced in time from 0 to T , and the values $\nu = 0.067$ and $T = 6$ were used in the numerical simulation. The initial condition is given by $u(0) = 1 + 0.2 \sin(x) + 0.02 \sin^2(x - 1) + \phi \sin(x - 1)$, where ϕ is the "control parameter". The function of interest is defined as $\mathcal{J}(\phi) = \int_0^T \int_{4.5}^{5.0} |u(x, t)|^2 dx dt$. The relative errors of both the FD and CSD approximations of $d\mathcal{J}/d\phi$ at $\phi = 0$ are shown in Figure 3.2. The exact value is obtained with a "direct" method, formed by writing a separate numerical code (sometimes called a "tangent linear" code) which explicitly calculates the perturbation equation (for further discussion, see, e.g., Adelman & Haftka (1986)). As

seen in Figure 3.2, for large ϵ , the errors of both the FD and CSD approximations are seen to scale with ϵ in a manner which is similar to the case evaluated in Figure 3.1.

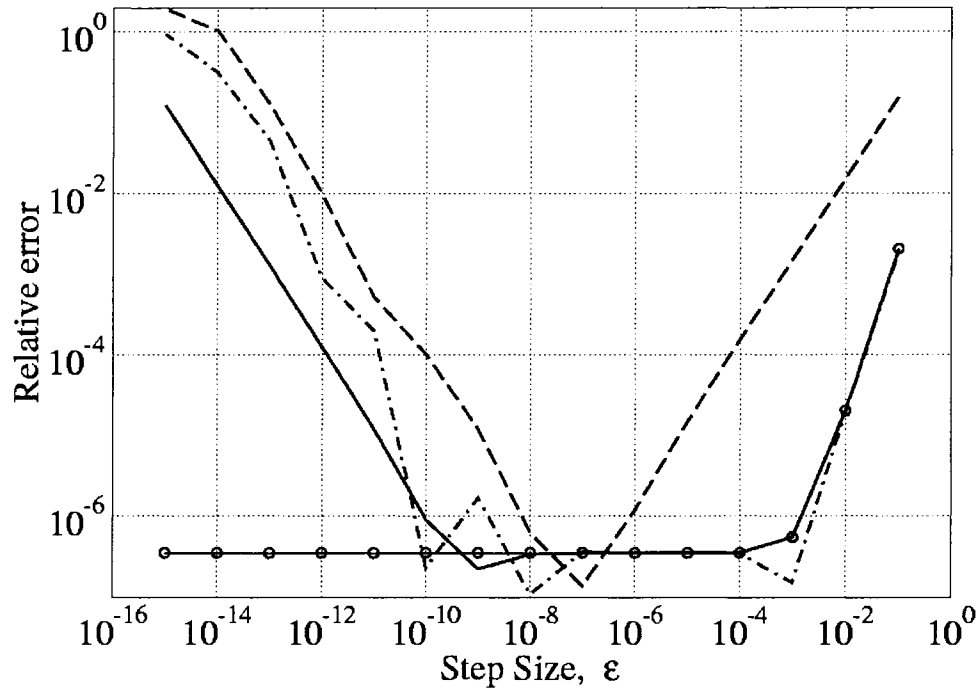


Figure 3.2: Relative error of (dashed) first-order and (dot-dashed) second-order FD approaches and the CSD approach using (solid) strategy A and (solid-circle) strategy B.

The CSD approach using strategy B (in which the real and imaginary parts are transformed separately) achieves an asymptotic value of the error for small ϵ in the same manner as it did in Figure 3.1. The selection of ϵ thus has no effect on the accuracy of the calculation, as long as it is chosen to be sufficiently small.

However when using strategy A (in which the real and imaginary parts are transformed simultaneously using a complex $\rightarrow$ complex transform), the error of the CSD method does not approach an asymptotic value. Instead, it scales like $O(1/\epsilon)$ for small ϵ . This is due to the fact that complex $\rightarrow$ complex Fourier transforms combine the real and imaginary parts of the analysis. Thus, the large (real) and small (imaginary) numbers in the analysis are combined by the FFTs, leading to numerical inaccuracies which are found to be comparable to the subtractive cancellation errors of the second-order FD approach.

3.4 Discussion and extension to other pseudospectral methods

The present analysis extends directly to other efficient pseudospectral methods based on fast transforms, such as those based on Chebyshev transforms, sine/cosine transforms, etc. Note that, at their core, most such methods make use of FFTs and therefore face issues quite similar to those explored above. More precisely, if the CSD calculation anywhere involves a linear combination of the state and perturbation components of the analysis, the CSD method will fail to provide an asymptotic (small) error as ϵ is made small.

However, via a minor modification of the wrappers on the forward and inverse transformation algorithms, it appears always to be possible to transform the state and perturbation components of the analysis separately and to keep these two components isolated during the process of spatial differentiation in the transformed representation, thereby retaining the remarkable accuracy of the CSD approach for computing the directional derivative over a broad range of the perturbation magnitude via a minor modification of the original simulation code. The CSD method for computing the directional derivative has thus been found to extend to pseudospectral simulation codes in a favorable manner, providing an attractive and simple alternative to writing a separate “tangent linear” code from scratch for the computation of the directional derivative.

Acknowledgments

The authors gratefully acknowledge generous funding from the programs of Dr. Belinda King and Dr. John Schmisser at AFOSR.

The text of this chapter, in full, is a reprint of the material as it appears in Cerviño & Bewley, 2003, “On the extension of the complex-step derivative technique to pseudospectral algorithms”. *J. Comput. Phys.* 187, 544–549. The dissertation author was the primary author and the coauthor listed in this publication directed and supervised the research which forms the basis for this chapter.

References

- ADELMAN, H. & HAFTKA, R. 1986 Sensitivity analysis of discrete structural systems. *AIAA J.* **24** (5), 823–832.
- ANDERSON, W., NEWMAN, J., WHITFIELD, D. & NIELSEN, E. 1999 Sensitivity analysis for the navier-stokes equations on unstructured meshes using complex variables. In *AIAA Paper 99-3294*.
- LYNESS, J. 1967 Numerical algorithms based on the theory of complex variable. In *Proc. ACM 22nd Nat. Conf.*, pp. 124–134. Thomas Book Co., Washington DC.
- LYNESS, J. & MOLER, C. 1967 Numerical differentiation of analytic functions. *SIAM J. Numer. Anal.* **4**, 202–210.
- MARTINS, J., KROO, I. & ALONSO, J. 2000 An automated method for sensitivity analysis using complex variables. In *AIAA Paper 00-0689*.
- MARTINS, J., STURDZA, P. & ALONSO, J. 2001 The connection between the complex step derivative approximation and algorithmic differentiation. In *AIAA Paper 01-0921*.
- NEWMAN, J., ANDERSON, W. & WHITFIELD, D. 1998 Multidisciplinary sensitivity derivatives using complex variables. Tech. rep., Engineering Research Center Report mSSU-COE-ERC-98-08. Mississippi State University.
- SQUIRE, W. & TRAPP, G. 1998 Using complex variables to estimate derivatives of real functions. *SIAM Rev.* **40** (1), 110–112.

Chapter 4

Identification of time-periodic orbits in chaotic systems

4.1 Introduction

Many turbulent systems of physical interest, such as jets, wakes, or flow past wings (in air) and fins (in water), are dominated by the approximately time-periodic phenomenon of vortex shedding at the largest scales, with a large number of (chaotic) smaller vortices (over a range of scales) embedded. For control purposes, the precise location of these smaller vortices is relatively unimportant, though they do have a significant statistical influence on the motion of the large-scale vortices that is nontrivial and must be accounted for adequately. Identifying the fundamental periodic motions of the largest vortices in such flows will help us to better understand (as claimed by Waleffe (2001)), and ultimately optimize and control, such quasi-periodic turbulent flow systems, due to the simpler nature of these coherent structures.

Turbulent flows are typically characterized in numerical methods by marching the governing equations in time over a long time interval until all transients are “washed-out” of the domain and the flow approaches an approximately statistically time-periodic state. A common technique when attempting an optimization problem in this kind of systems is to state a cost function over a finite-time horizon, and to minimize this cost function via an

iterative gradient-based optimization over successive horizons (that is, receding-horizon model predictive control). Normally, the gradient is computed by iterative calculations of the state and adjoint fields, such as in Bewley *et al.* (2001). This technique, however, is in a way ill-suited for optimizing parameters (such as wall compliance properties, open-loop forcing schedule, etc...) to tune time-averaged statistics (as discussed in Luo (2004)), mainly because the cost function is a finite-time average. Another drawback of this technique for these systems is that the adjoint field grows exponentially (in backward time) in chaotic systems, so that the sensitivity of the parameter change is dominated by a narrow window of time. The exponential growth can be only partially compensated with exponential discount functions.

An alternative approach to this chaos control problem has been proposed by Bewley & Trenchea (2002). The idea is to modify the “skeleton” formed by the first several periodic orbits of the attractor, in order to favorably reconfigure the entire attractor. It is necessary in this approach to compute repeatedly the periodic orbits of the state and adjoint systems. Fig. 4.1 shows the skeleton of the Lorenz attractor formed by the first seven periodic orbits. The present work focuses on the extraction of the periodic orbits with period $T \leq T_{max}$, a necessary step in the optimization of chaotic systems when following this approach.

At the same time, the problem of finding *all* time-periodic orbits in a chaotic attractor with period less than or equal to T_{max} is an interesting problem that has attracted some interest in the literature for low-dimensional attractors. Viswanath (2003) used an extension of the Lindstedt-Poincaré method to obtain periodic orbits with 14 accurate digits. In high-dimensional systems, gradient-based methods are preferred. Kazantsev (1998) computes the gradient of a cost function measuring the periodicity of the orbits with the computation of the adjoint field, as is presented here in §4.3. There, a Newton method is used to perform the optimization. The method is checked by applying it to the low-dimensional Lorenz attractor. The difficulty when using the Newton method is the necessity of having a good initial guess of the periodic orbit. We avoid this problem by using conjugate gradient method (§4.3). Kawahara & Kida (2001) find two time-periodic

orbits in a plane Couette flow. Lan & Cvitanović (2004) use a variational principle to obtain periodic orbits of the Kuramoto-Sivashinsky system. Waleffe (2001) finds the so-called “exact coherent structures” in channel flow applying Newton’s method, obtaining the initial guess by homotopy from free-slip to no-slip boundary conditions.

Due to the ergodic nature of chaotic systems, a time-evolving simulation on a chaotic attractor often passes “close” to the many time-periodic orbits of period less than or equal to some T_{max} immersed within the attractor (it is well known that periodic orbits are densely embedded in the attractor of a chaotic system; see, e.g., Ott (2002)), as seen in Fig. 4.1. Thus, a time-evolving simulation of a chaotic system may be used to estimate its periodic orbits. An efficient algorithm for accomplishing this is presented in §4.2.

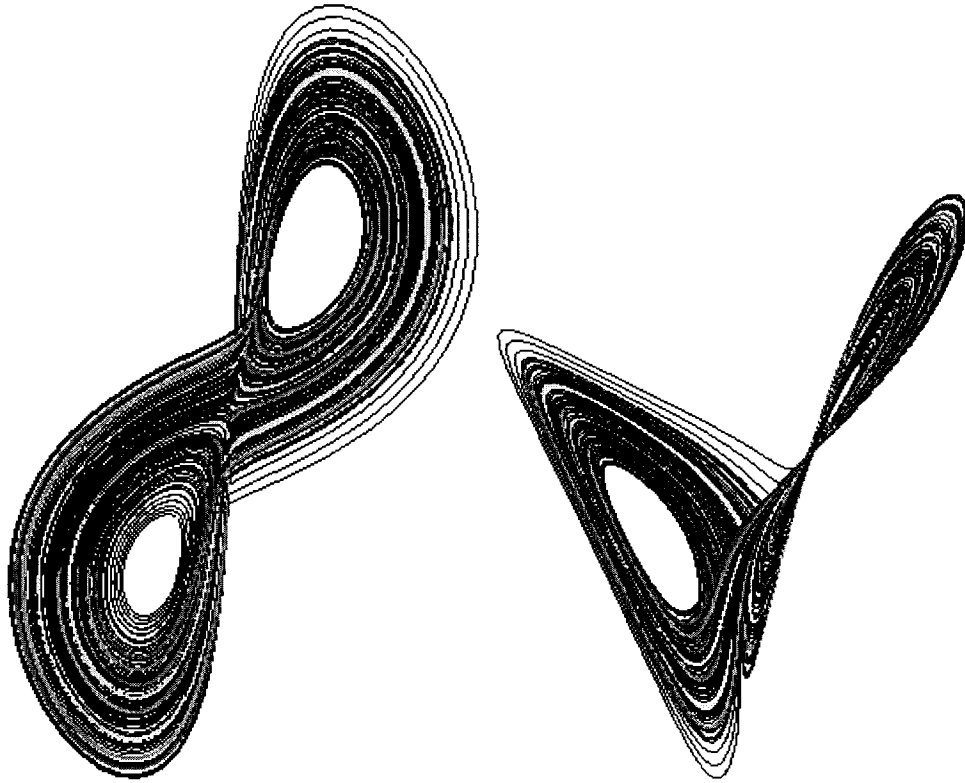


Figure 4.1: Two different views of the “skeleton” formed by the periodic orbits in the attractor (colored tubes), and a generic trajectory (thin black line). The generic trajectory passes close to the periodic orbits due to the ergodic nature of the system.

Once we have obtained several approximately time-periodic trajectories of period less than or equal to T_{max} , we may refine both the trajectories and their corresponding periods to see if there is an exactly time-periodic orbit of the dynamic system nearby. Two algorithms for accomplishing this refinement are presented. The first, in §4.3, considers iterates which all exactly (that is, to the numerical precision specified) satisfy the governing equation and are only approximately time periodic, with the degree to which the trajectory satisfies the periodic boundary conditions improved as the iteration proceeds. The second, described in §4.4, considers iterates which are all exactly time periodic and only approximately satisfy the governing equation, with the degree to which the periodic orbit satisfies the governing equation improved as the iteration proceeds.

Numerical results for the case of the Lorenz attractor are provided in §4.5, demonstrating the convergence of the initialization procedure described in §4.2 and comparing the performance of the two algorithms described in §4.3 and 4.4 for refining the approximately time-periodic trajectories. Note that the algorithms explored in this work all scale well with the dimension of the state, as they only require the explicit calculation of vectors, and never even require the solution of a system of simultaneous linear equations. Thus, the algorithm developed should extend readily to chaotic attractors of much higher state dimension, such as turbulent flows. This extension is currently under investigation, and will be reported in future work.

4.2 Obtaining several approximately time-periodic trajectories

As mentioned in the Introduction, we are interested in finding orbits in the chaotic attractor which are approximately time periodic, and in refining these approximations in both a time-evolving framework and a time-periodic framework. In this section we present an efficient algorithm for extracting up to j_{max} approximately time-periodic trajectories (with period $\tilde{T}$ less than or equal to T_{max}) from a time-evolving simulation of a chaotic system. In the following algorithm, we denote the state as $\mathbf{q}$, the integer enumer-

ating the timestep as k , and the (fixed) length of each timestep as Δt . To avoid duplicates in the resulting database, each of the approximately time-periodic trajectories found by the following algorithm will have a period different from the other approximately time-periodic trajectories found such that

$$|\tilde{T}_j - \tilde{T}_{j'}|/(\tilde{T}_j + \tilde{T}_{j'}) \geq T_{tol}/2 \quad \text{for } j \neq j'; \quad (4.1)$$

an appropriate value for the parameter T_{tol} might thus be about $T_{tol} = 0.01$. The algorithm follows:

- A. Perform a time-evolving calculation of the chaotic system $\mathbf{q}$ from random initial conditions. Once a sufficient amount of time has passed that this time-evolving simulation is near the attractor, continue the simulation for an additional time $t = T_{max}$, saving this trajectory segment in memory in an array of length $k_{max} = T_{max}/\Delta t$, denoting the k 'th timestep in this array as $\mathbf{s}_k$. Also initialize an (initially empty) table of length j_{max} to save the initial conditions, the time period, and the norm of the initial conditions minus the final conditions of the approximately time periodic trajectories found, denoting the j 'th entry of this table as $\{\tilde{\mathbf{q}}_j, \tilde{T}_j, \Delta\tilde{\mathbf{q}}_j\}$, and initialize $j_{saved} = 0$.
- B. Continue the simulation of the chaotic system $\mathbf{q}$ for an additional time step and increment $k \leftarrow k + 1$. Define $k_{new} = \text{mod}(k - 1, k_{max}) + 1$. Compute the $k' \in [1, k_{max}]$ which minimizes $\Delta\mathbf{q}_{new} = |\mathbf{q}_k - \mathbf{s}_{k'}|^2$, and compute the corresponding period $T_{new} = [\text{mod}(k_{new} - k' - 1, k_{max}) + 1]\Delta t$. Then check the following:

if $(\min_{j \in [1, j_{saved}]} [|T_{new} - \tilde{T}_j|/(\tilde{T}_j + T_{new})]) < T_{tol}/2$ then

if $(\Delta\mathbf{q}_{new} < \Delta\tilde{\mathbf{q}}_j)$ then

$$\{\tilde{\mathbf{q}}_j, \tilde{T}_j, \Delta\tilde{\mathbf{q}}_j\} \leftarrow \{\mathbf{s}_{k'}, T_{new}, \Delta\mathbf{q}_{new}\}$$

if $(\min_{j' \in [1, j_{saved}], j' \neq j} [|T_{new} - \tilde{T}_{j'}|/(\tilde{T}_{j'} + T_{new})]) < T_{tol}/2$ then

if $(\Delta\tilde{\mathbf{q}}_j < \Delta\tilde{\mathbf{q}}_{j'})$ then

$$\{\tilde{\mathbf{q}}_{j'}, \tilde{T}_{j'}, \Delta\tilde{\mathbf{q}}_{j'}\} \leftarrow \{\tilde{\mathbf{q}}_{j_{saved}}, \tilde{T}_{j_{saved}}, \Delta\tilde{\mathbf{q}}_{j_{saved}}\}$$

else

$$\{\tilde{\mathbf{q}}_j, \tilde{T}_j, \Delta\tilde{\mathbf{q}}_j\} \leftarrow \{\tilde{\mathbf{q}}_{j_{saved}}, \tilde{T}_{j_{saved}}, \Delta\tilde{\mathbf{q}}_{j_{saved}}\}$$

end

```


$$j_{saved} \leftarrow j_{saved} - 1$$

end
end
elseif ( $j_{saved} < j_{max}$ ) then

$$j_{saved} \leftarrow j_{saved} + 1$$


$$\{\tilde{\mathbf{q}}_{j_{saved}}, \tilde{T}_{j_{saved}}, \Delta \mathbf{q}_{j_{saved}}\} \leftarrow \{\mathbf{s}_{k'}, T_{new}, \Delta \mathbf{q}_{new}\}$$

elseif ( $\Delta \mathbf{q}_{new} < \max_{j \in [1, j_{max}]} \Delta \tilde{\mathbf{q}}_j$ ) then

$$\{\tilde{\mathbf{q}}_j, \tilde{T}_j, \Delta \mathbf{q}_j\} \leftarrow \{\mathbf{s}_{k'}, T_{new}, \Delta \mathbf{q}_{new}\}$$

end

```

C. Assign $\mathbf{s}_{k_{new}} \leftarrow \mathbf{q}_k$ and continue from step B until the algorithm gives diminishing returns.

Note that, if the state is high dimensional, it might be necessary to save the array containing the trajectory segment $\mathbf{s}_k$ for $k = [1, \dots, k_{max}]$ to disk. In this case, it is useful to save an auxiliary array of length k_{max} in memory containing one or more relevant statistics (e.g., the L_1 , L_2 , and/or L_∞ norm of the state and one or two representative elements of the state vector to indicate the phase of the system) at each timestep in order to characterize the state and eliminate the majority of the candidate k' before recalling select entries of $\mathbf{s}_{k'}$ from disk to calculate the k' which minimizes $\Delta \mathbf{q}_{new} = |\mathbf{q}_k - \mathbf{s}_{k'}|^2$ in step B of the above algorithm.

Two significant limitations of this algorithm are:

1. It is not exhaustive, and does not guarantee to find an approximately time-periodic trajectory near each of the actual time-periodic orbits of the system with period less than or equal to T_{max} . In our numerical tests, however, the algorithm was found to be fairly efficient.

2. It cannot find two different time-periodic orbits with periods which are closer than the tolerance set by Eq. (4.1) for the particular value of T_{tol} chosen. However, such orbits may often be found by simple symmetry arguments. For example, in the Lorenz problem, if a certain $\{q_1(t), q_2(t), q_3(t) \text{ on } t \in [0, T]\}$ is a time-periodic orbit of the system, then

$\{-q_1(t), -q_2(t), q_3(t) \text{ on } t \in [0, T]\}$ is also a time-periodic orbit of the system.

4.3 Refining the trajectories with time-evolving iterates

Given an initial guess for the periodic orbit itself, $\mathbf{q}(\tau)$ on $\tau \in [0, 2\pi]$, and for the frequency of the periodic orbit, $\Omega = 2\pi/T$, where the time $t = \tau/\Omega$, we now attempt to solve the problem governed by the equation

$$\mathbf{n}(\mathbf{q}, \Omega) = 0 \quad \text{on } \tau \in [0, 2\pi] \quad \text{with } \mathbf{q}(0) = \mathbf{q}(2\pi) \quad (4.2)$$

by appropriate determination of $\mathbf{q}(\tau)$ on $\tau \in [0, 2\pi]$ and Ω . In the Lorenz system, for example, the nonlinear operator $\mathbf{n}(\mathbf{q}, \Omega)$ is defined by

$$\mathbf{n}(\mathbf{q}, \Omega) = \begin{pmatrix} \Omega \frac{dq_1}{d\tau} - \sigma(q_2 - q_1) \\ \Omega \frac{dq_2}{d\tau} + q_2 + q_1 q_3 \\ \Omega \frac{dq_3}{d\tau} + bq_3 - q_1 q_2 + br \end{pmatrix}. \quad (4.3)$$

Problem Eq. (4.2) amounts to a high dimensional nonlinear root finding problem, which in general can be rather difficult.

We will present two techniques for restating this problem as a tractable minimization problem that might or might not lead to an exact solution of the actual problem of interest, as stated above. The first, outlined in this section, will iteratively refine both the initial conditions $\mathbf{q}(0) = \mathbf{q}_0$ and the frequency Ω in an attempt to solve the above problem. At each iteration of this approach, the system trajectory proceeding from $\mathbf{q}_0$ will be calculated directly, then the initial conditions $\mathbf{q}_0$ and frequency Ω adjusted with a gradient-based search in an attempt to satisfy the desired periodic conditions on the state, $\mathbf{q}(2\pi) = \mathbf{q}_0$. The algorithm developed in this section is based on the calculation of an adjoint field that is marched backward in time over the interval in question, and thus is quite similar to that used in weather forecasting (sometimes called 4D-VAR).

The second technique, outlined in §4.4, will iteratively refine the entire trajectory $\mathbf{q}(\tau)$ on $\tau \in [0, 2\pi]$, in addition to the frequency Ω , in an attempt to solve the above

problem. At each iteration of this approach, the system trajectory will be constructed from a Fourier series, so the periodic conditions on the state $\mathbf{q}(0) = \mathbf{q}(2\pi)$ will be satisfied exactly; with this approach, the trajectory $\mathbf{q}(\tau)$ on the entire interval $\tau \in [0, 2\pi]$ and the frequency Ω will be adjusted at each iteration with a gradient-based search in an attempt to satisfy the desired governing equation, $\mathbf{n}(\mathbf{q}, \Omega) = 0$ on $\tau \in [0, 2\pi]$.

4.3.1 Derivation of gradients

Defining the *control variable* ϕ , which contains the parameters to optimize, in this case as $\phi = \{\mathbf{q}_0; \Omega\}$, we seek the best ϕ to minimize the *cost function*

$$J(\phi) = \frac{1}{2} |\mathbf{q}(2\pi) - \mathbf{q}_0|^2, \quad (4.4)$$

which is a measurement of the periodicity of the orbit. This cost function is the same as the one in Kazantsev (1998), and the method to obtain the gradient they use is the same as here, although the descending procedure they use is the truncated Newton method, which might be a bit more problematic in high-dimensional systems, due to the need of a "good" initial guess, which is more difficult to find in such systems. In the present work we use Conjugate Gradient method as the descent procedure.

The *state variable* $\mathbf{q}$ is related to ϕ by the *state equation*, which in this case takes the form

$$\mathbf{n}(\mathbf{q}, \Omega) = 0 \quad \text{on } \tau \in [0, 2\pi] \quad \text{with } \mathbf{q}(0) = \mathbf{q}_0, \quad (4.5)$$

where again, in the example of the Lorenz problem, the nonlinear operation $\mathbf{n}(\mathbf{q}, \Omega)$ is given in Eq. (4.3). Note that the difference between Eq. (4.5) and Eq. (4.2) is that in Eq. (4.5) periodicity is not imposed. The selection of the norm in the cost function may be tuned in order to emphasize the components of the system of primary concern in the optimization problem. For the Lorenz problem, we might define this norm as simply

$$|\mathbf{q}(2\pi) - \mathbf{q}_0|^2 \triangleq (\mathbf{q}(2\pi) - \mathbf{q}_0)^T (\mathbf{q}(2\pi) - \mathbf{q}_0), \quad (4.6)$$

where $()^T$ denotes transpose.

Consider a small perturbation ϕ' to the control variable ϕ and the concomitant small perturbation $\mathbf{q}'$ to the state variable $\mathbf{q}$. Substituting $\phi \leftarrow \phi + \phi'$ (that is, $\{\mathbf{q}_0; \Omega\} \leftarrow \{\mathbf{q}_0; \Omega\} + \{\mathbf{q}'_0; \Omega'\}$) and $\mathbf{q} \leftarrow \mathbf{q} + \mathbf{q}'$ in Eq. (4.5), neglecting terms that are quadratic in the perturbation, and applying Eq. (4.5), we may write the *perturbation equation* as

$$N_{\mathbf{q}}\mathbf{q}' + N_{\Omega}\Omega' = 0 \quad \text{on } \tau \in [0, 2\pi] \quad \text{with } \mathbf{q}'(0) = \mathbf{q}'_0, \quad (4.7)$$

where, from Eq. (4.3), the operators on the LHS may be written for the case of the Lorenz equations as

$$N_{\mathbf{q}}\mathbf{q}' = \begin{pmatrix} \Omega \frac{dq'_1}{d\tau} - \sigma(q'_2 - q'_1) \\ \Omega \frac{dq'_2}{d\tau} + q'_2 + q_1 q'_3 + q_3 q'_1 \\ \Omega \frac{dq'_3}{d\tau} + b q'_3 - q_1 q'_2 - q_2 q'_1 \end{pmatrix} = \begin{pmatrix} \Omega \frac{d}{d\tau} + \sigma & -\sigma & 0 \\ q_3 & \Omega \frac{d}{d\tau} + 1 & q_1 \\ -q_2 & -q_1 & \Omega \frac{d}{d\tau} + b \end{pmatrix} \begin{pmatrix} q'_1 \\ q'_2 \\ q'_3 \end{pmatrix}, \quad (4.8a)$$

$$N_{\Omega}\Omega' = \begin{pmatrix} \Omega' \frac{dq_1}{d\tau} \\ \Omega' \frac{dq_2}{d\tau} \\ \Omega' \frac{dq_3}{d\tau} \end{pmatrix} = \begin{pmatrix} \frac{dq_1}{d\tau} \\ \frac{dq_2}{d\tau} \\ \frac{dq_3}{d\tau} \end{pmatrix} \Omega'. \quad (4.8b)$$

Note also that, for the Lorenz problem, combining Eq. (4.4) and Eq. (4.6) leads to

$$J(\phi + \phi') \approx J(\phi) + (\mathbf{q}(2\pi) - \mathbf{q}_0)^T (\mathbf{q}'(2\pi) - \mathbf{q}'_0), \quad (4.9)$$

where $\mathbf{q}'(2\pi)$ is linearly related to $\phi' = \{\mathbf{q}'_0; \Omega'\}$ by the perturbation equation Eq. (4.7).

In order to re-express the linear relationship between $\mathbf{q}'(2\pi)$ and $\phi' = \{\mathbf{q}'_0; \Omega'\}$ given by Eq. (4.7), and thereby rewrite Eq. (4.9) in a more useful manner that only depends on the control perturbation $\phi' = \{\mathbf{q}'_0; \Omega'\}$, we make use of an adjoint identity. Defining the inner product $\langle \mathbf{r}, \mathbf{q}' \rangle = \int_0^{2\pi} \mathbf{r}^* \mathbf{q}' d\tau$, we may write the identity

$$\langle \mathbf{r}, N_{\mathbf{q}}\mathbf{q}' \rangle = \langle N_{\mathbf{q}}^* \mathbf{r}, \mathbf{q}' \rangle + b. \quad (4.10)$$

where, noting Eq. (4.8a) and simply integrating by parts, it follows for the Lorenz problem

that

$$\begin{aligned}
 N_{\mathbf{q}}^* \mathbf{r} &= \begin{pmatrix} -\Omega \frac{d}{d\tau} + \sigma & q_3 & -q_2 \\ -\sigma & -\Omega \frac{d}{d\tau} + 1 & -q_1 \\ 0 & q_1 & -\Omega \frac{d}{d\tau} + b \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \\
 &= \begin{pmatrix} -\Omega \frac{dr_1}{d\tau} + \sigma r_1 + q_3 r_2 - q_2 r_3 \\ -\Omega \frac{dr_2}{d\tau} - \sigma r_1 - q_1 r_3 + r_2 \\ -\Omega \frac{dr_3}{d\tau} + q_1 r_2 + b r_3 \end{pmatrix}, \tag{4.11a}
 \end{aligned}$$

$$b = \Omega \mathbf{r}^T \mathbf{q}' \Big|_{\tau=2\pi} - \Omega \mathbf{r}^T \mathbf{q}' \Big|_{\tau=0}. \tag{4.11b}$$

We now define the *adjoint variable* $\mathbf{r}$ according to the following *adjoint equation*

$$N_{\mathbf{q}}^* \mathbf{r} = 0 \quad \text{on } \tau \in [0, 2\pi] \quad \text{with } \mathbf{r}(2\pi) = \mathbf{q}(2\pi) - \mathbf{q}_0. \tag{4.12}$$

Applying Eq. (4.12) and Eq. (4.7) to the identity given in Eq. (4.10) yields

$$-\langle \mathbf{r}, N_{\Omega} \Omega' \rangle = \Omega (\mathbf{q}(2\pi) - \mathbf{q}_0)^T \mathbf{q}'(2\pi) - \Omega \mathbf{r}^T(0) \mathbf{q}'_0,$$

and thus we may rewrite Eq. (4.9) as

$$J(\phi + \phi') \approx J(\phi) + [\mathbf{r}(0) - \mathbf{q}(2\pi) + \mathbf{q}_0]^* \mathbf{q}'_0 - \left[\frac{1}{\Omega} \int_0^{2\pi} N_{\Omega}^* \mathbf{r} d\tau \right]^* \Omega', \tag{4.13}$$

where, noting Eq. (4.8b),

$$N_{\Omega}^* \mathbf{r} = \begin{pmatrix} \frac{dq_1}{d\tau} & \frac{dq_2}{d\tau} & \frac{dq_3}{d\tau} \end{pmatrix} \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} = \frac{dq_1}{d\tau} r_1 + \frac{dq_2}{d\tau} r_2 + \frac{dq_3}{d\tau} r_3. \tag{4.14}$$

Performing a Taylor series expansion of $J(\phi + \phi')$ near ϕ , we may write

$$J(\phi + \phi') \approx J(\phi) + \left\langle \frac{DJ}{D\phi}, \phi' \right\rangle. \tag{4.15}$$

We thus see that the "gradient" $DJ/D\phi = \{DJ/D\mathbf{q}_0; DJ/D\Omega\}$ is defined by the inner product selected in Eq. (4.15). Appropriate selection of this inner product is a way of *preconditioning* the optimization process. The selection of the inner product defining the gradient may be tuned as explained in Protas *et al.* (2004) (in the case of PDE systems, by incorporating spatial filters) in order to focus the initial steps of the optimization on selected components (in the PDE case, length scales) of the system. For the Lorenz problem, we might define the gradient such that

$$\left\langle \frac{DJ}{D\phi}, \phi' \right\rangle \triangleq \left[\frac{DJ}{D\mathbf{q}_0} \right]^T \mathbf{q}'_0 + \rho \left[\frac{DJ}{D\Omega} \right]^T \Omega'. \quad (4.16)$$

We are free to adjust the parameter ρ in this definition, which ultimately has the effect of increasing or decreasing the relative speed at which Ω will be adjusted during the optimization of ϕ . Combining Eq. (4.15) and Eq. (4.16) and comparing with Eq. (4.13) leads immediately to

$$\frac{DJ}{D\mathbf{q}_0} = \mathbf{r}(0) - \mathbf{q}(2\pi) + \mathbf{q}_0 \quad \text{and} \quad \frac{DJ}{D\Omega} = -\frac{1}{\rho\Omega} \int_0^{2\pi} N_{\Omega}^* \mathbf{r} d\tau. \quad (4.17)$$

4.3.2 Estimation of descent parameter

We now consider the problem of updating the control ϕ in some direction ϕ' (gradient direction in the *Steepest Descent* method, or gradient direction modified by a momentum term in the *Conjugate Gradient* method), scaling this update by a factor α which we are free to choose. We would like to be able to select the value of the scalar α to minimize the resulting cost function, $J(\phi + \alpha\phi')$; in other words, we would like to perform a *line minimization* with respect to the parameter α along the direction of ϕ' . Because this line minimization is an iterative procedure, where multiple function evaluations are required (one function evaluation on a dynamic system such as the Lorenz attractor is a time evolution of the system over T_{max}), most of the computer time is spent here. It is important for this reason to have a method to estimate α such that it is as close to the optimal value as possible. A new method for estimating the parameter α , explained in the Appendix, and which has proven to save a considerable amount of computation time, is used. We briefly summarize the method.

The minimization may be obtained by considering the Taylor series expansion for $J(\phi + \alpha\phi')$ near $J(\phi)$ and keeping terms up to those which are quadratic in the perturbations, then setting $dJ(\phi + \alpha\phi')/d\alpha = 0$ in the resulting expression and solving for α , as shown below.

The Taylor series expansion for $J(\phi + \alpha\phi')$, up to second order terms is

$$J(\phi + \alpha\phi') = J(\phi) + \alpha J'(\phi; \phi') + \frac{\alpha^2}{2} J''(\phi; \phi') + O(\alpha^3). \quad (4.18)$$

It is straightforward now to obtain the value of α that would minimize the cost function in case of having a quadratic function (which becomes a good approximation as the control parameters get close to the values that minimize the cost function):

$$\alpha_q = -\frac{J'(\phi; \phi')}{J''(\phi; \phi')}. \quad (4.19)$$

The perturbation cost function J' can be obtained by perturbing the cost function in Eq. (4.4):

$$J'(\phi; \phi') = (\mathbf{q}(2\pi) - \mathbf{q}_0)^T (\mathbf{q}'(2\pi) - \mathbf{q}'_0), \quad (4.20)$$

and the second perturbation cost function J'' by perturbing Eq. (4.20), and taking into account that $\mathbf{q}_0'' = 0$:

$$J''(\phi; \phi') = (\mathbf{q}(2\pi) - \mathbf{q}_0)^T \mathbf{q}''(2\pi) + (\mathbf{q}'(2\pi) - \mathbf{q}'_0)^T (\mathbf{q}'(2\pi) - \mathbf{q}'_0). \quad (4.21)$$

The unknowns in the perturbation cost functions given by Eq. (4.20) and Eq. (4.21) are $\mathbf{q}'(2\pi)$ and $\mathbf{q}''(2\pi)$, which can be obtained from two evolution equations, the perturbation equation Eq. (4.7) and the corresponding second perturbation equation

$$N_{\mathbf{q}} \mathbf{q}'' = -2 \left(\frac{\partial N_{\Omega}}{\partial \mathbf{q}} \mathbf{q}' \right) \Omega' - \left(\frac{\partial N_{\mathbf{q}}}{\partial \mathbf{q}} \mathbf{q}' \right) \mathbf{q}' \quad \text{on } \tau \in [0, 2\pi] \quad \text{with } \mathbf{q}''(0) = 0. \quad (4.22)$$

Notice that perturbing Eq. (4.7) gives a term in Ω'' which is 0, leading to Eq. (4.22). $\mathbf{n}(\mathbf{q}, \Omega)$ is defined in Eq. (4.3), $N_{\mathbf{q}}$ and N_{Ω} , in Eq. (4.8a) and Eq. (4.8b),

$$\left(\frac{\partial N_{\Omega}}{\partial \mathbf{q}} \mathbf{q}' \right) \Omega' = \begin{pmatrix} \frac{dq'_1}{d\tau} \\ \frac{dq'_2}{d\tau} \\ \frac{dq'_3}{d\tau} \end{pmatrix} \Omega', \quad \text{and} \quad \left(\frac{\partial N_{\mathbf{q}}}{\partial \mathbf{q}} \mathbf{q}' \right) \mathbf{q}' = \begin{pmatrix} 0 & 0 & 0 \\ q'_3 & 0 & q'_1 \\ -q'_2 & -q'_1 & 0 \end{pmatrix} \begin{pmatrix} q'_1 \\ q'_2 \\ q'_3 \end{pmatrix}.$$

The value of the descent parameter α_q obtained by Eq. (4.19) is compared at every iteration with the value of α that would correspond to the minimum of the cost function obtained by inverse parabolic interpolation, as will be explained in more detail in §3.3. After both values are close to a certain precision during a specified number of iterations, the value given by Eq. (4.19) is taken as the optimum value, saving in this way computational time. However the accuracy of this assumption has to be checked every few iterations.

4.3.3 Summary of numerical algorithm

We now summarize the Polak-Ribiere variant of the conjugate gradient algorithm to minimize the cost function $J(\phi)$ when $\phi = \{\mathbf{q}_0; \Omega\}$, as given in Eq. (4.4). For further discussion of the conjugate gradient algorithm used here, please see, e.g., Chapter 5 of Bewley (Preprint).

- A. Obtain an initial guess for the “control” variable $\phi = \{\mathbf{q}_0; \Omega\}$ corresponding to a time-periodic orbit of the system in Eq. (4.2)-Eq. (4.3) using, for example, the algorithm presented in §2, and initialize $i = 1$ and $i_{reset} = 20$.
- B. Calculate the state $\mathbf{q}$ from the state equation Eq. (4.5) (a forward march in time on $\tau = 0 \rightarrow 2\pi$).
- C. Calculate the adjoint $\mathbf{r}$ from the adjoint equation Eq. (4.12) (a backward march in time on $\tau = 2\pi \rightarrow 0$).
- D. Calculate the gradient $\mathbf{g} \triangleq DJ/D\phi = \{DJ/D\mathbf{q}_0; DJ/D\Omega\}$ using Eq. (4.17).
- E. Compute the control update direction ϕ' according to the Polak-Ribiere variant of the conjugate gradient algorithm:

$$res = \mathbf{g}^T \mathbf{g}$$

if $i = 1$ then

$$\phi' = -\mathbf{g}$$

else

$$\beta = \frac{res - \mathbf{g}_{old}^T \mathbf{g}}{res_{old}}$$

$$\phi' \leftarrow -\mathbf{g} + \beta\phi'$$

end

$$res_{old} = res$$

$$\mathbf{g}_{old} = \mathbf{g}$$

- F. Given $\mathbf{q}$ from step B and $\phi' = \{\mathbf{q}'_0; \Omega'\}$ from step E, calculate $\mathbf{q}'$ from Eq. (4.7) and then $\mathbf{q}''$ from Eq. (4.22) (both of these calculations are forward marches in time on $\tau = 0 \rightarrow 2\pi$). Noting the definition of $J'(\phi; \phi')$ and $J''(\phi, \phi')$ in Eq. (4.20) and Eq. (4.21), calculate α_q using the formula given in Eq. (4.19).
- G. Take the initial guess for $\alpha = \alpha_q$ from step F and a guess $\{0, \alpha_q, 1.6\alpha_q\}$ for a triplet of values for α that brackets the actual minimum of $J(\phi + \alpha\phi')$. Use mnbrak and Brent's method (see Press *et al.* (1992)) to refine the value of α that minimizes $J(\phi + \alpha\phi')$ to some prespecified tolerance (say, 1%) by repeated calculation of $J(\phi + \alpha\phi')$, using Eq. (4.4) and Eq. (4.5), for several trial values of α .
- H. Using the converged value of α from step G, update $\phi \leftarrow \phi + \alpha\phi'$, $i \leftarrow \text{mod}(i + 1, i_{reset})$ and repeat from step B until convergence.

Note that step G of the above algorithm insures convergence of the optimization even when the initial guess of ϕ is poor. As convergence is approached, the value of α_q determined in step F from the local quadratic approximation becomes an increasingly accurate approximation of the minimizer of $J(\phi + \alpha\phi')$, and step G may eventually be dropped altogether, taking $\alpha = \alpha_q$ in step H. A simple check to determine whether or not step G is necessary is to compute three test values of $J(\phi + \alpha\phi')$ and then perform an inverse parabolic interpolation (see, e.g., §5.2.3 of Bewley (Preprint))

$$\alpha_t = \frac{1}{2} \frac{y_0(x_1 + x_2)(x_1 - x_2) + y_1(x_0 + x_2)(x_2 - x_0) + y_2(x_0 + x_1)(x_0 - x_1)}{y_0(x_1 - x_2) + y_1(x_2 - x_0) + y_2(x_0 - x_1)},$$

where, e.g., $\{x_0, y_0\} = \{0, J(\phi)\}$, $\{x_1, y_1\} = \{\alpha_q, J(\phi + \alpha_q\phi')\}$, and $\{x_2, y_2\} = \{1.6\alpha_q, J(\phi + 1.6\alpha_q\phi')\}$. If the test

$$|\alpha_t - \alpha_q|/|\alpha_t + \alpha_q| < \alpha_{tol}/2$$

passes (taking, say, $\alpha_{tol} = 0.01$), then step G may be safely skipped. Also note that, as this test increases or decreases confidence in the local quadratic approximation of $J(\phi)$

between the current value of ϕ and the minimizer, it may be used to increase or decrease i_{reset} accordingly.

4.4 Refining the trajectories with time-periodic iterates

In this section we present an alternative technique for solving the problem described at the beginning of §3, by iteration inside a time-periodic framework. At each iteration of the approach developed in this section, the system trajectory will be constructed from a Fourier series, so that periodic conditions on the state $\mathbf{q}(0) = \mathbf{q}(2\pi)$ will be satisfied exactly; the trajectory $\mathbf{q}(\tau)$ on $\tau \in [0, 2\pi]$ and the frequency Ω will be adjusted at each iteration with a gradient-based search in an attempt to satisfy the desired governing equation, $\mathbf{n}(\mathbf{q}, \Omega) = 0$ on $\tau \in [0, 2\pi]$.

4.4.1 Derivation of gradients

Defining the *control variable* ϕ in this case as $\phi = \{\mathbf{q}(\tau) \text{ on } \tau \in [0, 2\pi]; \Omega\}$, we seek the best ϕ to minimize the *cost function*

$$J(\phi) = \frac{1}{2} |\mathbf{n}(\mathbf{q}, \Omega)|^2, \quad (4.23)$$

where again, in the example of the Lorenz problem, the nonlinear operation $\mathbf{n}(\mathbf{q}, \Omega)$ is given in Eq. (4.3). This cost function measures how well a given trajectory $\mathbf{q}(\tau)$ satisfies the equation governing the system. The selection of the norm in the cost function may be tuned (by incorporating filters in time and, in the case of PDE systems, space) in order to emphasize the time scales (and, in the PDE case, length scales) of primary concern in the optimization problem. For the Lorenz problem, we may take simply

$$|\mathbf{n}|^2 \triangleq \int_0^{2\pi} \mathbf{n}^T \mathbf{n} d\tau. \quad (4.24)$$

Note that, although the control variable is now much higher dimensional, this optimization problem is fundamentally simpler than that described in §3.1, as the control variable ϕ appears *directly* in the cost function in Eq. (4.23), and is not related to the variables

appearing in the cost function by a dynamic evolution equation (a “state equation”). Thus, we can calculate the gradient directly, without evaluating an auxiliary dynamic evolution equation (an “adjoint equation”), as shown below.

Consider a small perturbation ϕ' to the control variable ϕ . Substituting $\phi \leftarrow \phi + \phi'$ (that is, $\{\mathbf{q}(\tau) \text{ on } \tau \in [0, 2\pi]; \Omega\} \leftarrow \{\mathbf{q}(\tau) \text{ on } \tau \in [0, 2\pi]; \Omega\} + \{\mathbf{q}'(\tau) \text{ on } \tau \in [0, 2\pi]; \Omega'\}$) in Eq. (4.23), applying Eq. (4.24), and neglecting terms that are quadratic in the perturbation, it follows that

$$\begin{aligned} J(\phi + \phi') &\approx J(\phi) + \int_0^{2\pi} \mathbf{n}^T (N_{\mathbf{q}} \mathbf{q}' + N_{\Omega} \Omega') d\tau = \\ &= J(\phi) + \int_0^{2\pi} (N_{\mathbf{q}}^* \mathbf{n})^T \mathbf{q}' d\tau + \left(\int_0^{2\pi} N_{\Omega}^* \mathbf{n} d\tau \right)^T \Omega', \end{aligned} \quad (4.25)$$

where, for the Lorenz example, the linear operations $N_{\mathbf{q}} \mathbf{q}'$ and $N_{\Omega} \Omega'$ are defined as in Eq. (4.8a) and Eq. (4.8b), and the linear operators $N_{\mathbf{q}}^* \mathbf{n}$ and $N_{\Omega}^* \mathbf{n}$ are defined as in Eq. (4.11a) and Eq. (4.14). Note that the boundary terms in Eq. (4.25) due to integration by parts of the $d/d\tau$ terms in $N_{\mathbf{q}}$ vanish immediately due to the time periodicity enforced (by construction) on $\mathbf{q}$ at each iteration [that is, $\mathbf{q}(0) = \mathbf{q}(2\pi)$, $\mathbf{q}'(0) = \mathbf{q}'(2\pi)$, $\mathbf{n}|_{\tau=0} = \mathbf{n}|_{\tau=2\pi}$, etc.]. Performing a Taylor series expansion of $J(\phi + \phi')$ near ϕ , we may write

$$J(\phi + \phi') \approx J(\phi) + \left\langle \frac{DJ}{D\phi}, \phi' \right\rangle. \quad (4.26)$$

We thus see that the “gradient” $DJ/D\phi = \{DJ/D\mathbf{q}(\tau) \text{ on } \tau \in [0, 2\pi]; DJ/D\Omega\}$ is defined by the inner product selected in Eq. (4.26). Appropriate selection of this inner product is a way of *preconditioning* the optimization process. The selection of the inner product defining the gradient may be tuned (by incorporating filters in time and, in the case of PDE systems, space) in order to focus the initial steps of the optimization on a certain range of time scales (and, in the PDE case, length scales) (see Protas *et al.* (2004)). For the Lorenz problem, we might define the gradient such that

$$\left\langle \frac{DJ}{D\phi}, \phi' \right\rangle \triangleq \int_0^{2\pi} \left(\frac{DJ}{D\mathbf{q}} \right)^T \mathbf{q}' d\tau + \rho \left(\frac{DJ}{D\Omega} \right)^T \Omega'. \quad (4.27)$$

We are free to adjust the parameter ρ in this definition, which ultimately has the effect of increasing or decreasing the relative speed at which Ω will be adjusted during the opti-

mization of ϕ . Thus, combining Eq. (4.26) and Eq. (4.27) and comparing with Eq. (4.25), we identify

$$\frac{DJ}{D\mathbf{q}(\tau)} = N_{\mathbf{q}}^* \mathbf{n}(\mathbf{q}(\tau), \Omega) \quad \text{on } \tau \in [0, 2\pi], \quad \text{and} \quad \frac{DJ}{D\Omega} = \frac{1}{\rho} \int_0^{2\pi} N_{\Omega}^* \mathbf{n}(\mathbf{q}(\tau), \Omega) d\tau. \quad (4.28)$$

4.4.2 Estimation of descent parameter

The estimation of the descent parameter in this case follows the same procedure as in §3.2. Simply writing the Taylor series expansion for $J(\phi + \alpha\phi')$ near $J(\phi)$ and keeping terms up to those quadratic in α gives

$$J(\phi + \alpha\phi') \approx J(\phi) + \alpha \int_0^{2\pi} \mathbf{n}^T (N_{\mathbf{q}} \mathbf{q}' + N_{\Omega} \Omega') d\tau + \quad (4.29a)$$

$$\begin{aligned} & + \frac{\alpha^2}{2} \int_0^{2\pi} \left[(N_{\mathbf{q}} \mathbf{q}' + N_{\Omega} \Omega')^T (N_{\mathbf{q}} \mathbf{q}' + N_{\Omega} \Omega') \right] d\tau \\ & + \frac{\alpha^2}{2} \int_0^{2\pi} \left[\mathbf{n}^T \left(2 \left(\frac{\partial N_{\Omega}}{\partial q} \mathbf{q}' \right) \Omega' + \left(\frac{\partial N_{\mathbf{q}}}{\partial q} \mathbf{q}' \right) \mathbf{q}' \right) \right] d\tau \\ & \triangleq J(\phi) + \alpha J'(\phi; \phi') + \frac{\alpha^2}{2} J''(\phi; \phi'), \end{aligned} \quad (4.29b)$$

where, for the Lorenz case, all operators are defined in §3. Note that in the time-periodic case, both $\mathbf{q}''$ and Ω'' are 0, since $\mathbf{q}$ and Ω are the control parameters, and the perturbations $\mathbf{q}'$ and Ω' are fixed. Minimizing this expression with respect to α gives, as before,

$$\frac{dJ(\phi + \alpha\phi')}{d\alpha} = 0 \quad \Rightarrow \quad \alpha_q = -\frac{J'(\phi)}{J''(\phi)}, \quad (4.29c)$$

where, by Eq. (4.29a)-Eq. (4.29b), J' and J'' are given by

$$J'(\phi; \phi') = \int_0^{2\pi} \mathbf{n}^T (N_{\mathbf{q}} \mathbf{q}' + N_{\Omega} \Omega') d\tau, \quad (4.30)$$

$$\begin{aligned} J''(\phi; \phi') = \int_0^{2\pi} & \left[(N_{\mathbf{q}} \mathbf{q}' + N_{\Omega} \Omega')^T (N_{\mathbf{q}} \mathbf{q}' + N_{\Omega} \Omega') + \right. \\ & \left. + \mathbf{n}^T \left(2 \left(\frac{\partial N_{\Omega}}{\partial q} \mathbf{q}' \right) \Omega' + \left(\frac{\partial N_{\mathbf{q}}}{\partial q} \mathbf{q}' \right) \mathbf{q}' \right) \right] d\tau. \end{aligned} \quad (4.31)$$

4.4.3 Summary of numerical algorithm

We now summarize the Polak-Ribiere variant of the conjugate gradient algorithm to minimize the cost function $J(\phi)$ when $\phi = \{\mathbf{q}(\tau) \text{ on } \tau \in [0, 2\pi]; \Omega\}$, as given in Eq. (4.23).

The algorithm is a bit simpler than that presented in §3.3, though the control variable is now much higher dimensional.

- A. Obtain an initial guess for the “control” variable $\phi = \{\mathbf{q}(\tau) \text{ on } \tau \in [0, 2\pi]; \Omega\}$ corresponding to a time-periodic orbit of the the system Eq. (4.2)-Eq. (4.3). Such an initial guess may be obtained using, for example, the algorithm presented in §2 to determine an appropriate $\{\mathbf{q}_0, \Omega\}$, then marching the state equation for $\mathbf{q}$ over the interval $\tau \in [0, 2\pi]$, then expanding the resulting trajectory using a Fourier series, keeping a finite number of terms in the expansion (perhaps gradually filtering out the higher frequencies). Initialize $i = 1$ and $i_{reset} = 25$.
- B. Calculate $\mathbf{n}(\mathbf{q}, \Omega)$ (for the case of Lorenz, from Eq. (4.3)) on the interval $\tau = [0, 2\pi]$.
- C. Calculate the gradient $\mathbf{g} \triangleq DJ/D\phi = \{DJ/D\mathbf{q}_0; DJ/D\Omega\}$ directly using Eq. (4.28).
- D. Compute the control update direction ϕ' according to the Polak-Ribiere variant of the conjugate gradient algorithm:

$$res = \mathbf{g}^T \mathbf{g}$$

if $i = 1$ then

$$\phi' = -\mathbf{g}$$

else

$$\beta = \frac{res - \mathbf{g}_{old}^T \mathbf{g}}{res_{old}}$$

$$\phi' \leftarrow -\mathbf{g} + \beta \phi'$$

end

$$res_{old} = res$$

$$\mathbf{g}_{old} = \mathbf{g}$$

- E. Given the control $\phi = \{\mathbf{q}(\tau) \text{ on } \tau \in [0, 2\pi]; \Omega\}$ and $\phi' = \{\mathbf{q}'(\tau) \text{ on } \tau \in [0, 2\pi]; \Omega'\}$ from step D, and noting the definition of $J'(\phi; \phi')$ and $J''(\phi, \phi')$ in Eq. (4.29b), calculate α_q using the formula given in Eq. (4.29c).

- F. Take the initial guess for $\alpha = \alpha_q$ from step E and a guess $\{0, \alpha_q, 1.6\alpha_q\}$ for a triplet of values for α that brackets the minimum of $J(\phi + \alpha\phi')$. Use `mnbrak` and Brent's method (see Press *et al.* (1992)) to refine the value of α that minimizes $J(\phi + \alpha\phi')$ to some prespecified tolerance (say, 1%) by repeated calculation of $J(\phi + \alpha\phi')$, using Eq. (4.4) and Eq. (4.5), for several trial values of α .
- G. Using the converged value of α from step F, update $\phi \leftarrow \phi + \alpha\phi'$, $i \leftarrow \text{mod}(i+1, i_{\text{reset}})$ and repeat from step B until convergence.

Note that, as discussed in detail in the last paragraph of §3.3, step F of the above algorithm may be dropped as convergence is approached.

4.5 Numerical results

The methods and algorithms presented in the previous sections have been tested in the Lorenz system, the evolution of which is governed by $\mathbf{n}(\mathbf{q}, \Omega) = 0$, with $\mathbf{n}(\mathbf{q}, \Omega)$ given in Eq. (4.3). The value given to the parameters in that equation is $\sigma = 10$, $r = 28$, $b = 8/3$, the same as in the original paper by Lorenz (1963). In Viswanath (2003), with the same values, some periodic orbits were computed with an accuracy of 14 digits, and will be used here as a reference.

As shown in Viswanath (2003), the system under consideration has periodic orbits with 7 different periods such that $T < 4$. The periods corresponding to these orbits are given in the first column of Table 4.1. 2D plots of the corresponding orbits are shown in Fig. 4.2. Note that the Lorenz system has a symmetry such that the transformation $q_1 \rightarrow -q_1$, $q_2 \rightarrow -q_2$, $q_3 \rightarrow q_3$ does not change the system, so that for any non-symmetric orbit shown in Fig. 4.2 there is a mirrored orbit which is also periodic and has the same period (as mentioned earlier, the method to obtain the approximately time-periodic orbits described in §4.2 does not capture two orbits with the same period). Table 4.1 shows the period of the orbits after convergence of the three methods described. The second column shows the estimation obtained with the algorithm described in §4.2. It can be seen that, although not exactly, the periods of the approximate orbits match very well the values

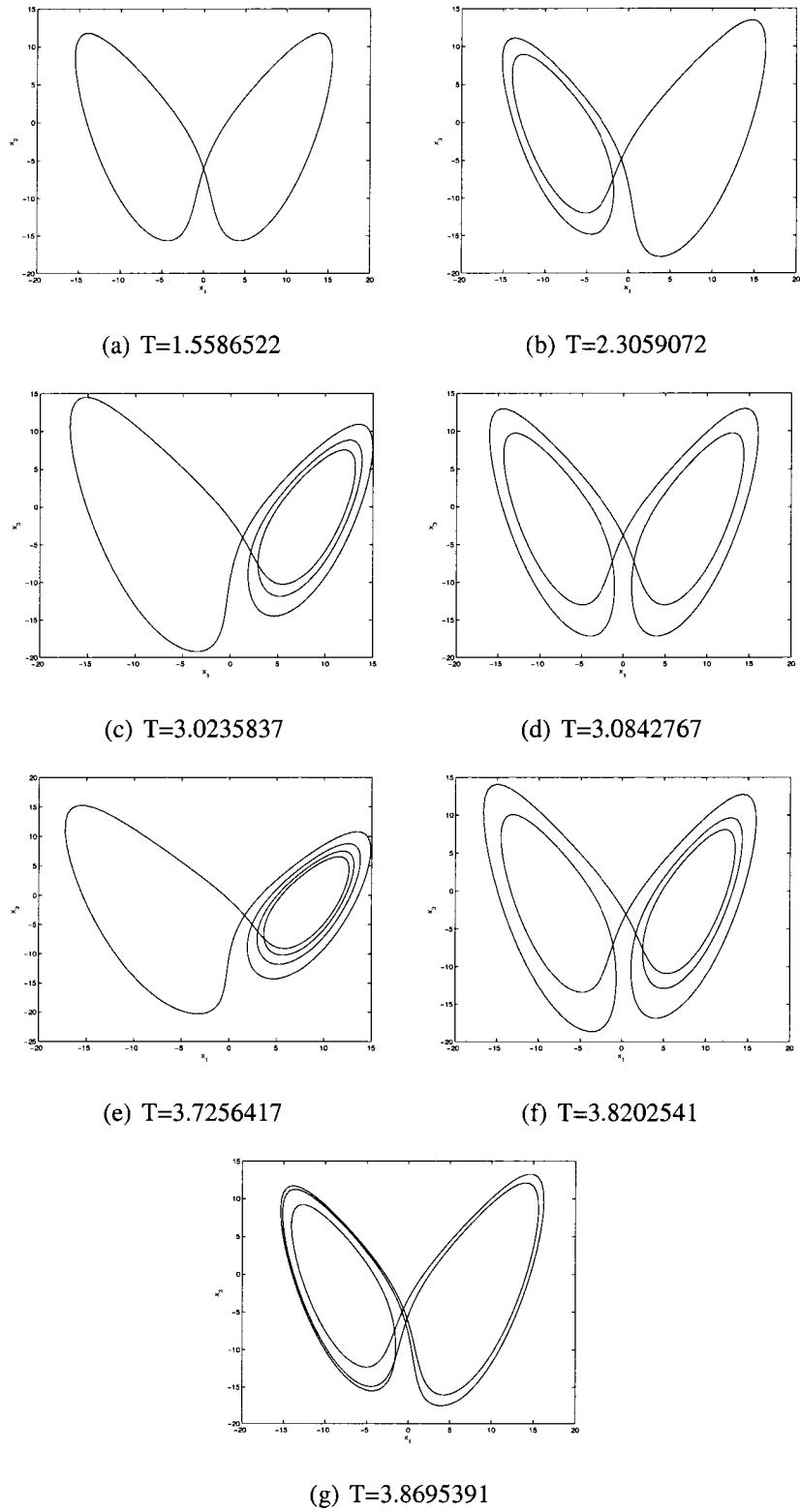


Figure 4.2: Periodic orbits with $T < 4$.

obtained in Viswanath (2003), up to the precision of the time step used in the simulation, and the tolerance imposed for convergence. Since these orbits are only approximate, sometimes two orbits with a similar period are obtained, which later in the refinement process converge to the same orbit. These approximately time-periodic orbits are used as initial conditions for the time-evolving iterates (final values of the period given in the third column of Table 4.1) and for the time-periodic iterates (final values of the period shown in the fourth column of Table 4.1).

It can be seen that the periods obtained by the time-evolving iterates converge to a final value that matches the values in Viswanath (2003). The convergence history of the cost function defined in Eq. (4.4) is shown in Fig. 4.3.

T (Viswanath)	T (approximate orbits)	T (time-evolving)	T (time-periodic)
1.5586522107152	1.560	1.5586522107202	1.5586538205782
2.3059072639398	2.305	2.3059072639673	2.3060675343917
3.0235837034339	3.020	3.0235837035343	3.0239132580456
3.0842767758221	3.085	3.0842767759418	3.0842797089088
3.7256417715558	3.725	3.7256417718239	3.7250350145471
3.8202541634368	3.820	3.8202541637764	3.8204379797438
3.8695391125646	3.905	3.8695391129544	3.8696469620383

Table 4.1: Period of the periodic orbits corresponding to $T < 4$.

A question that arises when performing the time-periodic refinement of the orbits is how many modes we need to keep. The more modes are used, the slowest is the method, due mainly to the Fast Fourier Transforms needed and the increase in the number of parameters to optimize. Fig. 4.4 shows the spectrum corresponding to the periodic orbit shown in Fig. 4.2(g) obtained by time-evolving refinement. It can be seen that approximately only the first 100 modes contribute to the spectrum, indicating that the time-evolution of the orbit, performed with a Runge-Kutta, is over-resolved (although over-resolving is necessary in the time-evolving framework for stability of the Runge-Kutta scheme). The rest of the orbits with $T < 4$ have also been analyzed and the same

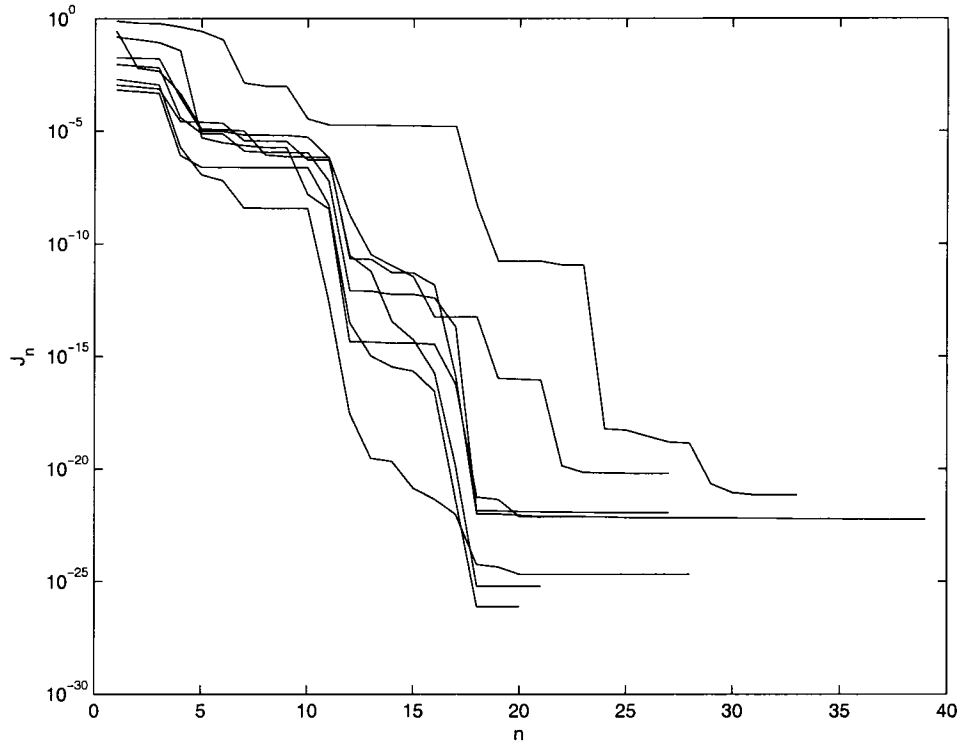


Figure 4.3: Convergence history of the cost function defined in Eq. (4.4) in the time-evolving setting as a function of the optimization iteration.

conclusion has been drawn. The time-periodic optimization can therefore be performed considering only the necessary modes. We have taken 192 modes for all the orbits.

The fourth column in Table 4.1 shows the time-periodic refinement of the periods obtained with the algorithm presented in §4.2. In this case the periods also match those obtained in Viswanath (2003), although not as accurately as the ones obtained by the time-evolving refinement, even though the criterion for convergence is the same (that is, stopping the optimization iteration when the relative change in the cost function from one iteration to the following one is less than a specified ϵ). However, in fluid systems, the approximation obtained with this method would be satisfactory, and not as much precision as that given by the time-evolving would be necessary. The history of the convergence for the different orbits as a function of the iteration number is shown in Fig. 4.5. We should remark that not necessarily the shortest orbits usually converge faster than the larger ones, since the number of modes is the same for all of them.

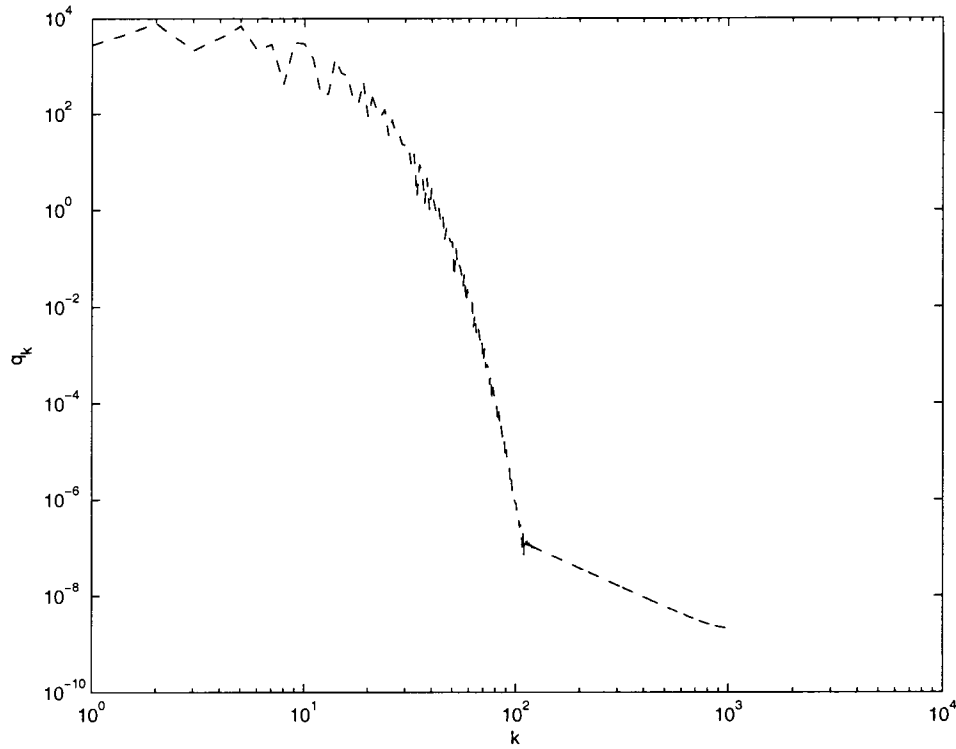


Figure 4.4: Spectrum of the periodic orbit shown in Fig. 4.2(g).

Finally, Fig. 4.1 shows the skeleton formed by the periodic orbits previously computed and a generic non-periodic orbit. As mentioned in the introduction, due to the ergodic nature of the attractor, the generic orbit passes close to all periodic orbits, as quantified in Fig. 4.6. In fact, periodic orbits are known to approach 100% as T_{max} is increased for any given value of ϵ , however small.

The skeleton formed by the periodic orbits reflects and gives information about the structure of the attractor.

4.6 Discussion

We have examined three different methods to obtain time-periodic orbits in a dynamic system. Although they have been tested on the Lorenz system, these methods are appropriate for higher-dimensional systems, as they are completely scalable. It is important to notice however, that in the time-periodic framework, the whole sequence of states

$\mathbf{q}(\tau)$ has to be stored, fact that might lead to the need of larger storage computers.

As mentioned in the previous sections, it is known that periodic orbits are dense in chaotic attractors (see, e.g., Ott (2002)); that is, for any ε , any given location on a chaotic attractor is within an ε neighborhood from a periodic orbit of period less than or equal to T_{max} , provided T_{max} is taken large enough (see Fig. 4.6). Examining the periodic orbits of period less than or equal to T_{max} in a chaotic system is an attractive technique of obtaining a “skeleton” of where the attractor lies. Once this skeleton is obtained, the entire attractor may be reconfigured by reconfiguring its time-periodic skeleton. This is basis of the control approach recently proposed by Bewley & Trenchea (2002), and will be examined numerically in a forthcoming paper.

4.7 Conclusions

The present paper presents a method to estimate periodic orbits in dynamic systems up to a period length of T_{max} , based on a time-evolving simulation of the system. Two methods to refine the estimated periodic orbits are discussed, one set in a time-evolving framework, where the periodicity of the orbit is optimized, and one set in a time-periodic framework, where the fidelity of the exactly periodic orbit to the governing equation is optimized. The three methods have been tested with the Lorenz chaotic system, where accurate results have been obtained.

The extension to higher-dimensional systems, as well as the necessity of finding periodic orbits in order to posteriorly perform an optimization have been discussed.

Appendix

Estimation of the descent parameter

4.A.1 Introduction

Gradient-based optimization methods perform a line minimization in the direction of the gradient (in the case of Steepest Descent) or the direction of the gradient with a momentum term (Conjugate Gradient):

$$\phi^{k+1} = \phi^k + \alpha^k h^k, \quad (4.32)$$

In both cases, an initial guess for the descending parameter α must be selected. In the present appendix, a new method to estimate an initial α that is closer to the optimum value than in the commonly used methods is explained. Results on the performance of the new estimation are given for the problem of parameter identification in a chaotic system.

4.A.2 Description of the problem

Consider the chaotic system described by the system of equations

$$\begin{aligned} \mathbf{n}(\mathbf{q}, \Omega) &= \mathbf{w}, \quad \mathbf{q}(t=0) = \mathbf{q}_0, \\ \mathbf{y}_M &= C\mathbf{q} + \mathbf{v}, \end{aligned} \quad (4.33)$$

where $\mathbf{y}_M$ represents the measurements obtained from the system. The problem under consideration is the identification of the parameters involved in this system by investigating whether it is possible to find $\phi = (\mathbf{q}_0, \theta, \mathbf{w}, \mathbf{v})$ such that the nominal model (that is, with $\mathbf{w} = 0, \mathbf{v} = 0$) is a very good approximation of the dynamics exhibited by the system (that is, that the state equation reproduces the measurements observed for small values of $\mathbf{v}$ and $\mathbf{w}$). Optimal values of $\mathbf{q}_0, \theta, \mathbf{w}, \mathbf{v}$ (our control variables) will be found via an iterative optimization procedure.

Lorenz attractor

The problem under study in this article is the Lorenz attractor, which is governed by the following system of equations

$$\begin{aligned} \dot{x}_1 &= \sigma(x_2 - x_1) + w_1 \\ \dot{x}_2 &= -x_2 - x_1x_3 + w_2 \\ \dot{x}_3 &= -b(x_3 + r) + x_1x_2 + w_3 \end{aligned} \quad (4.34)$$

where we define the state variable as $\mathbf{q} = (x_1, x_2, x_3)^T$, the vector as $\theta = (\sigma, b, r)^T$, and the system noise as $\mathbf{w} = (w_1, w_2, w_3)^T$. The vector y_M consists of a measurement of the second and third states, corresponding to temperature measurements on the system ($C = (0, 1, 1)$).

We have considered as the nominal state the one with $\theta = (\sigma, b, r)^T = (6, 2, 60)^T$ and $\mathbf{w} = \mathbf{0}$ for all times.

4.A.3 Derivation of gradients

An appropriate cost function for this problem, which penalizes the deviation of the model system from the nominal system is

$$J(\phi) = \frac{1}{2} \int_0^T (|\mathbf{v}|^2 + \rho |\mathbf{w}|^2) dt \quad . \quad (4.35)$$

We are free to adjust the value of ρ , which increases or decreases the relative weight of the penalization on $\mathbf{w}$ and $\mathbf{v}$. In order to proceed with the optimization, we need to obtain the gradient of this function with respect to the control variables. An adjoint-based gradient is obtained in this section.

First consider a small perturbation to the control variables

$$\begin{aligned} \mathbf{w} &\leftarrow \mathbf{w} + \mathbf{w}', & \mathbf{q}_0 &\leftarrow \mathbf{q}_0 + \mathbf{q}'_0, \\ \mathbf{v} &\leftarrow \mathbf{v} + \mathbf{v}', & \theta &\leftarrow \theta + \theta', \end{aligned} \quad (4.36)$$

with the accompanying small perturbation $\mathbf{q}'$ to the state variable $\mathbf{q}$. Substituting Eq. (4.36) and $\mathbf{q} \leftarrow \mathbf{q} + \mathbf{q}'$ in the state equation Eq. (4.33), neglecting terms that are

quadratic in the perturbations, and applying Eq. (4.33), we may write the perturbation equation for this problem as

$$\begin{aligned} L\mathbf{q}' &= B_\theta\theta' + \mathbf{w}', \quad \mathbf{q}'(t=0) = \mathbf{q}'_0, \\ 0 &= C\mathbf{q}' + \mathbf{v}', \end{aligned} \quad (4.37)$$

where, in the case of the Lorenz attractor given in Eq. (4.34), the matrix B_θ is given by

$$B_\theta = \begin{pmatrix} x_2 - x_1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & -(x_3 + r) & -b \end{pmatrix}. \quad (4.38)$$

The corresponding change in the cost function is

$$J' = \int_0^T (\mathbf{v}^*\mathbf{v}' + \rho\mathbf{w}^*\mathbf{w}') dt. \quad (4.39)$$

Note that $\mathbf{v}'$ is linearly related to $(\mathbf{q}_0, \theta, \mathbf{w})$ via the perturbation equation Eq. (4.37). We will re-express this relationship in a useful manner using an adjoint identity. For this purpose we first define the inner product $\langle \mathbf{r}, \mathbf{q}' \rangle = \int_0^T \mathbf{r}^* \mathbf{q}' dt$ and consider the adjoint identity

$$\langle \mathbf{r}, L\mathbf{q}' \rangle = \langle L^* \mathbf{r}, \mathbf{q}' \rangle + b, \quad (4.40)$$

from which we can derive the adjoint operator L^* and b (after some algebra, it is found that $b = \mathbf{r}^* \mathbf{q}'|_{t=T} - \mathbf{r}^* \mathbf{q}'|_{t=0}$).

The next step is to define the right-hand side of the adjoint equation. This definition is arbitrary, and it consists normally of a term which will later on yield an expression of the gradient which we can evaluate. In the present problem, an adequate form of this term is

$$L^* \mathbf{r} = -C^* \mathbf{v}, \quad \mathbf{r}(t=T) = 0. \quad (4.41)$$

Inserting the perturbation and adjoint equations into the adjoint identity yields

$$\langle \mathbf{r}, B_\theta\theta' + \mathbf{w}' \rangle = \langle -C^* \mathbf{v}, \mathbf{q}' \rangle + 0 - \mathbf{r}^*(0)\mathbf{q}'_0,$$

from where a proper relationship between $\mathbf{v}'$ and the perturbation in the parameters is obtained:

$$\int_0^T \mathbf{v}\mathbf{v}' dt = \left[\int_0^T B_\theta^* \mathbf{r} dt \right]^* \theta' + \int_0^T \mathbf{r}^* \mathbf{w}' dt + \mathbf{r}^*(0)\mathbf{q}'_0. \quad (4.42)$$

We may thus rewrite the expression for the perturbation cost function in Eq. (4.39) as

$$J' = \left[\int_0^T B_\theta^* \mathbf{r} dt \right]^* \theta' + \int_0^T (\mathbf{r} + \rho \mathbf{w})^* \mathbf{w}' dt + \mathbf{r}^*(0) \mathbf{q}_0'. \quad (4.43)$$

Defining the gradient such that

$$J' \triangleq \left[\frac{DJ}{D\theta} \right]^* \theta' + \left[\frac{DJ}{D\mathbf{w}} \right]^* \mathbf{w}' + \left[\frac{DJ}{D\mathbf{q}_0} \right]^* \mathbf{q}_0',$$

we finally obtain the desired gradients

$$\frac{DJ}{D\theta} = \int_0^T B_\theta^* \mathbf{r} dt, \quad \frac{DJ}{D\mathbf{w}} = \mathbf{r} + \rho \mathbf{w}, \quad \frac{DJ}{D\mathbf{q}_0} = \mathbf{r}(0). \quad (4.44)$$

Once the gradients are computed via the adjoint field, any gradient-based optimization procedure may be used. In the present paper the conjugate-gradient method will be used.

4.A.4 Estimation of the descent parameter

We now consider the problem of updating the control ϕ in some direction ϕ' (gradient direction in the *Steepest Descent* method, or gradient direction modified by a momentum term in the *Conjugate Gradient* method), scaling this update by a factor α which we are free to choose. We would like to be able to select the value of the scalar α to minimize the resulting cost function, $J(\phi + \alpha\phi')$; in other words, we would like to perform a *line minimization* with respect to the parameter α . Most of the computer time is spent in this line minimization, which is an iterative procedure (inside the iterative optimization procedure). For this reason it is important to have a method to estimate α such that it is as close to the optimal value as possible. A new method for estimating α is proposed below.

The minimization may be obtained by considering the Taylor series expansion for $J(\phi + \alpha\phi')$ near $J(\phi)$ and keeping terms up to those which are quadratic in the perturbations, then setting $dJ(\phi + \alpha\phi')/d\alpha = 0$ in the resulting expression and solving for α , as shown below.

The Taylor series expansion for $J(\phi + \alpha\phi')$, up to second order terms is

$$J(\phi + \alpha\phi') = J(\phi) + \alpha J'(\phi; \phi') + \frac{\alpha^2}{2} J''(\phi; \phi') + O(\alpha^3). \quad (4.45)$$

It is straightforward now to obtain the value of α that would minimize the cost function in case of having a quadratic function (which becomes a good approximation as the control parameters get close to the values that minimize the cost function):

$$\alpha_q = -\frac{J'(\phi; \phi')}{J''(\phi; \phi')}. \quad (4.46)$$

The second perturbation cost function J'' can be obtained from the perturbation J' given in Eq. (4.39)

$$J'' = \int_0^T \left[\mathbf{v}'^* \mathbf{v}' + \mathbf{v}^* \mathbf{v}'' + \rho \mathbf{w}'^* \mathbf{w}' \right] dt \quad (4.47)$$

The unknowns in Eq. (4.39) and Eq. (4.47) are $\mathbf{v}'$ and $\mathbf{v}''$, which can be obtained from the measurements equation in Eq. (4.33)

$$\begin{aligned} \mathbf{v}' &= -C\mathbf{q}' \\ \mathbf{v}'' &= -C\mathbf{q}'' \end{aligned} \quad (4.48)$$

The perturbation variables $\mathbf{q}'$ and $\mathbf{q}''$ have to be obtained from an evolution equation. $\mathbf{q}'$ is obtained from the perturbation equation Eq. (4.37). An evolution equation for $\mathbf{q}''$ may be obtained by perturbing Eq. (4.37):

$$L\mathbf{q}'' + L'\mathbf{q}' = B'_0\theta' \quad , \quad \mathbf{q}''(t=0) = 0. \quad (4.49)$$

Note that the perturbation equation Eq. (4.37) is linear in the perturbation variables, whereas the second perturbation equation Eq. (4.49) is linear in the second perturbation variables, but not in the first ones.

This method of estimating the descent parameter α has been implemented in the optimization code, and compared to other methods in order to show its effectiveness.

The most straightforward method to use for the estimation of α would be to pick a constant value α_c at the beginning of all the optimization iterations. Two different values have been picked ($\alpha_c = 0.1$ and $\alpha_c = 0.01$), in order to show the difference in function evaluations depending on the value picked. For $\alpha_c = 0.1$, shown in Fig. 4.7, 344 optimization iterations were needed in order to make the cost function to be below 0.02, involving 3014 function evaluations (evolution of both state and adjoint systems). For

$\alpha_c = 0.01$, in Fig. 4.8, 341 optimization iterations were necessary, but involving only 2283 function evaluations. This difference shows that the estimation of α is very important in order to reduce optimization time. In the case of $\alpha_c = 0.01$, the estimation is closer to the actual optimal value for most of the iterations than for $\alpha_c = 0.1$, being the number of function evaluations much smaller (although the number of iterations in the optimization loop is very close).

Another way to estimate α is to use the value that gave the optimum in the previous iteration. This way of picking α will preserve the order of magnitude of the parameter α , which, as seen Fig. 4.7 does not change much. This preservation of the order of magnitude cannot be done when picking a constant value of the estimation, unless it is known *a priori*. With this method, in the present problem, 326 optimization iterations are needed in order to lower the cost function to the desired level, involving 2292 function evaluations, very similar to the case where $\alpha_c = 0.01$ was the initialization parameter in every iteration (because this constant value gives more or less the appropriate order of magnitude).

At last, the new method described above has been tested. The result is shown in Fig. 4.9. It can be seen that during the first iterations the estimation of α does not coincide with the optimal value, approaching this two values as the number of iteration increases (after a few iterations). The number of iterations needed with this method is 371, invoking 3521 iterations (this high number is mostly due to the fact that in every iteration both the perturbation and second perturbation equations are evolved, and then Brent's method performs more evaluations than probably needed).

The fact that the value of the estimated α as described above and the optimal value are close after a few iterations shows that the parabolic approximation becomes more valid. Defining a 'trust' variable which measures how good the parabolic approximation is (comparing the value obtained with the one given by inverse parabolic interpolation), it is possible to avoid further function evaluations, when this variable acquires a given value, by using the value of α estimated with Eq. (4.46). With this new variable, 264 iterations were needed, and 1073 function evaluations (counting the evolution of the

adjoint equation, the state equation, the perturbation equation and the second perturbation equation). The newly defined variable is restarted whenever the new cost function is larger than the old one.

In some cases, even when the new cost function is not larger than the old one, it is better to restart this variable that measures the proximity to the parabolic approximation every few optimization iterations, since once the trust variable acquires the desired value, no more comparisons with the parabolic interpolation result is performed. We have restarted the variable every 20 iterations. The new optimization takes 242 iterations, and involves 1046 function evaluations, a number much smaller than the first method of α_c .

In the first methods, the number of iterations in the optimization loop was similar. However in the last two cases, with the new method and assuming that the parabolic approximation is correct whenever a trust variable used has an appropriate value, many less iterations are needed, and the number of function evaluations is greatly reduced.

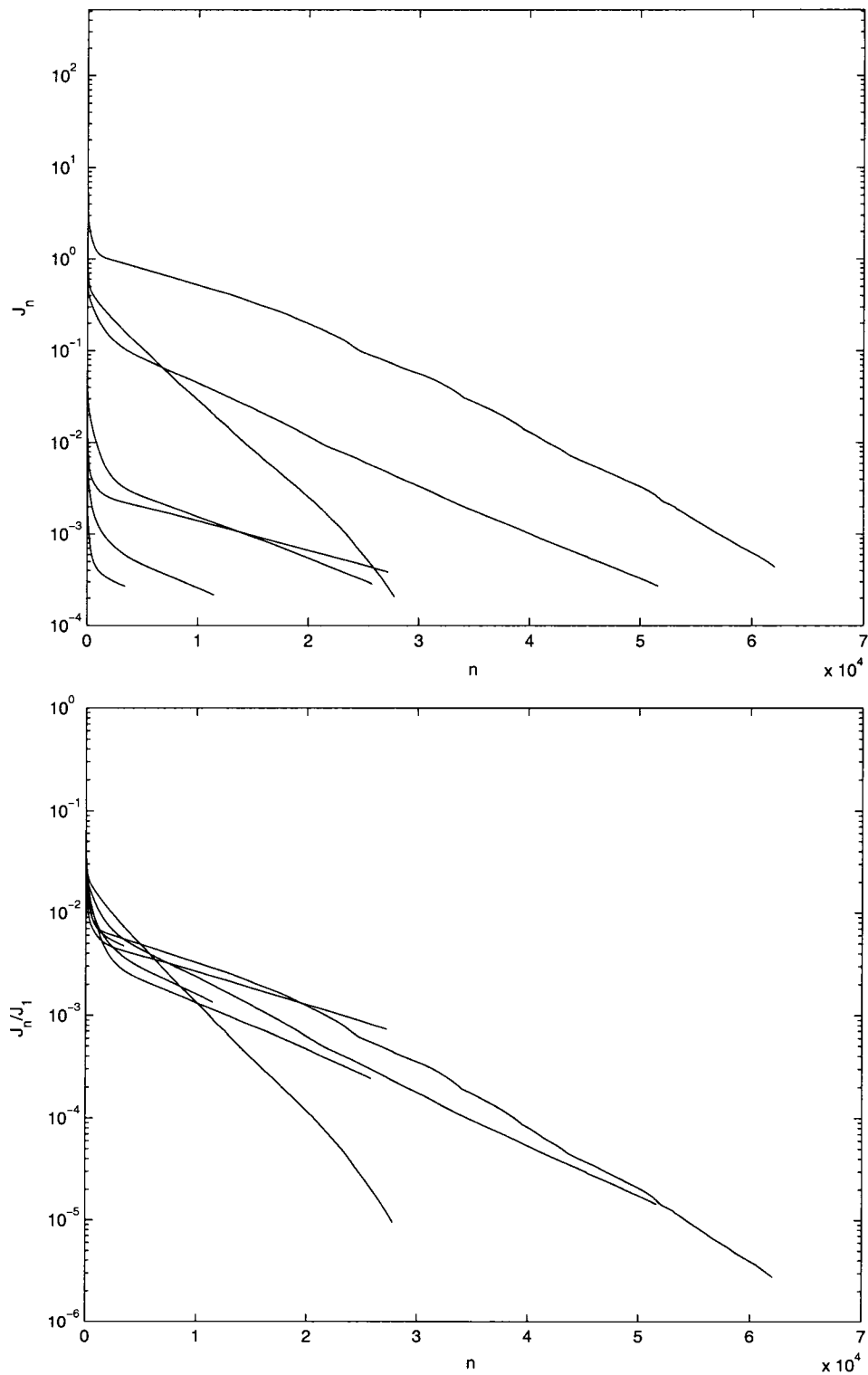


Figure 4.5: Convergence history of the cost function defined in the time-periodic setting as a function of the optimization iteration (top) and same cost function nondimensionalized with the initial value (bottom).

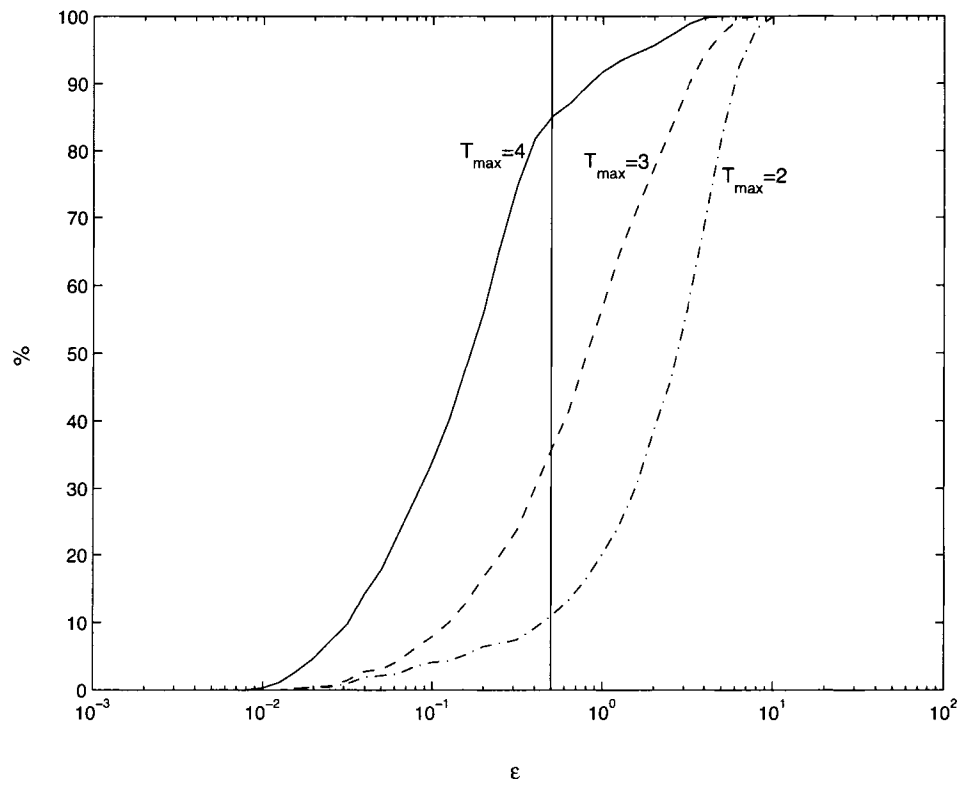


Figure 4.6: Proximity of the attractor to a set of periodic orbits.

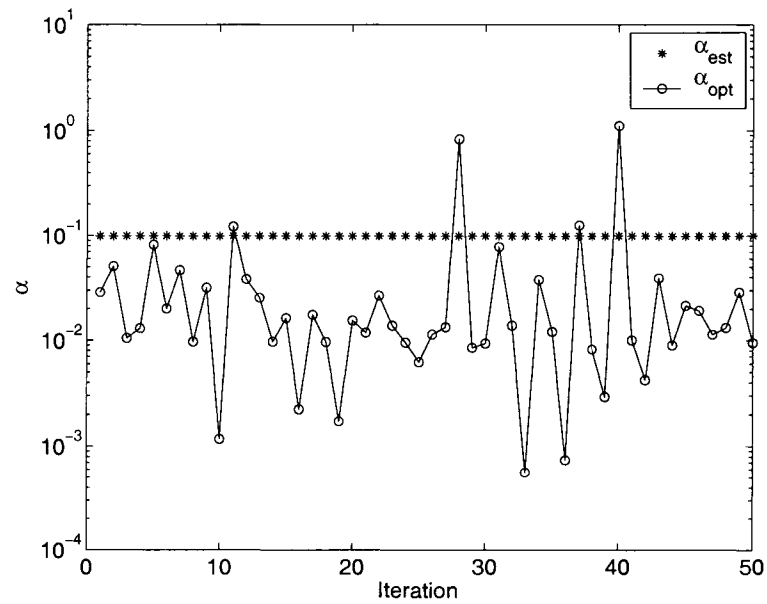


Figure 4.7: Estimated $\alpha_{\text{est}} = 0.1$ and optimal α_{opt} .

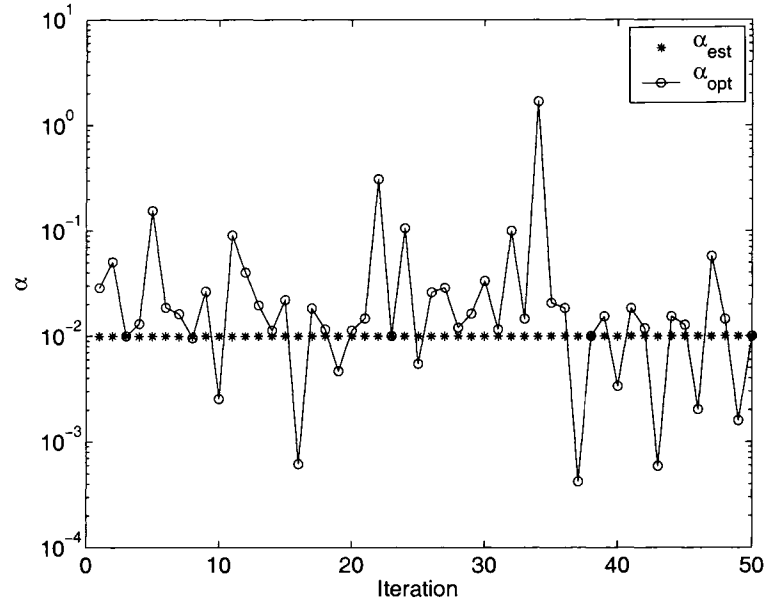


Figure 4.8: Estimated $\alpha_{est} = 0.01$ and optimal α_{opt}

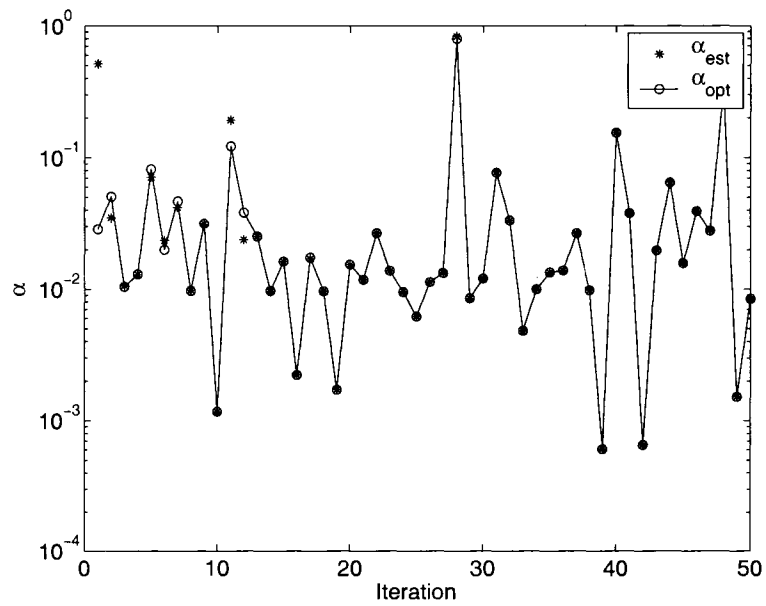


Figure 4.9: α estimated with the method described in Eq. (4.46).

References

- BEWLEY, T. Preprint Numerical methods in science and engineering. Available at http://turbulence.ucsd.edu/numerical_methods/notes.pdf.
- BEWLEY, T., MOIN, P. & TEMAM, R. 2001 DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms. *J. Fluid Mech.* **447**, 179–225.
- BEWLEY, T. & TRENCH, C. 2002 Noncooperative optimizations of controls for time-periodic navier-stokes systems with multiple solutions. In *AIAA Paper 02-2754*.
- KAWAHARA, G. & KIDA, S. 2001 Periodic motion embedded in plane couette turbulence: regeneration cycle and burst. *J. Fluid Mech.* **449**, 291–300.
- KAZANTSEV, E. 1998 Unstable periodic orbits and attractor of the lorenz model. INRIA Research Rapport RR-3344, Lorraine.
- LAN, Y. & CVITANOVIĆ, P. 2004 Variational method for finding periodic orbits in a general flow. *Phys. Rev. E* **69** (016217).
- LORENZ, E. N. 1963 Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141.
- LUO, H. 2004 The interaction of near-wall turbulence with compliant tensegrity fabrics: Modeling, simulation, and optimization. PhD thesis, UC San Diego.
- OTT, E. 2002 *Chaos in Dynamic Systems.*, 2nd edn. Cambridge University Press.
- PRESS, W., FLANNERY, B., TEUKOLSKY, S. & VETTERLING, W. 1992 *Numerical Recipes.* Cambridge University Press.
- PROTAS, B., BEWLEY, T. & HAGEN, G. 2004 A comprehensive framework for the regularization of adjoint analysis in multiscale pde systems. *J. Comput. Phys.* **195** (1), 49–89.
- VISWANATH, D. 2003 Symbolic dynamics and periodic orbits of the lorenz attractor. *Nonlinearity* **16**, 1035–1056.
- WALEFFE, F. 2001 Exact coherent structures in channel flow. *J. Fluid Mech.* **435**, 93–102.