

**COMPUTATIONAL METHODS IN CHEMICAL KINETICS
REDUCTION AND ANALYSIS FOR COMBUSTION MODELING
AND APPLICATIONS TO HYDROCARBON SYSTEMS**

Tesi di Dottorato di Ricerca in
Ingegneria Aerospaziale (XVII ciclo)

Scuola di Ingegneria Aerospaziale
Università di Roma “La Sapienza”
Via Eudossiana 16, 00184 Rome, Italy

Dottorando: Ing. Francesco Creta

Francesco Creta

May 2005

Preface

The recent years have called upon aerospace researchers to maximize their efforts towards the primary goal of developing solutions aimed at reducing the costs of access to orbit. The widely established short to medium-term option was found in the pursuit of a greater degree of reusability of either the entire launcher or part of it, so that the high development costs can be partly recovered. This situation has therefore prompted an increasing interest in exploring scientific and technological alternatives to the state of the art in space launchers, constituted by expendable (ELV) vehicles.

Among the most critical components of a launcher is its liquid rocket engine (LRE) and it is therefore towards the LRE that efforts should be focused in view of developing a reusable vehicle. Indeed, LRE's involve highly optimized machinery with technological challenges pushed towards the extreme performance limits, with operating pressures up to 200 bar and temperatures up to 3500 K. The large temperature gradients in LRE's stress materials both mechanically and thermally, resulting in thermal and mechanical coupled fatigue at best, cracks and hot corrosion at worst. This combination is unique to engineering, and intolerant of small flaws or defects, practical or conceptual. ELV's still have a 2% loss rate at best, this being a very high price to pay, dramatically distant from commercial airlines' standards, essentially due to the still low reliability of LRE. Analysis of the scientific aspects associated to state of the art liquid rocket engines indicates that LH₂/LOX engines are extremely expensive to develop and maintain operational. Only one reusable LH₂ engine exists (the SSME of the US Shuttle), which is, however, routinely refurbished completely after 30-35 missions at most, while commercial aeroengines need complete overhauling only after some 20,000 h. Kerosene LRE are much cheaper to develop, but have never been reused: their high wall heating due to soot formation during the combustion process makes them hard to design, if reusability is the primary issue.

Based on these premises, investigation is under way to probe the potential of storable liquid hydrocarbons LHC/LOX engines, with emphasis on methane (liquified natural gas, LNG). Propane (liquified petroleum gas, LPG) is also of interest. Costwise, LNG is a factor 4 cheaper than kerosene and a factor 7 cheaper than LPG. Moreover LHC density is far greater than LH₂ this implying lighter and smaller tanks. But the real issues connected with these choices are the physics and chemistry of combustion of these new fuels, for which experience is scarce. Conditions as close to a premixed prevaporized combustion regime in the combustion chamber are desireable for faster and more uniform mixing, resulting in shorter and lighter chambers, all these aspects concurring to a greater degree of reusability and lower costs. Injection of both fuel and oxidizer at supercritical conditions, for example, would prevent liquid jets or droplets to form, eliminating atomization and vaporization thereby making the combustion faster and the chamber shorter.

The issues just reviewed, only scratch the surface of a massive bulk of physico-chemical and technological problems that need to be investigated in great depth. Ever more detailed CFD (Computational Fluid Dynamics) modeling of the combustion process in the thrust chamber is therefore required to account for the coupling of an intrinsically turbulent flow field with chemical kinetics. State of the art computing power rules out a direct numerical simulation (DNS) of the entire spectrum of spatio-temporal scales involved in the turbulent combustion phenomena taking place in the thrust chamber. Even when turbulence modeling is used, such as LES (Large Eddy Simulation) or RANS (Reynolds-averaging with closure equations), it is still an exceedingly heavy computational task to resolve the widely spread chemical timescales (10^{-9} s - 10^1 s) emerging from a detailed chemical kinetics mechanisms for hydrocarbon combustion, typically comprising 100 to 1000 species. New theoretical and numerical tools are therefore mandatory if a realistic simulation of the phenomena occurring in a combustion chamber is to be attempted. In particular, a sound understanding of the following processes is needed: combustion instabilities (involving coupling between chamber acoustics and injection, atomization, chemical kinetics, mixing), heat transfer in thrust chambers, non-uniformity of mixture ratio distributions, micromixing.

State of the art turbulent combustion codes, forcefully utilise turbulence modeling together with severely reduced or simplified chemical kinetics mechanisms. This is a consequence of the extremely wide spatio-temporal scale spectrum on which the complex of turbulent combustion phenomena occur. The spatial resolution required, if no turbulence model is adopted, can be far from

tractable with the present-day computing power, especially at the high Reynolds numbers typical of combustion chambers. Moreover, the widely spread chemical timescales, which concur to the 'stiffness' of the governing equations, introduce severe stability requirements on the numerical methods employed. In particular, explicit schemes have stringent stability limits, which force them to advance the numerical integration on the smallest time-scale. Implicit schemes have less severe stability limits, their integration time step mostly controlled by accuracy requirements, but are less efficient computationally and more difficult to parallelize.

In this context, the present dissertation will analyse recent, and develop new methodologies for the dimensionality reduction and stiffness elimination of complex chemical systems.

Many spatially homogeneous kinetic systems, whose dynamics can be modeled by sets of ordinary differential equations, approach equilibrium slowly after the initial fast transients have died out, along surfaces of the phase space whose dimension is far smaller than the system's original dimension. These surfaces, referred to as "slow invariant manifolds", embed the asymptotic slow-evolving behavior of the system. Constraining the system onto such surfaces implies a reduction in the dimensionality of the system and, generally, a reduction in its stiffness. A proper definition of slow manifolds and their accurate identification is therefore crucial for the task of system reduction, whose ultimate goal is to provide smaller dimensional versions of the system which retain, as faithfully as possible, its asymptotic dynamical features.

The classical chemical kinetics reduction approaches, such as the Quasy-Steady-State Approximation (QSSA) or the Partial Equilibrium Approximation (PEA), have generally relied on experience and intuition to identify the quasy-steady state species and the fast reactions. As a result, the identification of slow manifolds is often an inaccurate and ambiguous process. More recently, other methods have employed a more rigorous geometric approach to generate the asymptotic dynamics of chemical kinetics systems and which generally yield greater accuracy in the identification of the low-dimensional manifold. Among them, the Computational Singular Perturbation (CSP) method developed by Lam and Goussis, the Intrinsic Low Dimensional Manifold (ILDM) method developed by Maas and Pope and the Fraser and Roussel method. The former two methods are essentially based on the identification of a suitable projection vector basis of the phase space, which is related to the system's dynamics, and which, relying on a sufficient separation between fast and slow time scales, defines a decomposition of such a space into a fast and a slow subspace. The slow

subspace will ultimately approximate the slow manifold. The Fraser and Roussel method, on the other hand, relies on invariance properties of the manifold with respect to the system's dynamics, and aims at generating such manifold exactly.

The present work is focused on the CSP method, its theoretical foundations, applications and developments. Extensive comparison with the ILDM and Fraser-Roussel methods is also given. A new automatic CSP-based mechanism simplification algorithm is then developed and applied to combustion processes including autoignition and premixed laminar flames in methane/air mixtures.

The visualization of a chemical kinetics system as a dynamical system, and the use of dynamical systems theory applied to reactive problems, allows for a broader view of the problems involved in system reduction. It is well established that non linear systems can exhibit asymptotic behaviors that are more complex than isolated equilibrium points. Indeed phenomena ranging from combustion kinetics to enzymatic reactions, often exhibit persistent oscillatory or even chaotic behaviors. In these cases, the conceptualization of asymptotic behavior with a slow motion constrained on the invariant manifold, leading to the equilibrium point, does not generally apply. Consequently, the application of reduction techniques whose prerequisite is the existence of slow manifolds, such as CSP or ILDM, is often destined to fail. In such a context, this work develops a new technique termed Natural Tangent Dynamics with Recurrent Biorthonormalizations (NTDRB), aimed at extending the concepts of system reduction to systems exhibiting non trivial asymptotic oscillatory or chaotic behaviors. Thus, NTDRB unifies into a single technique the analysis of dynamical systems possessing either a slow manifold associated with an equilibrium point, for which the main focus relies in the identification of the slow/fast timescales, or systems exhibiting non trivial behaviors, for which the central issue is the temporal dichotomy between unstable/stable components of the dynamics. The technique aims at finding the unique invariant phase space decomposition for the system and its associated intrinsic timescale spectrum by following the dynamics of vector elements, conceptually equivalent to perturbations of single trajectories, as it is naturally induced by the system. Because vectors have a natural tendency to progressively become aligned with the system's most unstable direction, an "unfolding" biorthonormalization procedure is recurrently enforced so as to induce each vector to align itself along each of the invariant system's directions.

The present work contains some introductory and illustrative material, based on existing literature, as well as new and original concepts. In particular the general implementation of an iterative CSP

basis vector "refinement" procedure, is presented for the first time, improving the decoupling between the fast and slow components of the dynamics, thus allowing for a more accurate estimation of the slow manifold. The CSP technique, typically applied to (finite dimensional) spatially homogeneous systems, is also extended to (infinite dimensional) non homogeneous convective-diffusive-reactive systems. The approximations and limits of applicability of such an extension have also been investigated and discussed. A new automatic chemical kinetics mechanism simplification algorithm, based on CSP data, is then developed and validated with respect to ignition phenomena and premixed flames. Finally, the NTDRB method is introduced, which, gathering established and original concepts in reduction methods and dynamical systems theory, represents, in its complex, a new technique.

The outline of the present dissertation is as follows. **Chapter 1** introduces the mathematical foundation of the CSP method comparing it to the ILDM and Fraser-Roussel methods by means of well known, low dimensional, analytical model problems and examples. It then moves on illustrating the CSP basis vectors refinement procedure which is based on a block power-method (introduced in Appendix A). A full description of the general implementation of the CSP refinements is also given. Finally a 3-species kinetic model is used as an illustrative example. **Chapter 2** introduces the CSP-based simplification algorithm. CSP "importance indices" and "pointers to CSP radicals" are illustrated, the complex of which constitutes a CSP database. The algorithm is then explained and subsequently applied to the autoignition of methane in air. Several simplified mechanisms are generated from the detailed GRI mech.3.0 and an extensive accuracy assessment is carried out. The diagnostic features of the CSP indices are also introduced and used to examine NO formation pathways. The CSP method is then extended to premixed laminar flames and comprehensive mechanisms extracted and validated. **Chapter 3** introduces the NTDRB technique. An introduction to the necessary dynamical systems theory and linear algebra tools is given. The concept of invariance with respect to a dynamical system is illustrated, and the definition of invariant subspace introduced by means of Oseledec's Multiplicative Ergodic Theorem. The founding "dominance principle" of NTDRB is then explained and finally the practical implementation of NTDRB is described. Several model examples are given illustrating the capability of NTDRB to capture the stable/unstable invariant subspaces for systems exhibiting limit cycles or chaotic attractors. Finally a practical application of NTDRB is given for a simplified H₂/O₂ system, comparing it with corresponding CSP results.

Related Publications

International Journals.

- [1] D.A. Goussis, M. Valorani, F. Creta, H.N. Najm , **Reactive and Reactive/Diffusive Time Scales in Stiff Reaction-Diffusion Systems**, *Progress In Computational Fluid Dynamics*, An International Journal (PCFD), Apr. 2004, in print.
- [2] A. Adrover, F. Creta, M. Giona, M. Valorani, V. Vitacolonna, **Natural Tangent Dynamics with Recurrent Biorthogonalizations: a geometric computational approach to dynamical systems exhibiting slow manifolds and periodic/chaotic limit sets**, *Physica D*, Feb. 2004, submitted.
- [3] A. Adrover, F. Creta, M. Giona, M. Valorani, **Biorthogonalization, invariant filtrations and Lyapunov spectra of chaotic dynamical systems**, *Physics Letters A*, Jan. 2004, in print.
- [4] M. Valorani, D. A. Goussis, F. Creta, H.N. Najm, **Higher Order Corrections in the Approximation of Low Dimensional Manifolds and the Construction of Simplified Problems with the CSP Method**, *J. Comput Phys*, Oct. 2004, in print.
- [5] M. Valorani, D. A. Goussis, F. Creta, H.N. Najm, **An Automatic Procedure for the Simplification of Chemical Kinetics Mechanisms based on CSP.**, *Combustion and Flame*, May. 2005, submitted.
- [6] M. Giona, A. Adrover, M. Valorani, F. Creta, S. Cerbelli, **Slow Manifold Structure and Bifurcation in Prototypical Models of Thermal Explosions.**, *Combustion Theory and Modelling*, Apr. 2005, submitted.

International Conferences and Meetings.

- [7] M. Valorani, F. Creta, D.A. Goussis, H.N. Najm, J. C. Lee, **Chemical Kinetics Mechanism Simplification via CSP**, *MIT Conference in Computational Fluid and Solid Mechanics*, 2005, submitted.
- [8] M. Valorani, F. Creta, **Systematic Investigation of CSP Generated Simplified Kinetics Mechanisms for GRI3.0 Methane/Air Mixtures**, *Joint Meeting of the Italian and Greek Sections of The Combustion Institute*, Paper No.019, Jun. 2004, Corfu, Greece.
- [9] F. Creta, M. Valorani, **Automatic Generation of Skeletal Kinetics for HC Fuels in LRE.**, *European Conference for Aero-Space Sciences (EUCASS)*, Jul. 2005, Moscow, Russia.
- [10] D.A. Goussis, M. Valorani, F. Creta, H.N. Najm, **Inertial Manifolds with CSP**, *Second MIT Conference in Computational Fluid and Solid Mechanics*, Jun. 17-20 2003, vol. 2, pp. 1951-1954.
- [11] M. Valorani, F. Creta, D.A. Goussis, H.N. Najm, **Local and Global Manifolds in Stiff Reaction-Diffusion Systems**, *Second MIT Conference in Computational Fluid and Solid Mechanics*, Jun. 17-20 2003, vol. 2, pp. 1548-1551.

Other International/National Conference and Meeting Participations.

H. N. Najm., M. Valorani, D. A. Goussis, F. Creta, **Analysis of Multidimensional Reacting Flow with CSP**, *SIAM Conference on Applications of Dynamical Systems*, May 2003, Snowbird, Utah, USA.

D.A. Goussis, H. Lam., M. Valorani, F. Creta, **Fundamentals of the Computational Singular Perturbation Method**, *SIAM Conference on Applications of Dynamical Systems*, May 2003, Snowbird, Utah, USA.

M. Valorani, F. Creta, D. Goussis, H. Najm, **Approximating Inertial Manifolds and Constructing Simplified Problems with the CSP Method**, *55th Annual Meeting of the Division of Fluid Dynamics, The American Physical Society*, Nov. 2002, Dallas, Texas, USA.

Acknowledgments

My gratitude goes out to Prof. M.Valorani, Prof. M.Giona, Prof. A.Adrover and Dr. S. Cerbelli of the University of Rome "La Sapienza", Dr. D.A.Goussis of ICEHT-FORTH, Rio Patras Greece, and Dr. H.N.Najm of the Combustion Research Facility, Sandia National Laboratories, Livermore CA, for their precious guidance and friendship. This work has profited from the suggestions, criticisms and review of Prof. S.H.Lam, of Princeton University, to whom I express my gratitude. This work was partly supported by the Italian Space Agency (ASI) and the Italian Ministry of Education, University and Research (MIUR).

Contents

Preface	iii
Related Publications	viii
Acknowledgments	xi
List of Tables	xv
List of Figures	xvii
1 Computational Singular Perturbation	1
1.1 Introduction	1
1.2 A definition of timescale	5
1.3 Slow invariant manifolds for some model problems	7
1.4 Steady-state and equilibrium assumptions	10
1.5 The Fraser Roussel method	11
1.6 The ILDM method	14
1.7 An Ideal choice for basis vectors	18
1.8 The CSP method.	22
1.8.1 The CSP refinement procedure	25
1.8.2 CSP refinement: general implementation	29
1.8.3 Basis Vectors Time Derivatives	33
1.8.4 CSP application examples	34
2 Chemical Kinetics Simplification by CSP	48
2.1 Introduction	48
2.2 Theoretical Background	48
2.3 Pointers to CSP Radicals	53
2.4 Importance Indices	55
2.4.1 Importance indices vs. conventional sensitivity coefficients	56

2.5	Simplification of chemical kinetics	57
2.6	Results	60
2.6.1	Comprehensive mechanism	69
2.6.2	Accuracy assessment	71
2.6.3	Extraction of an optimal mechanism	76
2.7	Some intermediate conclusions	80
2.8	Extension of the CSP method to reactive flows.	83
2.9	Applications to premixed laminar flames	85
2.9.1	Model simplification for premixed laminar flames.	89
2.9.2	Results	90
2.9.3	CSP data analysis.	92
2.10	Applications to counterflow diffusion flames	95
3	Natural Tangent Dynamics with Recurrent Biorthonormalizations	103
3.1	Introduction	103
3.2	Linear dynamics	108
3.3	Nonlinear dynamics	115
3.4	Vector dynamics and invariance properties	115
3.5	A Lagrangian approach	123
3.6	Oseledec's Multiplicative Ergodic Theorem	127
3.7	A dominance principle	129
3.8	The NTDRB technique to estimate invariant filtrations and Lyapunov exponents	135
3.9	A CSP-based approach using NTDRB	141
3.10	Non biorthogonality of the invariant system	144
3.11	Model reduction	153
3.12	Applications of NTDRB to reactive systems.	155
Appendix A		
	The block-power method	160
Appendix B		
	CSP Refinements: Evaluation of derivatives	167
Bibliography		173

List of Tables

1.1 Successive iterations of the Fraser-Roussel method for the Michaelis-Menten system, starting from the initial function $c_0(s) = s$. Shown are the first four coefficients $c_n^{(i)}$ in the expression $c_n(s) = c_n^{(0)} + \varepsilon c_n^{(1)} + \varepsilon^2 c_n^{(2)} + \varepsilon^3 c_n^{(3)} + \dots$	12
1.2 Successive iterations of the Fraser-Roussel method for the Davis-Skodje system, starting from the initial function $y_0(z) = z$. Shown are the first four coefficients $y_n^{(i)}$ in the expression $y_n(z) = y_n^{(0)} + \varepsilon y_n^{(1)} + \varepsilon^2 y_n^{(2)} + \varepsilon^3 y_n^{(3)} + \dots$	13
1.3 Effect of A_r -refinements starting from basis (1.103)-(1.104).	37
1.4 Effect of B^r -refinements starting from basis (1.103)-(1.104).	39
1.5 Effect of B^r -refinements without time derivatives, starting from basis (1.125)-(1.126).	39
1.6 CSP manifolds obtained from successive B^r -refinements without time derivatives, starting from basis (1.125)-(1.126). The CSP manifold $\mathcal{K}(k, m)$ is expressed in terms of its asymptotic expansion $y = \psi^{(k, m)}(z) = \psi_0^{(k, m)}(z) + \varepsilon \psi_1^{(k, m)}(z) + \varepsilon^2 \psi_2^{(k, m)}(z) + \dots$	40
1.7 Effect of B^r -refinements including time derivatives, starting from basis (1.125)-(1.126).	40
1.8 CSP manifolds obtained from successive B^r -refinements including time derivatives, starting from basis (1.125)-(1.126). The CSP manifold $\mathcal{K}(k, m)$ is expressed in terms of its asymptotic expansion $y = \psi^{(k, m)}(z) = \psi_0^{(k, m)}(z) + \varepsilon \psi_1^{(k, m)}(z) + \varepsilon^2 \psi_2^{(k, m)}(z) + \dots$	41
1.9 List of refinement strategies (methods) and of the associated algorithmic options.	45
2.1 Structure of the simplification algorithm	60

2.2	Example of simplified mechanisms $R(tol, [CH_4, NO])$ obtained for stoichiometric (left) and lean (right) combustion at atmospheric pressure. Reactions are listed in order of decreasing peak importance index as found by the iterative algorithm. Such peak indices are shown on the left of the relative reaction and, in brackets, the species they refer to together with the specification of whether it is a (s)low or (f)ast index. Tolerance levels were calibrated to yield the same number of active reactions in both cases. Nitrogen chemistry was regrouped at the bottom for clarity. The stoichiometric case highlights the importance of thermal NO and partially of prompt NO, while for the lean case the N ₂ O-intermediate mechanism is predominant.	66
2.3	Summary of the main features of the asymptotic behaviour of methane combustion. Number in brackets in the slow species column is the value of the CSP pointer. Numbers in the last column are the values of the participation indices to the driving slow mode of the specified reaction	68
2.4	Comprehensive simplified mechanism obtained by cross-referencing the CSP databases of the stoichiometric and lean cases.	72
2.5	Left: $S_{rem}(tol, [NO, CH_4])$ set containing species progressively removed from simplified mechanisms as tol is increased. Right: $S_{rem}(tol, [CO_2, CH_4])$	76
2.6	Performance summary for optimal mechanisms in the overall average error ranges [1%,10%] and [10%,100%].	78
2.7	Modified simplification algorithm	90
2.8	Summary of comprehensive mechanisms used for validation. Also shown, set S_{rem} containing species progressively removed from the simplified mechanisms as tol is increased.	93
2.9	Summary of comprehensive mechanisms used for the counterflow diffusion flame validation. Also shown, set S_{rem} of species progressively removed from the simplified mechanisms as tol is increased.	96
3.1	Liapunov exponents Λ_h calculated with the exterior algebra technique for the FT system.	154
3.2	Chemkin format H ₂ combustion simplified mechanism. For each reaction, A is the pre-exponential factor, β the temperature exponent and E the activation energy in the Arrhenius forward rate constant $k = AT^\beta e^{-E/(RT)}$	155

List of Figures

1.1	Davis-Skodje model for $\varepsilon = 0.1$. Various trajectories are shown in the plane (z, y) . Also shown is the reduced manifold \mathcal{M}_0 coinciding with \mathcal{M}_ε .	8
1.2	Michaelis-Menten-Henry system for enzyme kinetics. Various trajectories are shown in the plane (s, c) for $\varepsilon = 0.8$ and $\eta = 0.5$. Also shown are the reduced manifold \mathcal{M}_0 (coinciding with the SSA), the slow invariant manifold \mathcal{M}_ε and the EA.	10
1.3	Successive iterations of the Fraser-Roussel method: Michaelis-Menten system (a)-(b), starting from the initial function $c_0(s) = s$. Shown are the first few iterative functions $c_n(s)$ approximating the SIM \mathcal{M}_ε . (a) $\varepsilon = 5, \eta = 0.9$ (b) $\varepsilon = 0.9, \eta = 0.9$. Davis-Skodje system (c)-(d) $\varepsilon = 0.1$. (c) $y_0(z) = z$, (d) $y_0(z) = 1/z$.	13
1.4	Limit-cycle for the Franceschini-Tebaldi model, $r = 33.6$, on the (x_1, x_3) plane.	17
1.5	FT system: (a) ILDM modal projections $ f^i $, $i = 1, 2, \dots, 5$ of the source term on the eigenvectors. (b) Real part of the eigenvalues $Re(\lambda_i)$.	18
1.6	Reciprocals of the absolute instantaneous value of the eigenvalues of J ; $\varepsilon = 0.01$, $X(0)=0.5$, $Y(0)=0.5$, $Z(0)=0.5$.	42
1.7	All solution trajectories in the phase space are attracted by a 1-D manifold; markers are drawn at constant time intervals; $\varepsilon = 0.01$.	43
1.8	One-dimensional manifold defined as the intersection of the two surfaces of Eq. (1.134)	44
1.9	Methods 3 (left) and 4 (right). An initial identity or random basis converge to the eigenvectors of J (method 1) under the action of consecutive refinements.	45
1.10	Method 2. Single refinement of the eigenvector basis with fixed \tilde{M} .	46
2.1	Number of active species and reactions in sets $S_h(t_n, tol)$ and $R_h(t_n, tol)$.	59
2.2	Left two figures: Explosion limits in terms of ignition and reaction times as a function of initial temperature for a methane/air mixture at $p = 1 \text{ atm}$. Right two figures: Equilibrium temperatures and maximum heat release.	62

2.3	Percentage of original reactions found active in simplified mechanisms $R(tol, S_0)$, $R_1(tol, S_0)$ and corresponding percentage of active species in $S(tol, S_0)$, $S_1(tol, S_0)$ as a function of the tolerance threshold tol imposed on importance indices.	63
2.4	Shown as continuous arrows is the CH_4 depletion pathway for the $\Phi = 1$ simplified mechanism $R(tol = 0.15, [\text{CH}_4, \text{NO}])$. Superimposed as dotted arrows is the same pathway for the $R(tol = 0.09, [\text{CH}_4, \text{NO}])$ which is clearly more complete.	65
2.5	Active species sets $S_h(t_n, tol)$ obtained upon convergence for the $\Phi = 1.0$ case with $S_0 = [\text{CH}_4, \text{NO}]$ and $tol = 0.024$. The time interval shown is $0.1ms$ corresponding to the maximum heat release, or explosion, region. Each filled dot indicates an active species. Also shown, with smaller empty dots, the pointers to CSP radicals.	68
2.6	Active reactions sets $R_h(t_n, tol)$ for each time record t_n obtained upon convergence for the $\Phi = 1.0$ case (top figure) (with $S_0 = [\text{CH}_4, \text{NO}]$ and $tol = 0.15$) and for the $\Phi = 0.3$ case (bottom figure) (with $S_0 = [\text{CH}_4, \text{NO}]$ and $tol = 0.13$). The explosion region is $0.1ms$ long for the $\Phi = 1.0$ case and $1.0ms$ for the $\Phi = 0.3$ case.	69
2.7	Slow importance indices $(I_k^{NO})_{slow}$ for nitric oxide relative to reactions in thermal and N_2O nitric oxide mechanisms. The upper figure refers to the stoichiometric case, the lower to the lean case. Both refer to the explosion region.	70
2.8	Slow importance indices $(I_k^{NO})_{slow}$ for nitric oxide relative to reactions in thermal and N_2O nitric oxide mechanisms. The upper figure refers to the stoichiometric case, the lower to the lean case. Both refer to the near-equilibrium region.	71
2.9	Temperature and mole fractions for $\text{H}_2\text{O}, \text{CO}_2, \text{HCO}$ for various simplified mechanisms (stoichiometric case). Mechanisms are indicated by their tolerance value. Shown in brackets are the number of active species and reactions in each mechanism	73
2.10	Time shifted evolution of X_{NO} for mechanisms (stoichiometric case) yielding up to 42% $E(X_{NO})$ rms errors.	74
2.11	Ignition and reaction times % errors wrt detailed mech. for stoichiometric mechanisms $R(tol, [\text{CH}_4, \text{NO}])$. Ignition and reaction times are defined as the time for the temperature to rise 5% and 95% of the peak temperature respectively.	74
2.12	Top: Rms % errors $E(X_{\text{CH}_4})$ and $E(X_{NO})$ relative to CH_4 and NO instantaneous mole fractions and $E(T)$ as a function of active reactions in the stoichiometric simplified mechanisms. Also shown is the speed-up factor of each mechanism. Bottom: similar analysis on a larger pool of species.	77

2.13 Percentage error $10^2 X_{simp} - X_{GRI3.0} / X_{GRI3.0}$ in equilibrium values for temperature and NO,CO ₂ ,H ₂ O as a function of number of active reactions in simplified mechanism (stoichiometric case)	78
2.14 Top: Log-Log plot of CPU speed-up as a function of the percentage of active reactions and species in simplified (stoichiometric) mechanisms. Bottom: rms error on NO mole fract. $E(X_{NO})$ as a function of the speed-up factor.	79
2.15 rms % errors $E(X)$, % error on reaction times and speed-up factors for optimal mechanisms belonging to the $R[tol, \text{NO}, \text{CH}_4]$ and $R[tol, \text{CO}_2, \text{CH}_4]$ families. Also shown are the error ranges within which mechanisms are optimal.	80
2.16 Shifted solutions for the optimal mechanism $R(tol = 0.04, [\text{CO}_2, \text{CH}_4])$. Solid lines refer to GRI 3.0.	81
2.17 Number of active reactions in sets $R(tol, S_{user})$ (simplified mechanisms without reaction recovery) and in $R(tol, S_{user}) \cup R_{rec}(tol, S_{user})$ (simplified mechanisms with reaction recovery). Also shown is the no. of active species which is the same in both cases.	91
2.18 Premixed stoichiometric ($\Phi = 1.00$) laminar flame speed S_L resulting from a 22-species comprehensive mechanism with variable number of active reactions. Dashed horizontal line represents the flame speed $S_L = 38.8\text{cm/s}$ resulting from GRI-Mech 1.2.	91
2.19 Premixed laminar flame speeds $S_L(\text{cm/s})$ resulting from the detailed GRI Mech. 1.2 mechanism and from the comprehensive simplified mechanisms. The first number in brackets is the number of active species in the mechanism tested, the second is the number of reactions and the third number is the average speed-up factor with respect to the detailed mechanism.	94
2.20 Geometric configuration of the counterflow diffusion flame.	95
2.21 Solution and importance indices for the $\Phi = 1.0$ laminar premixed methane-air flame.	98
2.22 Solution and importance indices for the $\Phi = 0.7$ laminar premixed methane-air flame.	99
2.23 Detailed Gri 1.2 and simplified solutions ($tol = 1.9 \times 10^{-1}$, 19 species, 77 reactions) for the $\Phi = 1.0$ laminar premixed methane-air flame.	100
2.24 (cont'd) Detailed Gri 1.2 and simplified solutions ($tol = 1.9 \times 10^{-1}$, 19 species, 77 reactions) for the $\Phi = 1.0$ laminar premixed methane-air flame.	101

2.25 (a),(b) Peak temperature T_{max} [K] and peak CH_3 mole fraction for various simplified mechs. at different average normal strain rate values $a_n(s^{-1})$ resulting from inflow fuel/oxydizer velocities $u_F = u_O \in [35 \text{ cm/s}, 80 \text{ cm/s}]$; numbers in brackets are active species and reactions in simplified mechanisms; (c)-(f) molar composition along $x(\text{cm})$ for the $u_F = u_O = 70 \text{ cm/s}$ case; (c) CH_4 and O_2 , (d) CH_3 , (e) HCO , (f) H_2O and CO_2	102
3.1 The Hartman-Grobman theorem. (a) stable and unstable manifolds. (b) slow and fast stable manifolds	105
3.2 The role of non-commutativity along a trajectory for the Lorenz system	116
3.3 Vector dynamics for a discrete system	117
3.4 Time behavior of the fast component $ f^1 $ along a system trajectory (starting point $(y_o, z_o)^T = (1, 10)^T$) for the Davis-Skodje model Eq. (3.37) (with $\varepsilon = 0.01$). (a) using the invariant set of vectors/covectors, ;(b) using eigenvectors of the local Jacobian matrix; (c) using CSP vectors/covectors with 1 refinement (starting from eigenvectors of the local Jacobian matrix); (d) using CSP vectors/covectors with 2 refinements (starting from eigenvectors of the local Jacobian matrix). The broken line represents the behavior of $ y(t) - \bar{y}(t) _{\mathcal{M}}$	123
3.5 (a) Solutions to Eq(3.53) and (3.60) with $h = 2$ and $f^2(\mathbf{x}_o) = -z_0 = -1$. (b) Renormalization coefficient $\langle J\mathbf{e}_2, \mathbf{e}_2 \rangle$ and its time average $\langle \langle J\mathbf{e}_2, \mathbf{e}_2 \rangle \rangle$ along a trajectory with $(y_0, z_0) = (1, 1)$	126
3.6 Linear system Eq.(3.81). (a)Behavior of $(J\mathbf{a}_i, \mathbf{a}_i)$ and $(\mathbf{b}^i J, \mathbf{b}^i)$ along a trajectory stemming from $\mathbf{x}_o = (1, 1, 1)^T$ (b)Corresponding behavior of $f^i(\mathbf{x}(t))$, the dotted straight line represents e^{-t} . (c)Folding of vectors $\{\mathbf{a}_i\}$ onto \mathbf{v}_1 and covectors $\{\mathbf{b}^i\}$ onto \mathbf{w}^3	132
3.7 Lorenz system. $\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(t)$ and $\langle (\mathbf{b}^i J, \mathbf{b}^i) \rangle(t)$ along the system's trajectory starting from $\mathbf{x}_o = (10, 10, 10)^T$	134
3.8 Lorenz system. The quantity $\frac{1}{t} \log \ \mathbf{g}(\mathbf{x}(t)) \ \rightarrow 0$ as $t \rightarrow \infty$ ($\mathbf{x}_o = (1, 1, 1)^T$). The system's vector field $\mathbf{g}(\mathbf{x})$ spans the central invariant subspace $E_C(\mathbf{x})$ for which $\Lambda_C = 0$	135
3.9 Application of NTDRB: linear system Eq.(3.81). Figures on the left refer to $\tau = 10^{-4}$, figures on the right to $\tau = 0.1$. (a),(b) Behavior of $(J\mathbf{a}_i, \mathbf{a}_i)$ and $(\mathbf{c}^i J, \mathbf{c}^i)$ along a trajectory stemming from $\mathbf{x}_o = (1, 1, 1)^T$. (c),(d) Folding of vectors \mathbf{a}_i onto \mathbf{v}_i and covectors \mathbf{c}^i onto \mathbf{w}^i	143

3.10 Application of NTDRB: linear system Eq.(3.81). Figure on the left refers to $\tau = 10^{-4}$, figure on the right to $\tau = 0.1$. (a),(b) behavior of $f^i(\mathbf{x}(t))$, the dotted straight lines represent e^{-t} , e^{-3t} and e^{-5t} .	143
3.11 Projection $f^1(t) = \mathbf{c}^1(\mathbf{x})\mathbf{e}_C(\mathbf{x})$ for the Lorenz system along a trajectory. The fact that $\mathbf{c}^1\mathbf{e}_C \neq 0$ is evidence of the non biorthogonality of the system $\{\{E^i(\mathbf{x})\}, \{F^i(\mathbf{x})\}\}$.	145
3.12 Lorenz system. (a) $\langle(J\mathbf{a}_i, \mathbf{a}_i)\rangle(t)$ (b) $\langle(\mathbf{c}^i J, \mathbf{c}^i)\rangle(t)$, $i = 1, 2, 3$ as estimated via NTDRB.	145
3.13 Lorenz system. (a) components (a_1^1, a_1^2, a_1^3) . (b) components (n_1^1, n_1^2, n_1^3) of $\mathbf{n}_1 = \mathbf{a}_1 \times \mathbf{a}_2 / \ \mathbf{a}_1 \times \mathbf{a}_2\ $.	148
3.14 Lorenz system: exterior algebra method for estimating Liapunov exponents (continuous lines). (a) $\langle\mu_1^G\rangle(t) = \langle(J\mathbf{a}_1, \mathbf{a}_1)\rangle(t)$. (b) $\langle\mu_2^G\rangle(t) - \langle\mu_1^G\rangle(t)$. (c) $\langle\mu_3^G\rangle(t) - \langle\mu_2^G\rangle(t)$. Dotted lines refer to vector norm growth rates $\langle(J\mathbf{a}_i, \mathbf{a}_i)\rangle(t)$ as already shown in Fig. 3.12 (a).	150
3.15 Lorenz system. Projections $f^h(\mathbf{x})/\ \mathbf{g}(\mathbf{x})\ = \mathbf{c}^h(\mathbf{x})\mathbf{g}(\mathbf{x})/\ \mathbf{g}(\mathbf{x})\ $.	151
3.16 Franceschini-Tebaldi system. (a) and (b) projections of the limit set on the (x_1, x_3) plane for the periodic and chaotic cases. (c) and (d) corresponding projections $f^h(\mathbf{x})$.	153
3.17 Modal amplitudes $f^6(t)$ through $f^9(t)$ calculated with NTDRB vectors and CSP vectors (eigenvectors).	156
3.18 NTDRB and CSP timescale $\tau_8(t)$.	157

Chapter 1

Computational Singular Perturbation

1.1 Introduction

This chapter will be devoted to the description of the *Computational Singular Perturbation* (CSP) [12, 13, 14, 15, 16, 17, 18, 10, 19] method for the reduction and diagnostics of large systems of chemical kinetics equations. It is a well established fact that detailed and accurate reaction mechanisms, may they be in combustion [20], atmospheric science [21], enzyme kinetics or biochemistry [22], involve a large number of species and reactions acting over a wide spectrum of disparate timescales which can range from nanoseconds to minutes. The large number of species (which can reach the order of the 1000's in hydrocarbon combustion) and the presence of such diverse timescales is a source of severe complexity both in terms of the overall dimension of the system and in its stiffness. This can constitute an unbearable burden on present technology computers in terms of memory storage capabilities and computation time. Stiffness, for example, forces explicit numerical integration schemes to proceed with timesteps of the order of the fastest timescale if stability is to be preserved. Implicit schemes on the other hand allow larger time steps at the cost of a greater algebraic burden. It is to be further pointed out that such a degree of complexity exists in the simplest case of a spatially homogeneous problem where no spatial coordinate is present and the governing equations are constituted by a finite system of ordinary differential equations (finite dimensional dynamical systems). Reactive flow problems such as a turbulent flow field interacting with chemical kinetics pose an even greater challenge. Governing equations in this case are constituted by a system of partial differential equations which generate a so called infinite dimensional dynamical system. Here an extremely fine time *and* space discretization must be devised to accurately solve the fluid field and which, in most cases, is computationally prohibitive when a direct simulation is attempted.

The above reasoning demands the development of methods capable of decreasing the size and complexity of kinetic systems. Numerous such methods have been proposed and implemented over the past years each falling in the broad category of 'reduction methods', where a number of variables (usually termed "slow") can be singled out as evolving on a low dimensional "slow invariant manifold" (SIM) and the remaining (fast) are somehow slaved to the slow variables. Because the dimension of the slow invariant manifold, which can be thought of as an exponentially attractive invariant¹ subspace (hypersurface) of the composition space defining the asymptotic behavior of the system, is smaller than the dimension of the composition (phase) space itself, then the long term dynamics of the system will be governed by the dynamics on the manifold this resulting in a general reduction of complexity. This reduction will be manifested in terms of both a reduction in the overall dimension of the system and in a reduction of its stiffness. All reduction methods are therefore essentially focused on generating the SIM, or an approximation of it, for the system in question. In addition to the CSP reduction method, we will examine the quasi-steady-state approximation (QSSA) [23, 24], the partial-equilibrium approximation [25], the intrinsic low-dimensional manifold method (ILDM) [26, 27] and the iterative Fraser-Roussel method [28, 29, 30]. Other techniques have been proposed, but will not be analyzed here, such as the saddlepoint method [31] which can single out the relevant SIM as a manifold stemming from a saddlepoint of the system, if it exists, situated in an un-physical region of the phase space. Indeed in [6] global slow manifolds are defined as invariant manifolds connecting equilibrium points to saddles at infinity. Moreover a bifurcation of such saddles leads to the concept of 'inverting' manifolds which are common in combustion systems such as thermal explosions and chain branching kinetic systems.

Note that other methods (e.g. [32]) fall under the category of 'simplification methods' whereby reactions which already achieved almost equilibrium conditions and/or which have somehow been declared unimportant are simply discarded from the detailed mechanism. Chapter 2 will illustrate how CSP can be used to construct an automatic simplification algorithm.

Let us now illustrate the proper mathematical setting of reduction methods (a thorough analysis may be found in [33, 34, 35, 19]). A spatially homogeneous kinetic system is represented by a set of ODE's:

$$\frac{d\mathbf{x}}{dt} = \dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) \quad (1.1)$$

¹The term 'invariant' will generally stand for 'invariant with respect to the dynamical system in question'. If the dynamical system is thought of as a mapping of points of the phase space (states) into points at future time instants, then a point on an invariant structure will always be mapped into a point belonging to the same structure, hence its invariance.

where $\mathbf{x}(t)$ is the vector of species concentrations and $\mathbf{g}(\mathbf{x})$ represents the system's vector field whose entries are the species production rates. As we mentioned, reduction methods are effective when the variables (concentrations) are classified into fast and slow. This means that the set of variables evolve on two distinct and widely separated timescales. It is customary to measure such timescale separation by an arbitrarily small parameter ε . In this case Eq.(1.1) acquires a *singularly perturbed form*:

$$\frac{d\mathbf{x}_s}{dt} = \varepsilon \mathbf{g}_s(\mathbf{x}_s, \mathbf{x}_f, \varepsilon) \quad (1.2)$$

$$\frac{d\mathbf{x}_f}{dt} = \mathbf{g}_f(\mathbf{x}_s, \mathbf{x}_f, \varepsilon) \quad (1.3)$$

where $\mathbf{x}_s \in R^m$ is the vector of slow variables, $\mathbf{x}_f \in R^n$ the vector of fast variables and where we assumed that $\mathbf{g}_s \in R^m$ and $\mathbf{g}_f \in R^n$ as well as their derivatives are $\mathcal{O}(1)$. Note that t may be referred to as the fast time as it defines the time scale on which the fast variables evolve. For this reason the system (1.2)-(1.3) may be addressed to as the fast system. By defining the time scale of the slow variables as $\tau = \varepsilon t$ system (1.2)-(1.3) assumes an equivalent (or slow) form:

$$\frac{d\mathbf{x}_s}{d\tau} = \mathbf{g}_s(\mathbf{x}_s, \mathbf{x}_f, \varepsilon) \quad (1.4)$$

$$\varepsilon \frac{d\mathbf{x}_f}{d\tau} = \mathbf{g}_f(\mathbf{x}_s, \mathbf{x}_f, \varepsilon) \quad (1.5)$$

In the limit $\varepsilon \rightarrow 0$, system (1.4)-(1.5) reduces to

$$\frac{d\mathbf{x}_s}{d\tau} = \mathbf{g}_s(\mathbf{x}_s, \mathbf{x}_f, 0) \quad (1.6)$$

$$0 = \mathbf{g}_f(\mathbf{x}_s, \mathbf{x}_f, 0) \quad (1.7)$$

where the first equation describes the evolution of the slow variables and the second equation represents an algebraic constraint which confines the reduced slow system to a *reduced slow manifold* \mathcal{M}_0 . If h_0 is a single vector-valued function (defined in R^m) such that $\mathbf{g}_f(\mathbf{x}_s, h_0(\mathbf{x}_s), 0) = 0$ then the reduced slow manifold can be defined as:

$$\mathcal{M}_0 = \{(\mathbf{x}_s, \mathbf{x}_f)^T \in R^{m+n} | \mathbf{x}_f = h_0(\mathbf{x}_s)\}. \quad (1.8)$$

All reduction methods assume that each point $(\mathbf{x}_s, h_0(\mathbf{x}_s)) \in \mathcal{M}_0$ is an asymptotically stable fixed point for Eq. (1.3).

The Fenichel geometric singular perturbation theory (GSPT) [33, 34, 35, 19], under the above assumptions, guarantees the existence of a *slow invariant manifold* \mathcal{M}_ε for all sufficiently small ε , which is a perturbation of \mathcal{M}_0 , has the same dimensions and is invariant under the system (1.2)-(1.3). Moreover \mathcal{M}_ε is exponentially attractive for all nearby trajectories of system (1.2)-(1.3). The following theorem summarizes these conclusions:

Theorem 1.1.1 *For all sufficiently small ε , there is a function h_ε , such that the graph*

$$\mathcal{M}_\varepsilon = \{(\mathbf{x}_s, \mathbf{x}_f)^T \in R^{m+n} | \mathbf{x}_f = h_\varepsilon(\mathbf{x}_s)\}. \quad (1.9)$$

is invariant under the dynamics of system (1.2)-(1.3). \mathcal{M}_ε is said the slow invariant manifold for such system. The function h_ε admits an asymptotic expansion for $\varepsilon \rightarrow 0$,

$$h_\varepsilon(\mathbf{x}_s) = h_0(\mathbf{x}_s) + \varepsilon h_1(\mathbf{x}_s) + \varepsilon^2 h_2(\mathbf{x}_s) + \dots \quad (1.10)$$

The long-term dynamics of system (1.2)-(1.3) is governed by the equation (1.4)

$$\frac{d\mathbf{x}_s}{d\tau} = \mathbf{g}_s(\mathbf{x}_s, h_\varepsilon(\mathbf{x}_s), \varepsilon) \quad (1.11)$$

constrained on \mathcal{M}_ε , with $\tau = \varepsilon t$ ■.

Note that coefficients h_1, h_2, \dots can be determined by imposing the invariance of the manifold, i.e. substituting the defining Eq. (1.9) into Eq. (1.5) thereby obtaining the following *invariance equation*:

$$\varepsilon Dh_\varepsilon(\mathbf{x}_s) \mathbf{g}_s(\mathbf{x}_s, h_\varepsilon(\mathbf{x}_s), \varepsilon) = \mathbf{g}_f(\mathbf{x}_s, h_\varepsilon(\mathbf{x}_s), \varepsilon) \quad (1.12)$$

where we placed $d\mathbf{x}_f/d\tau = Dh_\varepsilon(\mathbf{x}_s)d\mathbf{x}_s/d\tau$ and where $Dh_\varepsilon(\mathbf{x}_s)$ is intended as the $n \times m$ matrix $\partial h_\varepsilon^i / \partial x_s^j$. If the original system does indeed possess a slow invariant manifold, it may therefore be reduced to a much smaller system for the evolution of the slow variables only, constrained on the manifold \mathcal{M}_ε . The fundamental question of any reduction method is therefore *how accurately it can approximate \mathcal{M}_ε* . We should point out that the explicit singularly perturbed form of system (1.2)-(1.3) is never available in complex chemical-physical systems where even the determination of the slow and fast variables and of the parameter ε is a substantial part of the reduction process. Reduction methods should therefore provide automatic algorithms capable of recovering \mathcal{M}_ε and the asymptotic reduced form (1.11). If, on the other hand, the system does not possess an invariant slow manifold, but rather a more complex asymptotic behavior confined to less trivial invariant sets

(examples of such sets are periodic or chaotic attractors) then conventional reduction methods based on such manifolds lose their effectiveness. Chapter 3 will explore many of such issues.

1.2 A definition of timescale

It is worth, at this preliminary stage, to emphasize the concept of timescales for a non linear system such as (1.1), a concept that will prove crucial in the task of system reduction. The singularly perturbed form (1.4)-(1.5) employs a particular representation of the vector field \mathbf{g} , projecting it onto the constant orthogonal standard basis of the phase (composition) space $\{\mathbf{e}_i\} = (0, \dots, 1, \dots, 0)^T$ where 1 is the i -th entry and $i = 1, n + m$. The vector field in this case will have coordinates $(g_s^1, \dots, g_s^m, \varepsilon^{-1}g_f^{m+1}, \dots, \varepsilon^{-1}g_f^{m+n})^T$ indicating, for each species, two sets of constant timescales of $\mathcal{O}(1)$ and $\mathcal{O}(\varepsilon^{-1})$ with the assumption that the entries of \mathbf{g} are $\mathcal{O}(1)$. In general, however, for systems that do not possess an explicit singularly perturbed form, such as those represented by Eq. (1.1), the parameter ε is not a constant and will be dependent upon the dynamics of the system. A different representation of the vector field will have to be sought requiring some sort of time dependent, not necessarily orthogonal unit basis $\{\mathbf{a}_i\}$ in which the vector field projects as a summation of 'modes' $\mathbf{g} = \sum_{i=1}^{m+n} \mathbf{a}_i f^i$ and its components (or modal amplitudes) are $(f^1, \dots, f^m, f^{m+1}, \dots, f^n)$ where $f^i = \mathbf{b}^i \mathbf{g}$, being $\{\mathbf{b}^i\}$ a set of linear functionals biorthonormal to $\{\mathbf{a}_i\}$ so that $\mathbf{b}^i \mathbf{a}_j = \delta_j^i$ (and $\sum_{i=1}^{m+n} \mathbf{a}_i \mathbf{b}^i = I$, where I is the identity matrix). Following the time evolution of the vector field modal amplitudes we obtain the following equation

$$\dot{f}^i = \dot{\mathbf{b}}^i \mathbf{g} + \mathbf{b}^i \dot{\mathbf{g}} = \dot{\mathbf{b}}^i \mathbf{g} + \mathbf{b}^i J \mathbf{g} = \sum_{j=1}^{m+n} (\dot{\mathbf{b}}^i + \mathbf{b}^i J) \mathbf{a}_j f^j = \Lambda_j^i f^j \quad (1.13)$$

$$\text{with } \Lambda_j^i = \sum_{j=1}^{m+n} (\dot{\mathbf{b}}^i + \mathbf{b}^i J) \mathbf{a}_j \quad (1.14)$$

where $J_j^i = \partial g^i / \partial x^j$ is the Jacobian matrix associated to $\mathbf{g}(\mathbf{x})$. For a linear system in which $\mathbf{g}(\mathbf{x}) = A\mathbf{x}$, A being a constant matrix, the most natural choice for the projection basis would be the set of constant normalized right eigenvectors of A , $\{\mathbf{v}_i\}$. The reason for this being that the time evolution of the vector field amplitudes $f^i = \mathbf{w}^i \mathbf{g}$, with $\{\mathbf{w}^i\}$ being the left eigenvectors of A , reduces to the set of uncoupled equations² $\dot{f}^i = \lambda^i f^i$, where $\{\lambda^i\}$ is the spectrum of eigenvalues of J . Because each amplitude $f^i \sim e^{\lambda^i t}$ evolves independently of each other, the spectrum $\{\lambda^i\}$

²The uncoupling follows from the fact that for a linear system $\mathbf{g}(\mathbf{x}) = A\mathbf{x}$ and therefore $J = A$, i.e the Jacobian is constant along with its eigenvectors. Thus $\mathbf{b}^i = \mathbf{w}^i = 0$ and $\Lambda_j^i = \mathbf{w}^i J \mathbf{v}_j = \text{diag}(\lambda^i)$ where $\{\mathbf{v}_j\}$ are the right eigenvectors of A .

constitutes an unambiguous spectrum of intrinsic timescales³. For non linear systems the effort should therefore focalize on *finding a suitable basis* $\{\mathbf{a}_i\}$ capable of rendering the mode amplitudes' evolution decoupled (or as decoupled as possible) from each other⁴ - in other words capable of eliminating mode mixing - as the scaling of such amplitudes in time provides a clear and unambiguous definition of *intrinsic timescale spectrum* for the system. Only when the latter task is achieved, and the proper basis $\{\mathbf{a}_i\}$ is found, can we perform a temporal ordering of modes which allows a decomposition of the phase space into an m -dimensional slow subspace E_s spanned by the subset of vectors $\{\mathbf{a}_i\}$, $i = 1, m$, associated to the slowest independently evolving amplitudes f_s^i , and an n -dimensional fast subspace E_f spanned by the remaining vectors $\{\mathbf{a}_i\}$, $i = m+1, m+n$ associated with the fastest evolving amplitudes f_f^i , so that $R^{m+n} = E_s \oplus E_f$. Supposing such an *ideal* vector basis is indeed available, the system's vector field may then be rewritten as a sum of independently evolving slow and fast modes: as

$$\frac{d\mathbf{x}}{dt} = \dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) = \mathbf{g}_s(\mathbf{x}) + \mathbf{g}_f(\mathbf{x}) = \sum_{i=1}^m \mathbf{a}_i f_s^i + \sum_{i=m+1}^{m+n} \mathbf{a}_i f_f^i \quad (1.15)$$

where $\mathbf{g}_s \in E_s$ and $\mathbf{g}_f \in E_f$ are the projections of the vector field in the slow and fast subspaces respectively. Clearly the idealized singularly perturbed form (1.4)-(1.5) would only be recovered if $\{\mathbf{a}_i\} = \{\mathbf{e}_i\}$ and if $f_s^i = g_s^i$, $f_f^i = \varepsilon^{-1} g_f^i$, this allowing us to discriminate between slow and fast variables⁵. By analogy, however, we can still define an invariant slow manifold in the general case as the locus of phase space points where the vector field bears no projection onto the fast subspace:

$$\mathcal{K} = \{\mathbf{x} \in R^{n+m} \mid \mathbf{g}_f(\mathbf{x}) = 0\} \quad (1.16)$$

or equivalently as $\mathcal{K} = \{\mathbf{x} \in R^{n+m} \mid f_f^i = 0, i = m+1, m+n\}$, so that the long-term dynamics of the system is governed, on the low-dimensional manifold \mathcal{K} , by the reduced slow system:

$$\frac{d(\mathbf{x} \mid \mathcal{K})}{dt} = \dot{\mathbf{x}} = \mathbf{g}_s(\mathbf{x} \mid \mathcal{K}). \quad (1.17)$$

The preceding definition therefore, extends the concept of the SIM to systems which are not cast in the ideal singularly perturbed fast-slow form. Such a definition however requires the knowledge

³The implicit assumption here, although not strictly needed, is that eigenvalues are real and negative so that each mode is a stable decaying mode.

⁴Clearly the set of eigenvectors of J is not necessarily a reasonable choice as it generally does not provide any decoupling for non linear systems.

⁵The standard basis is constant and each axis is parallel to the corresponding species axis of the phase space.

of the ideal set of projection vectors for the system which defines an invariant decomposition of the phase space and the associated system's spectrum of intrinsic time scales. The set of ideal vectors defines the modal amplitudes f^i which, because of the autonomous nature of the system, constitute a set of scalar fields $f^i(\mathbf{x}) = \mathbf{b}^i(\mathbf{x})\mathbf{g}(\mathbf{x})$. The SIM \mathcal{K} is the locus of points $\mathbf{x} \in R^{n+m}$ where the set of fast modal amplitudes are identically zero, i.e. $f_f^i(\mathbf{x}) = 0$, $i = m+1, m+n$, and thus coincides with the locus of phase space points where the system exhibits only the set of slow intrinsic timescales. From its invariance property it follows that the SIM is indeed a unique trajectory or a set of trajectories of the system. Using a 'lagrangian' approach one may loosely define the SIM as the asymptotically attractive phase space hypersurface for any trajectory $\mathbf{x}(t)$ of the system after the fastest timescales are exhausted and $f_f^i(\mathbf{x}(t)) \approx 0$, $i = m+1, m+n$.

1.3 Slow invariant manifolds for some model problems

Let us illustrate the preceding concepts by means of simple archetypal models. As a first example we consider the planar Davis-Skodje model [36, 10] which can be written as:

$$\begin{aligned} \frac{dz}{d\tau} &= -z \\ \varepsilon \frac{dy}{d\tau} &= -y + \frac{z}{1+z} - \frac{z\varepsilon}{(1+z)^2}. \end{aligned} \quad (1.18)$$

This model is cast in a singularly perturbed form with τ as the slow time, z as the slow variable and y as the fast variable for $0 < \varepsilon < 1$. The system possesses a slow invariant manifold \mathcal{M}_ε stemming from the equilibrium point $(z^*, y^*)^T = (0, 0)^T$ and identified by the graph $y = h_\varepsilon(z) = h_0(z) + \varepsilon h_1(z) + \varepsilon^2 h_2(z) + \dots$. The latter can be found from the invariance Eq. (1.12) which becomes

$$\varepsilon(-z) \left(\frac{dh_0}{dz} + \varepsilon \frac{dh_1}{dz} + \varepsilon^2 \frac{dh_2}{dz} + \dots \right) = \left(-[h_0 + \varepsilon h_1 + \varepsilon^2 h_2 + \dots] + \frac{z}{1+z} \right) - \frac{z\varepsilon}{(1+z)^2}.$$

By equating terms $\mathcal{O}(\varepsilon^k)$ with $k = 0, 1, 2, \dots$ we obtain

$$\begin{aligned} h_0(z) &= \frac{z}{1+z} \\ h_i(z) &= 0 \quad i = 1, 2, \dots \end{aligned} \quad (1.19)$$

In this particular example, we therefore find $\mathcal{M}_\varepsilon \equiv \mathcal{M}_0 = \{(y, z)^T \in R^2 | y = \frac{z}{z+1}, z \geq 0\}$ as

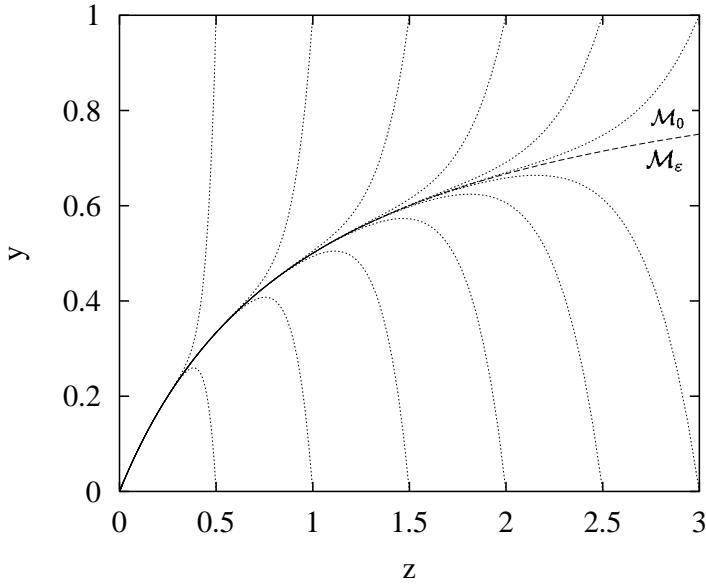
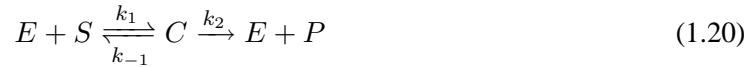


Fig. 1.1: Davis-Skodje model for $\varepsilon = 0.1$. Various trajectories are shown in the plane (z, y) . Also shown is the reduced manifold \mathcal{M}_0 coinciding with \mathcal{M}_ε .

shown in Fig. 1.1.

A second example is the Michaelis-Menten-Henry mechanism for enzyme kinetics introduced in [37, 38]:



Where E is enzyme, S is substrate, C is enzyme-substrate (ES) complex, and P is product. Taking the concentration vector as (E, S, C, P) the system of kinetic equations is

$$\begin{pmatrix} \dot{E} \\ \dot{S} \\ \dot{C} \\ \dot{P} \end{pmatrix} = \begin{pmatrix} -k_1 & k_{-1} + k_2 \\ -k_1 & k_{-1} \\ k_1 & -(k_{-1} + k_2) \\ 0 & k_2 \end{pmatrix} \begin{pmatrix} ES \\ C \end{pmatrix}. \quad (1.21)$$

Only two of the above equations are linearly independent. By adding the first and third we obtain the conservation of enzyme $\dot{E} + \dot{C} = 0$, implying $E(t) + C(t) = E(0) = E_0$. Adding the last three equations we obtain the conservation of substrate $\dot{S} + \dot{C} + \dot{P} = 0$, implying $S(t) + C(t) + P(t) =$

$S(0) = S_0$. Using such conservation equations the system reduces to a planar system:

$$\begin{pmatrix} \dot{S} \\ \dot{C} \end{pmatrix} = \begin{pmatrix} -k_1 E_0 & k_{-1} & k_1 \\ k_1 E_0 & -(k_{-1} + k_2) & -k_1 \end{pmatrix} \begin{pmatrix} S \\ C \\ SC \end{pmatrix}. \quad (1.22)$$

The singularly perturbed form is achieved by defining $K_E = k_{-1}/k_1$ and $K_S = (k_{-1} + k_2)/k_1$ and transforming time t into $\tau = (k_{-1} + k_2)\varepsilon t$ where $\varepsilon = E_0/K_S$. Moreover by defining $c = C/E_0$, $s = S/K_S$ and $\eta = k_2/(k_{-1} + k_2)$ we obtain

$$\begin{aligned} s' &= -s + (s + 1 - \eta)c \\ \varepsilon c' &= s - (s + 1)c \end{aligned} \quad (1.23)$$

where the apex ' indicates differentiation⁶ with respect to τ and the singular perturbation parameter ε satisfies $0 < \varepsilon < 1$. Note that s is the slow variable and c the fast variable. The reduced slow manifold is readily obtained in the limit as $\varepsilon \rightarrow 0$:

$$\mathcal{M}_0 = \{(c, s)^T \in R^2 | c = \frac{s}{s+1}, s \geq 0\}. \quad (1.24)$$

The invariant slow manifold $\mathcal{M}_\varepsilon = \{(c, s)^T \in R^2 | c = h_\varepsilon(s)\}$ is obtained through the invariance equation which in this case reads

$$\varepsilon \frac{dh_\varepsilon}{ds}(-s + (s + 1 - \eta)h_\varepsilon(s)) = s - (s + 1)h_\varepsilon(s) \quad (1.25)$$

where h_ε admits the asymptotic expansion $h_\varepsilon = h_0 + \varepsilon h_1 + \varepsilon^2 h_2 + \dots$. Again by equating terms

⁶Note that differentiation with respect to t (indicated by the dot ·) yields the equivalent fast system:

$$\begin{aligned} \dot{s} &= \varepsilon(k_{-1} + k_2)(-s + (s + 1 - \eta)c) \\ \dot{c} &= (k_{-1} + k_2)(s - (s + 1)c). \end{aligned}$$

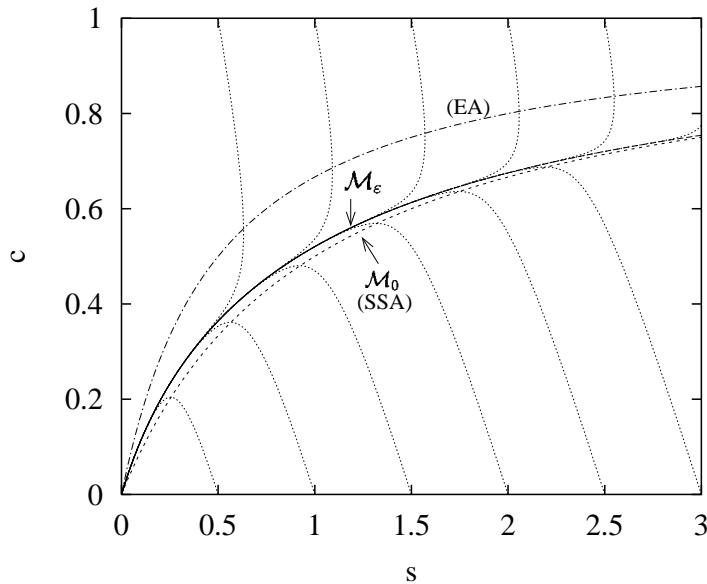


Fig. 1.2: Michaelis-Menten-Henry system for enzyme kinetics. Various trajectories are shown in the plane (s, c) for $\varepsilon = 0.8$ and $\eta = 0.5$. Also shown are the reduced manifold \mathcal{M}_0 (coinciding with the SSA), the slow invariant manifold \mathcal{M}_ε and the EA.

$\mathcal{O}(\varepsilon^k)$ with $k = 0, 1, 2, \dots$ we obtain coefficients h_0, h_1, h_2, \dots The first three coefficients are

$$\begin{aligned} h_0(s) &= \frac{s}{s+1} \\ h_1(s) &= \frac{\eta s}{(s+1)^4} \\ h_2(s) &= \frac{\eta s(2\eta - 3\eta s - s - 1)}{(s+1)^7}. \end{aligned} \tag{1.26}$$

The resulting slow invariant manifold \mathcal{M}_ε can be seen in Fig. 1.2.

1.4 Steady-state and equilibrium assumptions

We should observe that the classical steady-state assumption (SSA) may be applied to the Michaelis-Menten system by imposing a zero variation in the enzyme-substrate complex concentration, i.e. $\dot{C} = 0$, to system (1.22). This readily results in the relation $C = k_1 E_0 S / (k_1 S + (k_{-1} + k_2))$, which becomes $c = s/(s+1)$ after scaling. This means that the SSA confines the system to \mathcal{M}_0 rather than on the SIM \mathcal{M}_ε the two coinciding only in the limit as $\varepsilon \rightarrow 0$.

On the other hand, the equilibrium approximation (EA) requires that the first of the two reactions (1.20) be in equilibrium. This implies that $k_1 E S = k_{-1} C$ which results into $C = k_1 E_0 S / (k_{-1} +$

$k_1 S$). By scaling the variables, we obtain $c = s/(s + K_E/K_S)$ and being $K_E/K_S = 1 - \eta$ we have $c = s/(s + 1 - \eta)$. From Eqs. (1.23) we see that this coincides with the locus $\dot{s} = 0$ (achievable in the limit as $\varepsilon \rightarrow \infty$ in which case s becomes the fast variable).

It is possible to show [28] that the asymptotic behavior of the system, which is represented by the slow evolution on the SIM \mathcal{M}_ε , is bounded below by the SSA (represented by \mathcal{M}_0) and above by the EA, for each $\varepsilon > 0$ as can be seen in Fig.1.2.

1.5 The Fraser Roussel method

Fraser first introduced [39] and later developed with Roussel [29, 28] an iterative method for the determination of the slow invariant manifold \mathcal{M}_ε which is particularly suited for simple planar systems. The method relies on the iterative solution of the invariance equation relative to a given singularly perturbed system of the kind (1.4)-(1.5), i.e.:

$$\begin{aligned}\frac{d\mathbf{x}_s}{d\tau} &= \mathbf{g}_s(\mathbf{x}_s, \mathbf{x}_f, \varepsilon) \\ \varepsilon \frac{d\mathbf{x}_f}{d\tau} &= \mathbf{g}_f(\mathbf{x}_s, \mathbf{x}_f, \varepsilon).\end{aligned}$$

The slow invariant manifold is identified by the graph $\mathbf{x}_f = \mathbf{x}_f(\mathbf{x}_s, \varepsilon)$. Such relation translates into the identity:

$$\frac{d\mathbf{x}_f}{d\tau} = \frac{\partial \mathbf{x}_f}{\partial \mathbf{x}_s} \frac{d\mathbf{x}_s}{d\tau} \quad (1.27)$$

and by imposing invariance we obtain an invariance equation in the form:

$$\mathbf{g}_f(\mathbf{x}_s, \mathbf{x}_f, \varepsilon) = \varepsilon \mathbf{x}_{f,\mathbf{x}_s}(\mathbf{x}_s, \varepsilon) \mathbf{g}_s(\mathbf{x}_s, \mathbf{x}_f, \varepsilon). \quad (1.28)$$

In the case of planar systems on (x_s, x_f) , where $x_s(\tau)$ and $x_f(\tau)$ are scalar functions, the invariance equation can be easily rearranged into a functional relation of the kind:

$$x_f = x_f(x_s, \varepsilon) = F(x_s, x_{f,x_s}). \quad (1.29)$$

In the Fraser-Roussel method such an equation is solved iteratively, starting from an initial function $(x_f)_0$, so that after n iterations one obtains:

$$(x_f)_{n+1} = F(x_s, (x_f)_{n,x_s}). \quad (1.30)$$

Let us apply the method to the Michaelis-Menten system Eq. (1.23). The manifold \mathcal{M}_ε has the form $c = c(s)$ and substituting this into the invariance equation (1.28) we obtain

$$s - c - sc = \varepsilon c_{,s} (-s + (s + 1 - \eta)c). \quad (1.31)$$

Rearranging this we obtain the recursive expression similar to Eq. (1.30):

$$c_{n+1}(s) = \frac{s}{1 + s - \frac{\eta \varepsilon c_{n,s}(s)}{1 + \varepsilon c_{n,s}(s)}}. \quad (1.32)$$

By starting the iteration procedure from the initial trial function $c_0(s)$ we obtain a set of successive functions $c_n(s)$, $n = 1, 2, \dots$ which we can express as asymptotic expansions: $c_n(s) = c_n^{(0)} + \varepsilon c_n^{(1)} + \varepsilon^2 c_n^{(2)} + \varepsilon^3 c_n^{(3)} + \dots$. Choosing $c_0(s) = s$ we obtain the set of successive manifold approximations shown in Table 1.1 in terms of coefficients $c_n^{(i)}$. Such coefficients should be compared with coefficients $h_i(s)$ of Eq. (1.26) appearing in the expansion of $h_\varepsilon(s)$ which defines the SIM \mathcal{M}_ε . Note that each Fraser-Roussel iteration generates one order of the asymptotic expansion of \mathcal{M}_ε . Fig. 1.3 shows the first few iterates $c_n(s)$ for different values of ε .

$c_n(s)$	$\mathcal{O}(1)$	$\mathcal{O}(\varepsilon)$	$\mathcal{O}(\varepsilon^2)$	$\mathcal{O}(\varepsilon^3)$
$c_0(s)$	$c_n^{(0)}$	$c_n^{(1)}$	$c_n^{(2)}$	$c_n^{(3)}$
	s	0	0	0
$c_1(s)$	$\frac{s}{1+s}$	$\frac{s \eta}{(1+s)^2}$	$-\frac{s \eta (-1-s+\eta)}{(1+s)^3}$	$\frac{s (1+s-\eta)^2 \eta}{(1+s)^4}$
$c_2(s)$	$\frac{s}{1+s}$	$\frac{s \eta}{(1+s)^4}$	$-\frac{s \eta (1+s+(-2+s(-1+s+s^2)) \eta)}{(1+s)^7}$	$\frac{s \eta ((1+s)^2+(1+s)(-5+s^2(4+s(3+s))) \eta -(2+s(2+s))(-2+s^2(3+2s)) \eta^2}{(1+s)^{10}}$
$c_3(s)$	$\frac{s}{1+s}$	$\frac{s \eta}{(1+s)^4}$	$-\frac{s \eta (1+s-2 \eta+3 s \eta)}{(1+s)^7}$	$\frac{s \eta ((1+s)^2+(1+s)(-5+11s) \eta +(5+s(-16-8s+3s^3)) \eta^2)}{(1+s)^{10}}$
$c_4(s)$	$\frac{s}{1+s}$	$\frac{s \eta}{(1+s)^4}$	$-\frac{s \eta (1+s-2 \eta+3 s \eta)}{(1+s)^7}$	$\frac{s \eta ((1+s)^2+(1+s)(-5+11s) \eta +(5+3s(-8+5s)) \eta^2)}{(1+s)^{10}}$

Table 1.1: Successive iterations of the Fraser-Roussel method for the Michaelis-Menten system, starting from the initial function $c_0(s) = s$. Shown are the first four coefficients $c_n^{(i)}$ in the expression $c_n(s) = c_n^{(0)} + \varepsilon c_n^{(1)} + \varepsilon^2 c_n^{(2)} + \varepsilon^3 c_n^{(3)} + \dots$

We complete our survey of the Fraser-Roussel method with yet another example regarding the Davis-Skodje system seen in Eq. (1.18). This system admits an invariant slow manifold in the form $y = y(z)$ and an invariance equation which reads:

$$-y + \frac{z}{1+z} - \frac{z\varepsilon}{(1+z)^2} = -\varepsilon y_{,z} z, \quad (1.33)$$

$y_n(z)$	$\mathcal{O}(1)$	$\mathcal{O}(\varepsilon)$	$\mathcal{O}(\varepsilon^2)$	$\mathcal{O}(\varepsilon^3)$
$y_n^{(0)}$	$y_n^{(0)}$	$y_n^{(1)}$	$y_n^{(2)}$	$y_n^{(3)}$
$y_0(z)$	z	0	0	0
$y_1(z)$	$\frac{z}{1+z}$	$\frac{z^2(2+z)}{(1+z)^2}$	0	0
$y_2(z)$	$\frac{z}{1+z}$	0	$\frac{z^2(4+z(3+z))}{(1+z)^3}$	0
$y_3(z)$	$\frac{z}{1+z}$	0	0	$\frac{z^2(8+z(5+z(4+z)))}{(1+z)^4}$
$y_4(z)$	$\frac{z}{1+z}$	0	0	0

Table 1.2: Successive iterations of the Fraser-Roussel method for the Davis-Skodje system, starting from the initial function $y_0(z) = z$. Shown are the first four coefficients $y_n^{(i)}$ in the expression $y_n(z) = y_n^{(0)} + \varepsilon y_n^{(1)} + \varepsilon^2 y_n^{(2)} + \varepsilon^3 y_n^{(3)} + \dots$

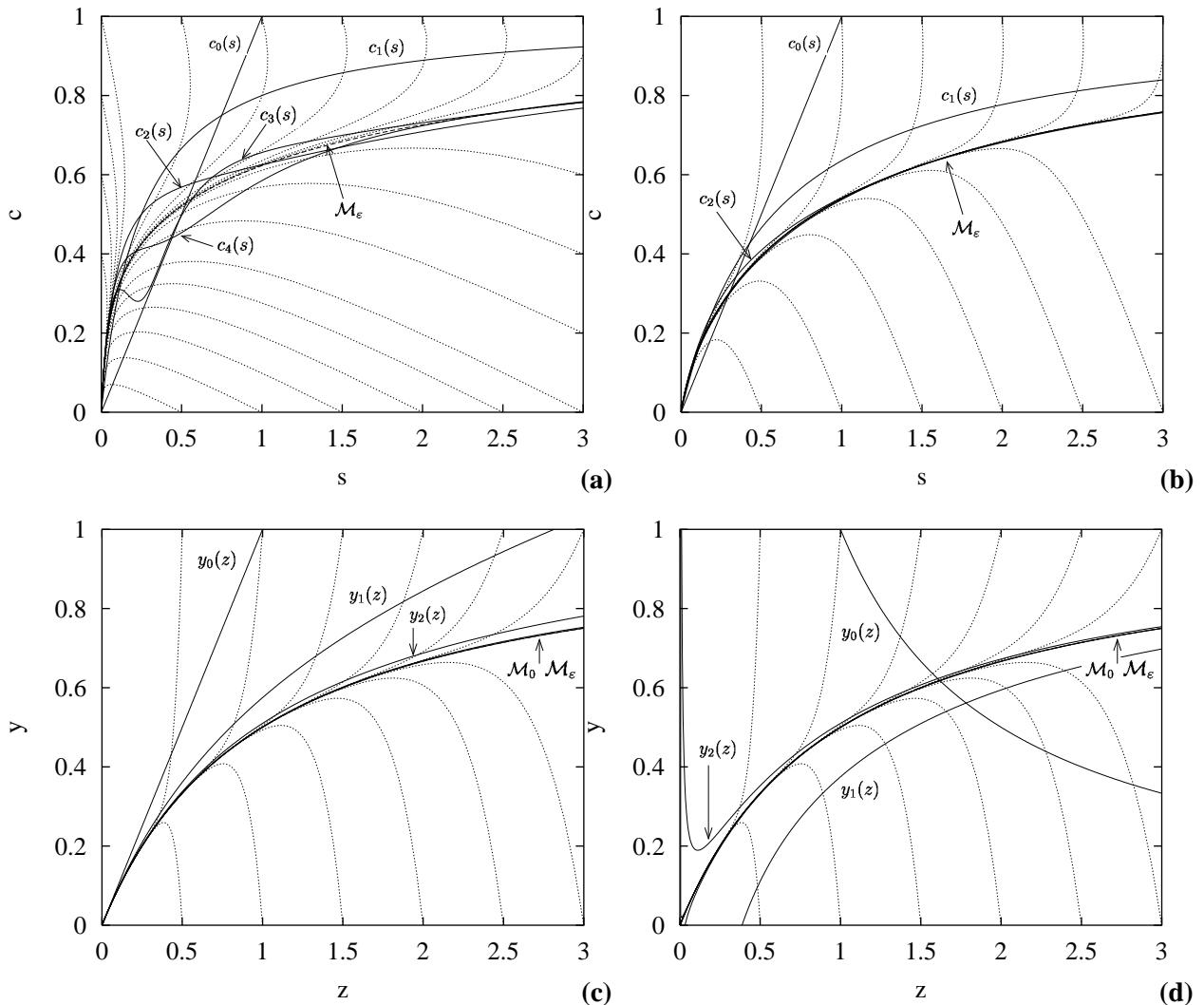


Fig. 1.3: Successive iterations of the Fraser-Roussel method: Michaelis-Menten system (a)-(b), starting from the initial function $c_0(s) = s$. Shown are the first few iterative functions $c_n(s)$ approximating the SIM \mathcal{M}_ε . (a) $\varepsilon = 5, \eta = 0.9$ (b) $\varepsilon = 0.9, \eta = 0.9$. Davis-Skodje system (c)-(d) $\varepsilon = 0.1$. (c) $y_0(z) = z$, (d) $y_0(z) = 1/z$.

which is readily solved for y to yield the following iterative equation:

$$y_{n+1}(z) = \varepsilon y_{n,z}(z) z - \frac{z\varepsilon}{(1+z)^2} + \frac{z}{1+z}. \quad (1.34)$$

As with the Michaelis-Menten system we can choose a linear initial trial function $y_0(z) = z$ and determine the successive iterates $y_n(z)$ in the form of asymptotic expansions $y_n(z) = y_n^{(0)} + \varepsilon y_n^{(1)} + \varepsilon^2 y_n^{(2)} + \varepsilon^3 y_n^{(3)} + \dots$. Results are shown in Table 1.2 and Fig. 1.3(c). Note that the Davis-Skodje system has a SIM \mathcal{M}_ε which coincides with the reduced manifold \mathcal{M}_0 (corresponding to $y = h_\varepsilon(z) = h_0(z) = z/(1+z)$) as shown by the coefficients in Eqs. (1.19). Every iteration, therefore, zeroes-out a higher order coefficient $y_n^{(i)}$, $i > 0$, as is clear in Table 1.2. Results do not change qualitatively if we choose a different initial trial function such as $y_0(z) = 1/z$, as can be seen in Fig. 1.3(d). We should note, however, that convergence is not guaranteed *a priori* with the Fraser-Roussel method. It can be shown[35], however, that for the special class of singularly perturbed planar systems and by choosing an initial trial function $x_{f0} = h_0(x_s)$ (being h_0 the first coefficient of the asymptotic expansion of \mathcal{M}_ε), the method recovers term by term the asymptotic expansion of \mathcal{M}_ε .

1.6 The ILDM method

The intrinsic low dimensional manifold (ILDM) was first introduced by Maas and Pope (1992) [26]. As with the other methods examined (and as in the case of CSP), ILDM is based on the existence of a significant separation between the slow and fast timescale spectra so that a slow manifold may be defined for the system. ILDM generally applies to systems represented by Eq. (1.1) that do not possess an explicit singularly perturbed form. As seen in section 1.2, for such systems the question arises as to how to define an intrinsic spectrum of timescales and an associated suitable projection basis so that a slow invariant manifold may be defined.

The ILDM method uses the right eigenvectors of the local Jacobian matrix ⁷ $J_j^i = \partial g^i / \partial x_j$ of the vector field $\mathbf{g}(\mathbf{x})$ as the projection basis. Such eigenvectors are ordered on the basis of the real part of the corresponding eigenvalue which, by assumption, is negative. ILDM, therefore, identifies the spectrum of intrinsic timescales of the system as the eigenvalue spectrum. Consequently, the set of eigenvectors associated to the, say n , fastest (most negative) timescales will span a fast subspace and those associated to the m slowest (least negative) timescales will span a slow subspace. The vector field will therefore decompose into a slow and a fast projection: $\mathbf{g}(\mathbf{x}) = \mathbf{g}_s(\mathbf{x}) + \mathbf{g}_f(\mathbf{x}) =$

⁷To be more precise the ILDM method uses the set of Schur vectors resulting from a Schur decomposition of J , i.e. $J = QNQ'$, where N is upper triangular.

$\sum_{i=1}^m \mathbf{a}_i f_s^i + \sum_{i=m+1}^{m+n} \mathbf{a}_i f_f^i$ and the ensuing slow manifold, similarly to what we have seen in section 1.2, will be defined as the locus $\mathbf{g}_f(\mathbf{x}) = 0$ which is the defining equation of the ILDM, i.e. $\mathcal{K} = \{\mathbf{x} \in R^{n+m} \mid f_f^i = 0, i = m+1, m+n\}$.

A rigorous mathematical assessment of the accuracy of the ILDM \mathcal{K} has been carried out [35] for an idealized singularly perturbed system as the one represented by Eqs. (1.4)-(1.5), which possesses an invariant slow manifold \mathcal{M}_ε represented by the graph $\mathbf{x}_f = h_\varepsilon(\mathbf{x}_s)$. Supposing the ILDM \mathcal{K} ($\mathbf{g}_f(\mathbf{x}) = 0$) admits an asymptotic solution in the form of a power series expansion as $\varepsilon \rightarrow 0$:

$$\mathbf{x}_f = \psi_\varepsilon(\mathbf{x}_s) = \psi_0(\mathbf{x}_s) + \varepsilon \psi_1(\mathbf{x}_s) + \varepsilon^2 \psi_2(\mathbf{x}_s) + \dots \quad (1.35)$$

this is an approximation of \mathcal{M}_ε , i.e. of its expansion $h_\varepsilon(\mathbf{x}_s)$ shown in Eq. (1.9), which is accurate up to *and including* the $\mathcal{O}(\varepsilon)$ terms. The error on the $\mathcal{O}(\varepsilon^2)$ term is proportional to the curvature of the reduced manifold \mathcal{M}_0 , i.e. of $h_0(\mathbf{x}_s)$. This can be restated by saying:

$$\begin{aligned} \psi_0 &= h_0 \\ \psi_1 &= h_1 \\ \psi_2 - h_2 &\propto D^2 h_0. \end{aligned}$$

As a first example we consider the Davis-Skodje system Eq. (1.18), re-written in the form of Eq. (1.1) as

$$\begin{aligned} \frac{dz}{dt} &= -z \\ \frac{dy}{dt} &= \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2} \end{aligned} \quad (1.36)$$

where t here indicates slow time. The local Jacobian of the associated vector field reads:

$$J(\mathbf{x}) = \begin{pmatrix} -1 & 0 \\ \frac{1+z-\varepsilon+z\varepsilon}{(1+z)^3\varepsilon} & -\frac{1}{\varepsilon} \end{pmatrix} \quad (1.37)$$

which has a constant 'slow' eigenvalue $\lambda_s = -1$ and a constant 'fast' eigenvalue $\lambda_f = -1/\varepsilon$. The associated slow and fast un-normalized eigenvectors are

$$\mathbf{v}_s = \mathbf{a}_1 = (1, ((1+z) - \varepsilon(1-z))(1-\varepsilon)^{-1}(1+z)^{-3})^T \quad ; \quad \mathbf{v}_f = \mathbf{a}_2 = (0, 1)^T \quad (1.38)$$

where the same notation of section 1.2 was used. The set of corresponding biorthonormal dual row vectors are:

$$\mathbf{b}^1 = (1, 0) \quad ; \quad \mathbf{b}^2 = ((1+z) - \varepsilon(1-z))(1-\varepsilon)^{-1}(1+z)^{-3}, 1). \quad (1.39)$$

The projections of $\mathbf{g}(\mathbf{x})$ onto the basis $\{\mathbf{a}_1, \mathbf{a}_2\}$ are:

$$f_s = \mathbf{b}^1 \cdot \mathbf{g} = -z \quad ; \quad f_f = \mathbf{b}^2 \cdot \mathbf{g} = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) + \frac{\varepsilon}{1-\varepsilon} \frac{2z^2}{(1+z)^3} \quad (1.40)$$

where $\mathbf{g}_s = \mathbf{a}_1 f_s$ and $\mathbf{g}_f = \mathbf{a}_2 f_f$.

This decomposition defines the ILDM \mathcal{K} as the locus $\mathbf{g}_f(y, z) = 0$ i.e. as the locus $f_f(y, z) = 0$ which can be solved for y to yield

$$y = \frac{z}{1+z} + \frac{\varepsilon^2}{1-\varepsilon} \frac{2z^2}{(1+z)^3} \quad (1.41)$$

whose asymptotic expansion is

$$\begin{aligned} y = \psi(z) &= \psi_0(z) + \varepsilon\psi_1(z) + \varepsilon^2\psi_2(z) + \dots = \\ &= \frac{z}{1+z} + \varepsilon^2 \frac{2z^2}{(1+z)^3} + \varepsilon^3 \frac{2z^2}{(1+z)^3} + \dots \end{aligned} \quad (1.42)$$

This must be compared with the Davis-Skodje invariant slow manifold \mathcal{M}_ε identified by $y = h_\varepsilon(z) = h_0(z) = z/(1+z)$ thus obtaining

$$\begin{aligned} \psi_0 &= h_0 = \frac{z}{1+z} \\ \psi_1 &= h_1 = 0 \\ \psi_2 - h_2 &= \frac{2z^2}{(1+z)^3}. \end{aligned}$$

We conclude therefore that the ILDM \mathcal{K} agrees with \mathcal{M}_ε up to and including $\mathcal{O}(\varepsilon)$ terms, and introducing an error $\psi_2 - h_2$ which is $\mathcal{O}(\varepsilon^2)$. Note that such error is proportional to the curvature $d^2 h_0/dz^2 = -2/(1+z)^3$ of $h_0(z)$.

The ILDM method has the clear advantage, over other methods seen thus far, of not needing the system to be cast into the explicit singularly perturbed form, hence its applicability is in principle extendable to any spatially homogeneous reactive system regardless of its complexity. The choice of the projection basis is however arguable. It is in essence founded on an analogy with linear

systems for which the eigenvectors of the Jacobian J of the vector field yield a complete decoupling of modes and is therefore an ideal choice, as seen in section 1.2. Extending such a choice to non linear systems is however not appropriate as this introduces a certain degree of mode mixing which renders the discrimination between fast and slow modes and the determination of a timescale spectrum ambiguous tasks and ultimately, as we have seen, leads to an error in the assessments of the slow invariant manifold. We will see in chapter 3 that the choice of the eigenbasis of J as a projection basis is ultimately inadequate because it does not exhibit the property of invariance with respect to the system, a property, on the other hand, exhibited by the system's slow invariant manifold. If we further extend this concept to systems not possessing an equilibrium point, like those seen so far, nor a slow manifold leading to it, but rather a richer asymptotic behavior such as a persistent oscillation on a one-dimensional limit cycle, then the choice becomes even more obscure. To illustrate this point let us consider the five-dimensional Franceschini-Tebaldi (FT) model[40], which will be re-examined in chapter 3:

$$\begin{aligned}\dot{x}_1 &= -2x_1 + 4x_2 x_3 + 4x_4 x_5 \\ \dot{x}_2 &= -9x_2 + 3x_1 x_3 \\ \dot{x}_3 &= -5x_3 - 7x_1 x_2 + r \\ \dot{x}_4 &= -5x_4 - x_1 x_5 \\ \dot{x}_5 &= -x_5 - 3x_1 x_4\end{aligned}$$

For $r = 33.6$, this system exhibits the limit-cycle oscillations shown in Fig.1.4.

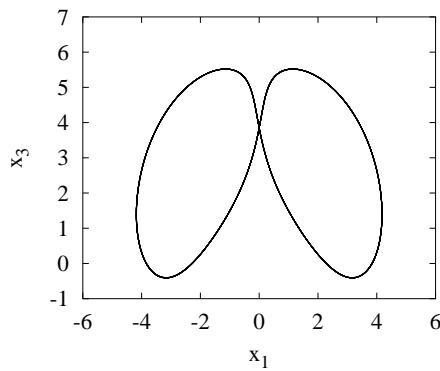


Fig. 1.4: Limit-cycle for the Franceschini-Tebaldi model, $r = 33.6$, on the (x_1, x_3) plane.

The ILDM method can be easily applied to the FT-system by projecting the source term onto the

eigenbasis of its Jacobian matrix. The ensuing modal amplitudes f^i along the periodic orbit can be observed in Fig. 1.5(a). It is clear that none of the amplitudes exhibits a distinct decaying behavior which would cause them to vanish. On the other hand the FT model exhibits a clear asymptotic behavior which is confined on an exponentially attractive one-dimensional subspace (manifold), the limit-cycle, which would suggest that four modal amplitudes should decay according to a set of intrinsic fast timescales. Assuming such an intrinsic timescale spectrum to be the eigenvalue spectrum of the Jacobian is misleading as it consists of both real and complex conjugate branches with several recurring bifurcation points (Fig 1.5(b)), making it impossible to determine a set of slow and fast timescales associated to stable decaying modes. The asymptotic features of the system are therefore ultimately not detected by the ILDM method, the reason being essentially due to the severe mode mixing in the evolution of the amplitudes, which is governed by Eq. (1.14), introduced by the very choice of the eigenbasis as a projection basis for the vector field. This suggests that some other choice must be pursued for the basis vectors to, at least, diminish such coupling.

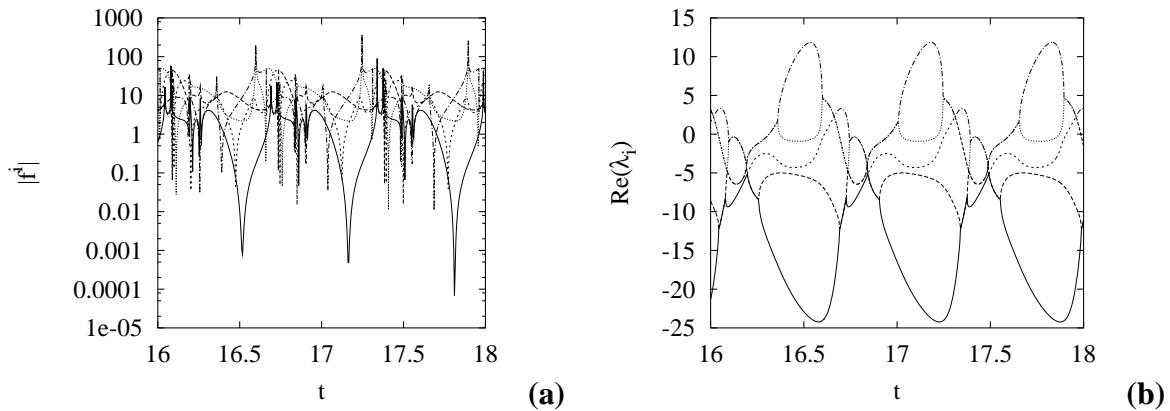


Fig. 1.5: FT system: (a) ILDM modal projections $|f^i|$, $i = 1, 2, \dots, 5$ of the source term on the eigenvectors. (b) Real part of the eigenvalues $\text{Re}(\lambda_i)$.

1.7 An Ideal choice for basis vectors

Let us now give an answer to the question: what is an ideal choice for a projection basis? We partly addressed the issue in section 1.2 and in the previous section; here we will give a formal definition and an example which will serve as an introduction to the CSP method.

Consider the dynamical system Eq. (1.1)

$$\frac{d\mathbf{x}}{dt} = \dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) \quad (1.43)$$

with $\mathbf{x} \in C \subseteq R^n$ and whose solution at time t starting from $\mathbf{x}(t = 0) = \mathbf{x}_0$ may be written as $\mathbf{x}(t) = \phi_t(\mathbf{x}_0)$. Let $\{\mathbf{a}_i(\mathbf{x})\}$, $i = 1, n$, be a set of n linearly independent column basis vectors (not necessarily orthogonal) spanning the tangent space $TC_{\mathbf{x}}$ at \mathbf{x} . Let $\{\mathbf{b}^i(\mathbf{x})\}$, $i = 1, n$, be a set of n linearly independent row basis covectors (linear functionals or 1-forms), spanning the dual of the tangent space $TC_{\mathbf{x}}^*$ so that the system $(\{\mathbf{a}_i(\mathbf{x})\}, \{\mathbf{b}^i(\mathbf{x})\})$ is a system of biorthonormal vectors and covectors, i.e.:

$$\mathbf{b}^i(\mathbf{x}) \mathbf{a}_j(\mathbf{x}) = \sum_{l=1}^n b_l^i a_j^l = \delta_j^i \quad (1.44)$$

where $\mathbf{b}^i = (b_1^i, b_2^i, \dots, b_n^i)$ and $\mathbf{a}_j = (a_j^1, a_j^2, \dots, a_j^n)^T$. Let us now project the system's vector field onto the set of basis vectors $\{\mathbf{a}_i(\mathbf{x})\}$. The system may then be re-written as:

$$\frac{d\mathbf{x}}{dt} = \mathbf{g}(\mathbf{x}) = \sum_{i=1}^n \mathbf{a}_i(\mathbf{x}) f^i(\mathbf{x}) \quad (1.45)$$

$$f^i(\mathbf{x}) = \mathbf{b}^i(\mathbf{x}) \mathbf{g}(\mathbf{x}), \quad (1.46)$$

being $f^i(\mathbf{x})$ the amplitude of the i -th mode $\mathbf{a}_i(\mathbf{x})$. Note that for the amplitude f^i to be a meaningful representative of the contribution of the i -th mode to the vector field $\mathbf{g}(\mathbf{x})$, the set $\{\mathbf{a}_i(\mathbf{x})\}$ should be a unit basis ($\|\mathbf{a}_i\| = 1$). For the sake of clarity the normalization of $\{\mathbf{a}_i(\mathbf{x})\}$ may be sometimes omitted in what follows. It should be kept in mind, however, that it is an essential procedure when a numerical algorithm is implemented.

Amplitudes f^i evolve in time. Differentiating Eq. (1.46) with respect to time along a solution $\mathbf{x}(t) = \phi_t(\mathbf{x}_0)$ of Eq. (1.43) we obtain a nonlinear system of ODE's:

$$\frac{df^i}{dt} = \sum_{j=1}^n \Lambda_j^i f^j, \quad \Lambda_j^i = \left(\frac{d\mathbf{b}^i}{dt} + \mathbf{b}^i \mathbf{J} \right) \mathbf{a}_j \quad (1.47)$$

where $\mathbf{J}(\mathbf{x})$ is the Jacobian $(\partial g^i / \partial x_j)$ of $\mathbf{g}(\mathbf{x})$. We should note at this stage that Eq. (1.47) is not subject to any approximation. On the contrary Eq. (1.47) is fully equivalent to Eq. (1.43), its nonlinearity being fully retained.

We define as *ideal* the system of basis vectors $(\{\mathbf{a}_i(\mathbf{x})\}, \{\mathbf{b}^i(\mathbf{x})\})$ such that the matrix Λ in Eq. (1.47) is diagonal, i.e. $\Lambda = \text{diag}(\mu_i)$. If this is the case then system (1.47) is decoupled and each amplitude evolves independently:

$$\frac{df^i}{dt} = \mu_i f^i \quad (1.48)$$

The elimination of mode coupling makes it possible to define a spectrum of intrinsic local timescales for the system as $\{1/\mu_i(t)\}$. In point of fact, we have from Eq. (1.48) that $f^i \sim e^{\int_0^t \mu_i(\tau) d\tau} =$

$e^{t\langle\mu_i(t)\rangle}$ so, more appropriately, the intrinsic timescale spectrum is $\{1/\langle\mu_i\rangle(t)\}$ where $\langle\cdot\rangle$ indicates time-average. The knowledge of an unambiguous timescale spectrum makes it possible to discriminate between fast and slow modes so that the asymptotic dynamics of the system, if it possesses a set of n fast and a set of m slow timescales, may be written, as seen in section 1.2, as:

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^m \mathbf{a}_i(\mathbf{x}) f^i(\mathbf{x}) \quad (1.49)$$

where the contribution of the n fast modes is neglected. Such asymptotic dynamics is constrained on a slow invariant manifold \mathcal{K} defined by the n relations:

$$f^i(\mathbf{x}) = \mathbf{b}^i(\mathbf{x}) \mathbf{g}(\mathbf{x}) = 0, \quad i = 1, \dots, n. \quad (1.50)$$

Let us now resort to the Davis-Skodje model, in the form of Eq. (1.36), to illustrate these concepts. It can be easily shown⁸ that such system possesses the following set of ideal basis vectors and covectors:

$$\begin{aligned} \mathbf{a}_1 &= (1, (1+z)^{-2})^T & \mathbf{b}^1 &= (1, 0) \\ \mathbf{a}_2 &= (0, 1)^T & \mathbf{b}^2 &= ((-1+z)^{-2}, 1) \end{aligned} \quad (1.51)$$

That these vectors are ideal follows from the resulting matrix Λ which attains the diagonal form

$$\Lambda = \begin{pmatrix} -1 & 0 \\ 0 & -\frac{1}{\varepsilon} \end{pmatrix} \quad (1.52)$$

thereby showing that the amplitudes' dynamics is uncoupled when we project on the above ideal basis. In particular we find that the amplitudes decay according to a slow timescale $1/\mu_1 = 1$ and a fast timescale $1/\mu_2 = \varepsilon$ (which in this particular case are constants):

$$\begin{aligned} f^1 &= \mathbf{b}^1 \mathbf{g} = -z = -z_o e^{-t} = f_0^1 e^{-t} \sim e^{-t} \\ f^2 &= \mathbf{b}^2 \mathbf{g} = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) = \frac{1}{\varepsilon} \left(-y_o + \frac{z_o}{1+z_o} \right) e^{-t/\varepsilon} = f_0^2 e^{-t/\varepsilon} \sim e^{-t/\varepsilon}. \end{aligned} \quad (1.53)$$

⁸An immediate way to find the ideal basis vectors is to 'cheat' by letting the equation of the manifold \mathcal{K} to be the known relation $y = z/(1+z)$. Because $\mathbf{b}^2 \mathbf{g}(\mathbf{x}|_{\mathcal{K}}) = \mathbf{0}$ from this we find $b_1^2/b_2^2 = -1/(1+z)^2$. Clearly this is not a generally viable approach.

The manifold is found by enforcing $f^2 = 0$, which yields $y = z/(1 + z)$. Thus the manifold-constrained asymptotic dynamics of the system reduces to $\dot{\mathbf{x}} = \mathbf{a}_1 f^1$, i.e.

$$\dot{z} = -z \quad , \quad \dot{y} = -z/(1 + z)^2, \quad (1.54)$$

which is solved trivially as $z = e^{-t} z_o$ and $y = e^{-t} z_o / (1 + e^{-t} z_o)$. Note that the complete Davis-Skodje system has an analytical phase flow given by

$$\mathbf{x}(t) = (z, y)^T = \phi_t(\mathbf{x}_o) = \left(e^{-t} z_o, \frac{e^{-t} z_o}{1 + e^{-t} z_o} + e^{-t/\varepsilon} \left(y_o - \frac{z_o}{1 + z_o} \right) \right)^T, \quad (1.55)$$

from which it is clear that the locus of initial conditions $y_0 = z_0/(1 + z_0)$ is invariant under the system's dynamics and coincides with the system's trajectory exhibiting only one slow timescale, which summarizes the definition of slow invariant manifold.

Because the manifold is unique, so are the ideal basis vectors. Any other choice of vectors would then induce some degree of mode coupling. If, for example, the eigenvectors of J were chosen for the Davis-Skodje system, as in ILDM, matrix Λ would read

$$\Lambda = \begin{pmatrix} -1 & 0 \\ -\frac{(2-\varepsilon)z(1+z)-\varepsilon 3z(1-z)}{(1-\varepsilon)(1+z)^4} & -\frac{1}{\varepsilon} \end{pmatrix} \quad (1.56)$$

showing residual coupling of the fast amplitude f^2 with the slow amplitude f^1 . The fundamental reason for this is the presence of term $d\mathbf{b}^i/dt$ in the operator Λ in Eq. (1.47). Had this term been neglected, as the ILDM method essentially does, the resulting operator Λ would indeed be diagonal (right and left eigenvectors diagonalise the Jacobian) but it would also be an approximation of the actual operator driving the amplitudes' dynamics. It would in fact be an operator *similar* to J which is indeed a linear approximation, this explaining why the ILDM never captures the terms $\mathcal{O}(\varepsilon^2)$ or higher in the asymptotic expansion of \mathcal{M}_ε . We will show that the term $d\mathbf{b}^i/dt$ depends upon the more significant term dJ/dt representing the variation of the Jacobian along a trajectory, which certainly is different from zero in non-linear problems. It is the retention of this term that allows, as we will see, the CSP method to capture nonlinearities and thus approximate the slow manifold up to any order of accuracy.

1.8 The CSP method.

The CSP method, originally developed by Lam and Goussis for kinetic problems, is generally effective in all multiple time scale problems possessing an asymptotically attractive invariant slow manifold leading to a stable equilibrium point. The general form of such problems is a set of stiff ODE's

$$\frac{dx}{dt} = \dot{x} = g(x) \quad (1.57)$$

where the stiffness is manifested by the presence of a gap in the eigenvalue spectrum of the Jacobian of the system's vector field. As repeatedly mentioned in section 1.2 and in the previous section 1.7, the goal of CSP is to express the vector field $g(x)$ in a suitable and hopefully ideal (in the meaning of the previous section) basis in which the modal amplitudes, i.e. the coordinates of g with respect to such basis, evolve independently of each other or in which, at most, the fast evolving amplitudes are decoupled from the slow and vice versa. CSP accomplishes the task of determining such basis through an iterative refinement procedure starting from an initial arbitrary basis. The effect of each refinement is to decrease, in norm, the off-diagonal blocks of matrix Λ in Eq. (1.47), governing the dynamics of amplitudes, by an amount of the order of the current time scale separation, indicated with ε , between the slower of the fast amplitudes and the fastest of the slow amplitudes, a measure of such timescales being the diagonal entries of Λ . Each basis vector refinement consists of two steps, the first dealing with the decoupling of the fast amplitudes from the slow (this increasing the accuracy of the current CSP manifold), the second dealing with the decoupling of the slow amplitudes from the fast (this decreasing the stiffness of the simplified manifold-constrained system). We will show that after q refinements the CSP method generates term by term the asymptotic expansion of the invariant slow manifold \mathcal{M}_ε up to and including $\mathcal{O}(\varepsilon^q)$ terms, thereby introducing an error of $\mathcal{O}(\varepsilon^{q+1})$.

In what follows, we will at first reformulate the CSP projection of g and the amplitude dynamics using a matrix notation which is coherent with the CSP literature. We will then introduce and illustrate the CSP basis vectors refinement procedure.

We will assume the total dimension of system (1.57) to be N (rather than $m + n$ of the previous sections) so that the composition vector $x \in R^N$ and the vector field $g \in R^N$. The latter, as it appears in (1.57), is represented in terms of the standard basis of the composition space. CSP uses another representation by projecting it on a CSP basis, represented by the $N \times N$ matrix A where

$$A = (\mathbf{a}_1, \dots, \mathbf{a}_N) \quad (1.58)$$

being $\{\mathbf{a}_i\}$ a system of N linearly independent basis column vectors. The resulting system is evidently

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) = A\mathbf{f} = \mathbf{a}_1 f^1 + \cdots + \mathbf{a}_n f^n \quad (1.59)$$

where $\mathbf{f} = (f^1, \dots, f^n)^T$ is the vector of modal amplitudes. Let matrix B be the inverse of A , so that $BA = AB = I$, and

$$B = (\mathbf{b}^1, \dots, \mathbf{b}^n)^T \quad (1.60)$$

being $\{\mathbf{b}^i\}$ a system of N linearly independent basis row vectors such that $\mathbf{b}^i \mathbf{a}_j = \delta_j^i$. The vector of modal amplitudes is given by

$$\mathbf{f} = B\mathbf{g} \quad (1.61)$$

so that each component is $f^i = \mathbf{b}^i \mathbf{g}$. Differentiating the above equation along a trajectory yields the amplitude evolution equation:

$$\frac{d\mathbf{f}}{dt} = \Lambda \mathbf{f} \quad (1.62)$$

where

$$\Lambda = \left(\frac{dB}{dt} + BJ \right) A \quad ; \quad \Lambda_j^i = \left(\frac{d\mathbf{b}^i}{dt} + \mathbf{b}^i \mathbf{J} \right) \mathbf{a}_j \quad (1.63)$$

with J being the Jacobian of \mathbf{g} .

The underlying assumption in CSP is that each mode $\mathbf{a}_i f^i$ is stable, this implying that each amplitude f^i along \mathbf{a}_i decays and eventually vanishes at the equilibrium point. As repeatedly pointed out, matrix Λ is not diagonal at a given point on a trajectory and therefore the amplitude evolution is coupled. This introduces some ambiguity in the definition of a proper local timescale for the evolution of each amplitude, so that fast and slow evolving amplitudes may be safely distinguished. An initial approximation of the local spectrum of intrinsic timescales may be assumed as the inverse of the diagonal entries of Λ , $\tau_i = 1/\Lambda_i^i$, these coinciding with the actual intrinsic timescales only if Λ were diagonal. This approximation on the other hand allows us to perform a classification of modes in terms of the absolute value of the corresponding τ_i . Supposing the amplitudes are sorted from fastest to slowest and the first M are considered fast and the remaining $N - M$ slow⁹, then the timescale spectrum has the following ordering:

$$\underbrace{|\tau_1| < \cdots < |\tau_M|}_{\text{fast timescales}} < \underbrace{|\tau_{M+1}| < \cdots < |\tau_n|}_{\text{slow timescales}} \quad (1.64)$$

⁹Note the change of notation: N is now the total dimension ($n + m$ in the previous sections) M is now the number of fast amplitudes (n in the previous sections) and $N - M$ the number of slow amplitudes (m in the previous sections)

where the spectral gap may be identified by the number $\varepsilon = |\tau_M|/|\tau_{M+1}| < 1$. The amplitudes will be split in two classes, $\mathbf{f} = (\mathbf{f}^r, \mathbf{f}^s)^T$, where $\mathbf{f}^r = (f^1, \dots, f^M)$ is an M -vector of fast amplitudes and $\mathbf{f}^s = (f^{M+1}, \dots, f^n)$ in an $(N - M)$ -vector of slow amplitudes. This splitting induces a splitting of basis A , and therefore a decomposition of the phase space $R^N = E^r \oplus E^s$, into a fast subspace E^r and a slow subspace E^s :

$$A = (A_r, A_s) ; \quad A_r = (\mathbf{a}_1, \dots, \mathbf{a}_M) ; \quad A_s = (\mathbf{a}_{M+1}, \dots, \mathbf{a}_n) \quad (1.65)$$

where A_r is the $N \times M$ matrix of vectors $\{\mathbf{a}_i\}_{i=1}^M$ spanning the fast subspace E^r and A_s is the $N \times (M - N)$ matrix of vectors $\{\mathbf{a}_i\}_{i=M+1}^N$ spanning the slow subspace E^s . The corresponding decomposition of B is:

$$B = (B^r, B^s)^T ; \quad B^r = (\mathbf{b}^1, \dots, \mathbf{b}^M)^T ; \quad B^s = (\mathbf{b}^{M+1}, \dots, \mathbf{b}^N)^T \quad (1.66)$$

where B^r is $M \times N$ and B^s is $(M - N) \times N$. Thus the vectors of fast and slow amplitudes read:

$$\mathbf{f}^r = B^r \mathbf{g} ; \quad \mathbf{f}^s = B^s \mathbf{g}. \quad (1.67)$$

The fast-slow partition¹⁰ allows us to rewrite the system as

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) = A_r \mathbf{f}^r + A_s \mathbf{f}^s = \mathbf{g}^r + \mathbf{g}^s. \quad (1.68)$$

Note that the projection of \mathbf{g} onto the fast subspace is given by $\mathbf{g}^r = \mathbf{g} - \mathbf{g}^s = (I - A_s B^s) \mathbf{g} = A_r B^r \mathbf{g}$, and similarly $\mathbf{g}^s = (I - A_r B^r) \mathbf{g} = A_s B^s \mathbf{g}$. Then $P^r = A_r B^r = (I - A_s B^s)$ is a projection operator $P^r : R^n \rightarrow E^r$ and similarly $P^s = A_s B^s = (I - A_r B^r)$ such that $P^s : R^n \rightarrow E^s$.

Similarly the amplitude evolution along a trajectory reads as

$$\begin{pmatrix} \dot{\mathbf{f}}^r \\ \dot{\mathbf{f}}^s \end{pmatrix} = \begin{pmatrix} \Lambda_r^r & \Lambda_s^r \\ \Lambda_r^s & \Lambda_s^s \end{pmatrix} \begin{pmatrix} \mathbf{f}^r \\ \mathbf{f}^s \end{pmatrix}. \quad (1.69)$$

The fast-slow mode mixing $M \times (M - N)$ block Λ_s^r and the slow-fast mode mixing $(M - N) \times M$ block Λ_r^s are generally non zero (and in principle of $\mathcal{O}(1)$) for an arbitrary initial choice of the initial

¹⁰Note that the orthonormality condition $AB=BA=I$, when the partition is applied, yields the following set of relations: $A_r B^r + A_s B^s = I$, $B^r A_r = I_r^r$, $B^s A_s = I_s^s$, $B^r A_s = 0_s^r$ and $B^s A_r = 0_r^s$.

basis A and the resulting B . The role of each CSP refinement is to modify A and B so that the off diagonal mode mixing blocks are reduced by $\mathcal{O}(\varepsilon)$ so that after q refinements $\Lambda_s^r = \mathcal{O}(\varepsilon^q)$ and $\Lambda_r^s = \mathcal{O}(\varepsilon^q)$. After each refinement we can impose the asymptotic condition

$$\mathbf{f}^r = B^r \mathbf{g} = 0 \quad (1.70)$$

which identifies the locus of points \mathcal{K} of the phase space where the fast amplitudes vanish when evaluated with respect to the current basis. The system's manifold \mathcal{M}_ε is indeed the locus of points where the fast amplitudes vanish when evaluated with respect to the ideal basis (for which $\Lambda_s^r = 0$ and $\Lambda_r^s = 0$). After q refinements the asymptotic condition (1.70) defining the CSP manifold \mathcal{K} will yield an approximation of the manifold which agrees with \mathcal{M}_ε up to and including terms of $\mathcal{O}(\varepsilon^q)$ (a rigorous proof may be found in [19]). At the same time, by virtue of condition (1.70), the system will take the simplified form:

$$\dot{\mathbf{x}} = A_s \mathbf{f}^s = (I - A_r B^r) \mathbf{g}, \quad (1.71)$$

which describes the evolution of the system on the CSP manifold \mathcal{K} . If the refinements have sufficiently reduced the coupling of the slow amplitudes from the fast, then such evolution is non-stiff, being governed merely by the set of slow timescales.

1.8.1 The CSP refinement procedure

In this section we will introduce the CSP refinement procedure. Such a procedure was developed by Lam and Goussis and it constitutes an extension of the iterative power method to compute the dominant eigenvalue/vector of a square matrix (where the least dominant are determined through deflation). The Lam-Goussis procedure is rather a 'block-power method' in that successive iterations reduce the matrix to a block-diagonal form. When considering the modal amplitude evolution operator Λ of Eq. (1.69), the technique can be termed 'refinement procedure of the CSP bases A and B '. Appendix A contains a complete illustration of the block-power method applied to a constant matrix.

A CSP refinement is composed of two distinct steps. In the first step, termed ' B^r -refinement', two refinement operators act on A and B so as to alter B^r (and A_s to preserve orthonormality) while leaving B^s and A_r unaltered. The effect of the first step is to reduce the upper right block Λ_s^r and hence the coupling of the fast amplitudes from the slow. This will result in a more accurate CSP manifold \mathcal{K} (see Eq. (1.70)), which will coincide with \mathcal{M}_ε after an infinite number of iterations.

In the second step, termed ' A_r -refinement', two further operators act on A and B so as to alter A_r

(and B^s to preserve orthonormality) while leaving A_s and B^r unaltered. This step will reduce the lower right block Λ_r^s and hence the coupling of the slow amplitudes from the fast. The result of this is an increasingly less stiff simplified problem (see Eq. (1.71)) as the iterations increase. Both steps preserve the orthonormality of the CSP basis vectors thus allowing the vector field projection after every step.

In what follows, each term will be accompanied by a pair of integer indices (k, m) where k will indicate the number of B^r -refinements (first step) and m the number of A_r -refinements (second step), as the two steps may not necessarily be performed sequentially.

The B^r -refinement is constituted by the following operations on A and B :

$$(A_r(k+1, m), A_s(k+1, m)) = (A_r(k, m), A_s(k, m)) \begin{pmatrix} I_r^r & -(\Lambda_r^r(k, m))^{-1} \Lambda_s^r(k, m) \\ 0_r^s & I_s^s \end{pmatrix} \quad (1.72)$$

$$\begin{pmatrix} B^r(k+1, m) \\ B^s(k+1, m) \end{pmatrix} = \begin{pmatrix} I_r^r & (\Lambda_r^r(k, m))^{-1} \Lambda_s^r(k, m) \\ 0_r^s & I_s^s \end{pmatrix} \begin{pmatrix} B^r(k, m) \\ B^s(k, m) \end{pmatrix}. \quad (1.73)$$

The A^r -refinement is constituted by the following operations on A and B :

$$(A_r(k, m+1), A_s(k, m+1)) = (A_r(k, m), A_s(k, m)) \begin{pmatrix} I_r^r & 0_s^r \\ \Lambda_r^s(k, m)(\Lambda_r^r(k, m))^{-1} & I_s^s \end{pmatrix} \quad (1.74)$$

$$\begin{pmatrix} B^r(k, m+1) \\ B^s(k, m+1) \end{pmatrix} = \begin{pmatrix} I_r^r & 0_s^r \\ -\Lambda_r^s(k, m)(\Lambda_r^r(k, m))^{-1} & I_s^s \end{pmatrix} \begin{pmatrix} B^r(k, m) \\ B^s(k, m) \end{pmatrix}. \quad (1.75)$$

By defining the two operators

$$U = \begin{pmatrix} 0_r^r & (\Lambda_r^r)^{-1} \Lambda_s^r \\ 0_r^s & 0_s^s \end{pmatrix} \quad ; \quad L = \begin{pmatrix} 0_r^r & 0_s^r \\ \Lambda_r^s(\Lambda_r^r)^{-1} & 0_s^s \end{pmatrix}, \quad (1.76)$$

then the B^r -refinement takes the form:

$$A(k+1, m) = A(k, m)(I - U(k, m)) \quad (1.77)$$

$$B(k+1, m) = (I + U(k, m))B(k, m) \quad (1.78)$$

and the A_r -refinement takes the form:

$$A(k, m+1) = A(k, m)(I + L(k, m)) \quad (1.79)$$

$$B(k, m+1) = (I - L(k, m))B(k, m). \quad (1.80)$$

The above refinement procedure stems, as mentioned, from the application of the block-power method aimed at the block-diagonalization of the amplitude evolution operator Λ . Following the notation of Appendix A this means applying the method to a matrix $C(k, m)$ such that

$$\begin{aligned} \Lambda(k, m) &= \left(\frac{dB(k, m)}{dt} + B(k, m)J \right) A(k, m) = \\ &= \begin{pmatrix} \Lambda_r^r(k, m) & \Lambda_s^r(k, m) \\ \Lambda_r^s(k, m) & \Lambda_s^s(k, m) \end{pmatrix} = B(k, m)C(k, m)A(k, m), \end{aligned} \quad (1.81)$$

where, matrix $C(k, m)$ is therefore

$$C(k, m) = A(k, m)\Lambda(k, m)B(k, m) = A(k, m) \left(\frac{dB(k, m)}{dt} + B(k, m)J \right). \quad (1.82)$$

By applying Eqs. (A.44)-(A.45) from Appendix A, the B^r -refinement becomes

$$\begin{aligned} B^r(k+1, m) &= (\Lambda_r^r(k, m))^{-1} B^r(k, m) C(k, m) = \\ &= (\Lambda_r^r(k, m))^{-1} \left(\frac{dB^r(k, m)}{dt} + B^r(k, m)J \right) \end{aligned} \quad (1.83)$$

$$A_s(k+1, m) = (I_r^r - A_r(k+1, m)B^r(k+1, m)) A_s(k, m) \quad (1.84)$$

$$B^s(k+1, m) = B^s(k, m) \quad (1.85)$$

$$A_r(k+1, m) = A_r(k, m) \quad (1.86)$$

where, in the first of the above, we substituted $B^r C = \Lambda_r^r B^r + \Lambda_s^r B^s = (dB^r/dt + B^r J)(A_r B^r + A_s B^s) = (dB^r/dt + B^r J)$.

Similarly by applying Eqs. (A.16)-(A.17) the A_r -refinement becomes

$$\begin{aligned} A_r(k, m+1) &= C(k, m)A_r(k, m)(\Lambda_r^r(k, m))^{-1} = \\ &= \left(-\frac{dA_r(k, m)}{dt} + JA_r(k, m) \right) (\Lambda_r^r(k, m))^{-1} \end{aligned} \quad (1.87)$$

$$B^s(k, m+1) = B^s(k, m) (I_r^r - A_r(k, m+1)B^r(k, m+1)) \quad (1.88)$$

$$A_s(k, m+1) = A_s(k, m) \quad (1.89)$$

$$B^r(k, m+1) = B^r(k, m). \quad (1.90)$$

The first of the above is obtained after some manipulation by substituting $CA_r = A_r\Lambda_r^r + A_s\Lambda_r^s = A_r(dB^r/dt)A_r + A_s(dB^s/dt)A_r + JA_r$ and using the relations $(dB^s/dt)A_r = -B^s(dA_r/dt)$ and $(dB^r/dt)A_r = -B^r(dA_r/dt)$ which are obtained from the orthonormality condition $AB = BA = I$.

It is interesting to note that if the basis vectors time derivatives in the above expressions are ignored, then, from Eq. (1.82), matrix $C(k, m)$ reduces to $C(k, m) = J$. This means that the refinement procedure reduces to a block-diagonalization of the local Jacobian $J(\mathbf{x})$ of the system's vector field $\mathbf{g}(\mathbf{x})$ resulting, as seen in Appendix A, in $A^r(k, m)$ spanning, at \mathbf{x} , the subspace spanned by the right eigenvectors of $J(\mathbf{x})$ possessing the greatest eigenvalues and $A_s(k, m)$ spanning the complementary subspace. As seen in section 1.7, this situation is equivalent to a linearization of the system and coincides with the ILDM approach which ultimately introduces an endemic error of $\mathcal{O}(\varepsilon^2)$ in the expression of the slow invariant manifold and proportional to its curvature.

A substantial difference between the refinement procedure and the block-power method described in Appendix A is the fact that matrix $C(k, m)$ is not constant but rather changes as the iteration proceeds being, from Eq. (1.82), a function of $A(k, m)$ and $B(k, m)$. As a result of this, matrix Λ is not transformed into a similar matrix after each iteration. This can be seen using Eq. (1.81) and Eqs. (1.77)-(1.80) from which we obtain:

$$\begin{aligned} \Lambda(k+1, m) &= (I + U(k, m))\Lambda(k, m)(I - U(k, m)) + \frac{dU(k, m)}{dt}(I - U(k, m)) \\ \Lambda(k, m+1) &= (I - L(k, m))\Lambda(k, m)(I + L(k, m)) + \frac{dL(k, m)}{dt}(I + L(k, m)) \end{aligned}$$

This may cause confusion as to the capability of the refinements to block diagonalise matrix $\Lambda(k, m)$. However, supposing $m = k$, it can be shown that the terms $\frac{dU(k, m)}{dt}$ and $\frac{dL(k, m)}{dt}$ are $\mathcal{O}(\varepsilon^{k+1})$ (or $\mathcal{O}(\varepsilon^{m+1})$) and hence do not compromise the off-diagonal blocks of $\Lambda(k, m)$ which are $\mathcal{O}(\varepsilon^k)$ (or $\mathcal{O}(\varepsilon^m)$). A rigorous proof can be found in [19].

The presence of time derivatives in the above expressions and indeed in the refinements Eqs. (1.83)-(1.90), underlines the fact that the CSP refinement procedure is capable of capturing the nonlinearity of the original system which manifests itself in higher order curvature terms in the expression of its slow invariant manifold. Because $A = A(\mathbf{x})$ and $B = B(\mathbf{x})$, A and B should be regarded as matrix-valued functions of phase space points and thus assume the value of fields (clearly $\{\mathbf{a}_i(\mathbf{x})\}$ and $\{\mathbf{b}^i(\mathbf{x})\}$ are vector fields). Their time derivatives are fields in their own right because they can be written as

$$\frac{dA}{dt} = \sum_{i=1}^n \frac{\partial A}{\partial x_i} \frac{dx_i}{dt} = \left(\frac{\partial A}{\partial x_1}, \dots, \frac{\partial A}{\partial x_n} \right) \mathbf{g} = (DA)\mathbf{g} \quad (1.91)$$

$$\frac{dB}{dt} = \sum_{i=1}^n \frac{\partial B}{\partial x_i} \frac{dx_i}{dt} = \left(\frac{\partial B}{\partial x_1}, \dots, \frac{\partial B}{\partial x_n} \right) \mathbf{g} = (DB)\mathbf{g} \quad (1.92)$$

where $\mathbf{g}(\mathbf{x})$ is the system's autonomous vector field.

When solving low-dimensional problems in the idealized singularly perturbed form (as are the Davis-Skodje or Michaelis-Menten systems), the computation of derivatives does not pose significant problems. The above relations can be used for any number of CSP refinements to evaluate analytical expressions for the derivatives, with the help, maybe of a symbolic manipulator. For actual chemical kinetic problems however time derivatives can only be evaluated for a limited number of refinements as discussed in the next sections.

1.8.2 CSP refinement: general implementation

In the following sections we will give examples of a fully analytical implementation of the CSP refinements Eqs. (1.83)-(1.90). These examples however are limited to simple low dimensional model problems for which time derivatives of higher order can realistically be sought. For higher dimensional problems, such as chemical kinetics problems, the evaluation of derivatives becomes prohibitive due to the complexity of the expressions to be used and the unavailability of constituent terms such as the higher order derivatives of the Hessian of $\mathbf{g}(\mathbf{x})$. Because of this, a maximum of two refinements is ultimately feasible. The implementation discussed here is composed of two distinct phases: (i) a first refinement, wherein the time derivatives of the basis vectors are neglected, yielding leading order accuracy (for the CSP manifold) and (ii) a second refinement, wherein the time derivatives are included, providing higher order accuracy.

Phase (i) is equivalent to one block-power refinement of the Jacobian J and, if carried out many times, the M fast and $N-M$ slow basis vectors will span the M -dimensional fast and $(N-M)$ -slow

eigenspaces of J , respectively. In essence this step provides a way of establishing the fast and slow eigenspaces of J starting from an arbitrary set of basis vectors. Clearly if the eigenvectors of J are chosen as initial basis vectors then Phase (i) is redundant. If derivatives are not included, however, the ensuing CSP manifold will always introduce an $\mathcal{O}(\varepsilon^2)$ error. Phase (ii), on the other hand, does include derivatives and can practically be carried out just once, due to the increasing cost of obtaining time derivatives of the basis vectors for additional to the first refinements.

Each of the two phases is composed of a B^r -refinement (identified by index k) and an A_r -refinement (identified by index m) respectively. Formulas for the evaluation of $B^r(k, m)$ and $A_r(k, m)$ for $k = 1, 2$ and $m = 1, 2$ will be presented together with expressions for the necessary basis vector time derivatives.

Phase (i)

We start from the iteration counters $k = m = 0$ and a set of arbitrary initial bases $B(0, 0)$ and $A(0, 0)$, already suitably partitioned into fast and slow vectors, and for which we enforce:

$$\frac{dB^r(0, 0)}{dt} = 0 \quad ; \quad \frac{dA_r(0, 0)}{dt} = 0 \quad (1.93)$$

B^r -refinement:

Under the assumptions (1.93), the B^r -refinement yields:

$$\begin{aligned} \Lambda_r^r(0, 0) &= \left(\frac{dB^r(0, 0)}{dt} + B^r(0, 0)J \right) A_r(0, 0) = B^r(0, 0)JA_r(0, 0) \\ \tau_r^r(0, 0) &= (\Lambda_r^r(0, 0))^{-1} \end{aligned}$$

$$\begin{aligned} B^r(1, 0) &= \tau_r^r(0, 0) \left(\frac{dB^r(0, 0)}{dt} + B^r(0, 0)J \right) = \tau_r^r(0, 0)B^r(0, 0)J \\ A_r(1, 0) &= A_r(0, 0) \\ B^s(1, 0) &= B^s(0, 0) \\ A_s(1, 0) &= (I - A_r(1, 0)B^r(1, 0)) A_s(0, 0) = (I - A_r(0, 0)B^r(1, 0)) A_s(0, 0) \end{aligned}$$

The end effect of this will be to lower the norm of the upper-right off-diagonal block:

$$\Lambda_s^r(1, 0) = B^r(1, 0)JA_s(1, 0) = O(\varepsilon\Lambda_s^r(0, 0))$$

by an order $\mathcal{O}(\varepsilon)$ (where $\varepsilon = |\tau_M|/|\tau M + 1|$ with $\tau_i = 1/\Lambda_i^i(0, 0)$), thereby making the fast modes

”purer” by decoupling them from the slow modes, while the lower left off-diagonal block is left unchanged:

$$\Lambda_r^s(1, 0) = B^s(0, 0)JA_r(0, 0) = \Lambda_r^s(0, 0)$$

A_r -refinement:

The A_r -refinement yields:

$$\begin{aligned}\Lambda_r^r(1, 0) &= \left(\frac{dB^r(1, 0)}{dt} + B^r(1, 0)J \right) A_r(1, 0) = B^r(1, 0)JA_r(0, 0) \\ \tau_r^r(1, 0) &= (\Lambda_r^r(1, 0))^{-1}\end{aligned}$$

$$\begin{aligned}A_r(1, 1) &= \left(-\frac{dA_r(1, 0)}{dt} + JA_r(1, 0) \right) \tau_r^r(1, 0) = JA_r(0, 0)\tau_r^r(1, 0) \\ B^r(1, 1) &= B^r(1, 0) \\ A_s(1, 1) &= A_s(1, 0) \\ B^s(1, 1) &= B^s(1, 0)(I - A_r(1, 1)B^r(1, 1)) = B^s(0, 0)(I - A_r(1, 1)B^r(1, 0))\end{aligned}$$

The end effect of this will be to lower the norm of the lower left off-diagonal block:

$$\Lambda_r^s(1, 1) = \left(\frac{dB^s(1, 1)}{dt} + B^s(1, 1)J \right) A_r(1, 1) = O(\varepsilon \Lambda_r^s(1, 0)) = O(\varepsilon \Lambda_r^s(0, 0))$$

by an $\mathcal{O}(\varepsilon)$ (with $\varepsilon = |\tau_M/\tau_{M+1}| < 1$, $\tau_i = 1/\Lambda_i^s(1, 0)$), thereby making the slow modes ”purer” by decoupling them from the fast modes, while the upper right off-diagonal block is left unchanged:

$$\begin{aligned}\Lambda_s^r(1, 1) &= \left(\frac{dB^r(1, 1)}{dt} + B^r(1, 1)J \right) A_s(1, 1) = \left(\frac{dB^r(1, 0)}{dt} + B^r(1, 0)J \right) A_s(1, 0) = \\ &= \Lambda_s^r(1, 0) = O(\varepsilon \Lambda_s^r(0, 0))\end{aligned}$$

Phase (ii)

In the refinements of Phase (ii) the basis vector time derivatives are calculated. The initial basis may be chosen as the basis resulting from Phase (i) or as the eigenvectors of J themselves (in which case Phase (i) is skipped).

B^r -refinement:

$$\begin{aligned}\Lambda_r^r(1,1) &= \left(\frac{dB^r(1,1)}{dt} + B^r(1,1)J \right) A_r(1,1) \\ \tau_r^r(1,1) &= (\Lambda_r^r(1,1))^{-1}\end{aligned}$$

$$B^r(2,1) = \tau_r^r(1,1) \left(\frac{dB^r(1,1)}{dt} + B^r(1,1)J \right) = \tau_r^r(1,1) \left(\frac{dB^r(1,0)}{dt} + B^r(1,0)J \right)$$

$$A_r(2,1) = A_r(1,1)$$

$$B^s(2,1) = B^s(1,1)$$

$$A_s(2,1) = (I - A_r(2,1)B^r(2,1)) A_s(1,1)$$

The end effect of this will be to lower the norm of the upper-right off-diagonal block:

$$\begin{aligned}\Lambda_s^r(2,1) &= \left(\frac{dB^r(2,1)}{dt} + B^r(2,1)J \right) A_s(2,1) = O(\varepsilon \Lambda_s^r(1,1)) = \\ &= O(\varepsilon \Lambda_s^r(1,0)) = O(\varepsilon^2 \Lambda_s^r(0,0))\end{aligned}$$

by an order $\varepsilon = |\tau_M/\tau_{M+1}| < 1$, thereby making the fast modes "purer" by decoupling them from the slow modes, while the lower left off-diagonal block is left unchanged:

$$\begin{aligned}\Lambda_r^s(2,1) &= \left(\frac{dB^s(2,1)}{dt} + B^s(2,1)J \right) A_r(2,1) = \left(\frac{dB^s(1,1)}{dt} + B^s(1,1)J \right) A_r(1,1) = \\ &= \Lambda_r^s(1,1) = O(\varepsilon \Lambda_r^s(1,0)) = O(\varepsilon \Lambda_r^s(0,0))\end{aligned}$$

A_r -refinement:

The A_r -refinement yields:

$$\begin{aligned}\Lambda_r^r(2,1) &= \left(\frac{dB^r(2,1)}{dt} + B^r(2,1)J \right) A_r(2,1) \\ \tau_r^r(2,1) &= (\Lambda_r^r(2,1))^{-1}\end{aligned}$$

$$A_r(2,2) = \left(-\frac{dA_r(2,1)}{dt} + JA_r(2,1) \right) \tau_r^r(2,1) = \left(-\frac{dA_r(1,1)}{dt} + JA_r(1,1) \right) \tau_r^r(2,1)$$

$$B^r(2,2) = B^r(2,1)$$

$$A_s(2,2) = A_s(2,1)$$

$$B^s(2,2) = B^s(2,1) (I - A_r(2,2)B^r(2,2)) = B^s(2,1) (I - A_r(2,2)B^r(2,1))$$

The end effect of this will be to lower the norm of the lower-left off-diagonal block:

$$\begin{aligned}\Lambda_r^s(2,2) &= \left(\frac{dB^s(2,2)}{dt} + B^s(2,2)J \right) A_r(2,2) = O(\varepsilon \Lambda_r^s(2,1)) = \\ &= O(\varepsilon^2 \Lambda_r^s(1,0)) = O(\varepsilon^2 \Lambda_r^s(0,0))\end{aligned}$$

by an order $\varepsilon = |\tau_M/\tau_{M+1}| < 1$, thereby making the slow modes "purer" by decoupling them from the fast modes, while the upper right off-diagonal block is left unchanged:

$$\begin{aligned}\Lambda_s^r(2,2) &= \left(\frac{dB^r(2,2)}{dt} + B^r(2,2)J \right) A_s(2,2) = \left(\frac{dB^r(2,1)}{dt} + B^r(2,1)J \right) A_s(2,1) = \\ &= \Lambda_s^r(2,1) = O(\varepsilon \Lambda_s^r(1,1)) = O(\varepsilon \Lambda_s^r(1,0)) = O(\varepsilon^2 \Lambda_s^r(0,0)).\end{aligned}$$

Note that in Phase (ii) we have assumed ε constant for simplicity. In fact it is not, being, before every refinement, equal to $\varepsilon(k, m) = |\tau_M(k, m)|/|\tau_{M+1}(k, m)|$ where $\tau_{M+1}(k, m) = \Lambda_i^i(k, m)$.

1.8.3 Basis Vectors Time Derivatives

The expressions for the time derivatives occurring in Phase (ii) are given here below. Their full derivation is given in Appendix B.

$$\frac{dB^r(1,1)}{dt} = \frac{dB^r(1,0)}{dt} = \tau_r^r(0,0)B^r(0,0)\frac{dJ}{dt}[I - A_r(0,0)B^r(1,0)] \quad (1.94)$$

$$\begin{aligned}\frac{dA_r(2,1)}{dt} &= \frac{dA_r(1,1)}{dt} = \\ &= [I - A_r(1,1)B^r(1,0)]\frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) + \\ &\quad - A_r(1,1)\frac{dB^r(1,0)}{dt}A_r(1,1) \quad (1.95)\end{aligned}$$

$$\begin{aligned}\frac{dB^r(2,1)}{dt} &= \tau_r^r(1,1)\left[\frac{dB^r(1,0)}{dt}J + B^r(1,0)\frac{dJ}{dt} + \frac{d^2B^r(1,0)}{dt^2}\right] \times \\ &\quad \times [I - A_r(1,1)B^r(2,1)] - B^r(2,1)\frac{dA_r(1,1)}{dt}B^r(2,1) \quad (1.96)\end{aligned}$$

where

$$\begin{aligned} \frac{d^2 B^r(1,0)}{dt^2} = & \left[\frac{d\tau_r^r(0,0)}{dt} B^r(0,0) \frac{dJ}{dt} + \tau_r^r(0,0) B^r(0,0) \frac{d^2 J}{dt^2} \right] [I - A_r(0,0) B^r(1,0)] + \\ & - \tau_r^r(0,0) B^r(0,0) \frac{dJ}{dt} A_r(0,0) \frac{dB^r(1,0)}{dt} \end{aligned} \quad (1.97)$$

Note that formula (1.97) involves dJ/dt and $d^2 J/dt^2$, while formulas (1.94), (1.95) and (1.96) involve only dJ/dt . This further implies that in Phase (ii) the B^r -refinement requires the availability of dJ/dt , while the A_r -refinement requires the availability of both dJ/dt and $d^2 J/dt^2$. These quantities can be evaluated from the expressions:

$$\frac{dJ}{dt} = \sum_{i=1,n} \frac{\partial J}{\partial y^i} \frac{dy^i}{dt} = \sum_{i=1,N_s} \frac{\partial J}{\partial y^i} g^i \quad (1.98)$$

$$\frac{d^2 J}{dt^2} = \sum_{i,j=1,n} \left[\frac{\partial^2 J}{\partial y^i \partial y^j} g^i + \frac{\partial J}{\partial y^i} J_j^i \right] g^j \quad (1.99)$$

which show that dJ/dt and $d^2 J/dt^2$ depend on \mathbf{x} only (i.e the state of the system). Such derivatives involve the terms $\partial J/\partial y^i$ (Hessian) and $\partial^2 J/\partial y^i \partial y^j$ which represent higher order terms in the local expansion of the system's vector field and hence capture its higher order nonlinearities.

Because all the quantities involved in a refinement, including derivative terms, are fundamentally a function of the local state vector \mathbf{x} , it is possible to imagine a tabulation of basis vectors in a physically meaningful domain of the phase space itself¹¹. Such a task could be carried out prior and independently of the system's integration which in turn would greatly benefit from the pointwise knowledge of a suitable projection basis.

1.8.4 CSP application examples

In this section we will illustrate the effect of CSP refinements on a fully analytical example first, the planar Davis-Skodje system, and then on a semi-analytical three dimensional example, a prototypical three species chemical kinetics system.

The Davis-Skodje system

The Davis-Skodje system was extensively used in the preceding sections to illustrate the Fraser-Roussel method in Sect.1.5 and the ILDM method in Sect.1.6. Thus it will prove a good testing

¹¹This technique is used in the ILDM method.

ground for the CSP method as well. Consistently with the fact that CSP notation places fast components first and slow second (both in basis vectors and amplitude vectors), we rewrite the system in the ILDM form Eq. (1.36) with the fast variable first and the slow second ($\mathbf{x} = (y, z)^T$):

$$\begin{aligned}\frac{dy}{dt} &= \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2} \\ \frac{dz}{dt} &= -z\end{aligned}\tag{1.100}$$

Following the considerations of Sect.1.7, the system in this form has the following set of ideal basis vectors

$$\begin{aligned}\mathbf{a}_1 &= (1, 0)^T & ; \quad \mathbf{a}_2 &= ((1+z)^{-2}, 1)^T \\ \mathbf{b}^1 &= (1, -(1+z)^{-2}) & ; \quad \mathbf{b}^2 &= (0, 1)\end{aligned}\tag{1.101}$$

for which the resulting matrix Λ attains the diagonal form

$$\Lambda = \begin{pmatrix} -\frac{1}{\varepsilon} & 0 \\ 0 & -1 \end{pmatrix}.\tag{1.102}$$

In this section will at first project the system's vector field on an arbitrary basis and examine the effects of A_r -refinements and B_r -refinements when time derivatives are included. The CSP basis in this case will converge to the ideal basis seen in Sect.1.7. Secondly we will examine the effect of neglecting the time derivatives on the accuracy of the CSP manifold. We will see that CSP refinements without derivatives will yield a manifold identical to the ILDM, whereas by including the derivatives the CSP manifold will converge to the exact expression of the SIM.

Consider the initial, i.e. $(k, m) = (0, 0)$, arbitrary basis vectors:

$$\mathbf{a}_1(0, 0) = (1, 1)^T ; \quad \mathbf{a}_2(0, 0) = (0, -1)^T\tag{1.103}$$

$$\mathbf{b}^1(0, 0) = (1, 0) ; \quad \mathbf{b}^2(0, 0) = (1, -1)\tag{1.104}$$

The resulting matrix $\Lambda(0, 0)$ reads

$$\Lambda(0, 0) = \begin{pmatrix} -\frac{1}{\varepsilon}\Psi & -\frac{1}{\varepsilon}F_z \\ -\frac{1}{\varepsilon}\Psi + 1 & -\frac{1}{\varepsilon}F_z - 1 \end{pmatrix}\tag{1.105}$$

where

$$F_z = \frac{1}{(1+z)^2} - \varepsilon \frac{1-z}{(1+z)^3} ; \quad \Psi = 1 - F_z. \quad (1.106)$$

The off-diagonal blocks Λ_2^1 and Λ_1^2 (which in this planar case are simply scalars) are both $\mathcal{O}(\varepsilon^{-1})$, the same order of the diagonal blocks. This clearly implies severe mode coupling which makes it impossible to discriminate between fast and slow timescales. The modal amplitudes read

$$f^1(0,0) = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2} \quad (1.107)$$

$$f^2(0,0) = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2} + z \quad (1.108)$$

The CSP manifold $\mathcal{K}(0,0)$ (although this terminology is unappropriate as no refinements were yet carried out) is obtained as the locus of points where the fast amplitudes vanish. By considering this locus as, indifferently, $f^1(0,0) = 0$ or $f^2(0,0) = 0$ one obtains

$$y = \psi(z) = \psi_0(z) + \varepsilon \psi_1(z) + \varepsilon^2 \psi_2(z) + \dots = \frac{z}{1+z} + \mathcal{O}(\varepsilon) \quad (1.109)$$

indicating an $\mathcal{O}(\varepsilon)$ error on the Davis-Skodje manifold $\mathcal{M}_\varepsilon = \mathcal{M}_0$ which is $y = z/(1+z)$. Note that the CSP simplified form for the manifold-constrained system would read: $\dot{\mathbf{x}} = f^2 \mathbf{a}_2$, i.e.

$$\frac{d}{dt} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} \left(\frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2} + z \right). \quad (1.110)$$

The eigenvalues of the Jacobian of the vector field for this system are 0 and $\mathcal{O}(\varepsilon^{-1})$, indicating that it is still dominated by fast timescales.

Let us now examine the effect of CSP refinements.

Effect of A_r -refinements

Performing one A_r -refinement as described in Sect.1.8.1 yields:

$$\mathbf{a}_1(0,1) = (1, \varepsilon \Psi^{-1})^T ; \quad \mathbf{a}_2(0,1) = (0, -1)^T \quad (1.111)$$

$$\mathbf{b}^1(0,1) = (1, 0) ; \quad \mathbf{b}^2(0,1) = (-\varepsilon \Psi^{-1}, 1) \quad (1.112)$$

where time derivatives were calculated analytically. Note that only \mathbf{a}_1 and \mathbf{b}^2 were affected. The

resulting matrix $\Lambda(0, 1)$ reads

$$\Lambda(0, 1) = \begin{pmatrix} -\frac{1}{\varepsilon} \left(1 - \frac{\varepsilon F_z}{\Psi}\right) & \frac{1}{\varepsilon} F_z \\ \frac{1}{\Psi} - \varepsilon \frac{1}{\Psi^2 + G_z} & -\frac{1}{\Psi} \end{pmatrix} \quad (1.113)$$

where

$$G_z = \frac{(2 - \varepsilon)z}{(1+z)^3} - \varepsilon \frac{3z(1-z)}{(1+z)^4}. \quad (1.114)$$

We readily note that the bottom left block, which determines the extent of the coupling of the slow mode with the fast, is now $\Lambda_1^2(0, 1) = \mathcal{O}(1)$, i.e. $\Lambda_1^2(0, 1) = \mathcal{O}(\varepsilon \Lambda_1^2(0, 0))$. On the other hand the term $\Lambda_2^1(0, 1)$ is not reduced. Moreover the modal amplitudes read

$$f^1(0, 1) = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2} \quad (1.115)$$

$$f^2(0, 1) = \frac{1}{\Psi} \left[\left(-y + \frac{z}{1+z} \right) - \frac{\varepsilon z}{(1+z)^2} \right] - z \quad (1.116)$$

Thus $f^1(0, 1) = f^1(0, 0)$ and is therefore unchanged (after all $f^1 = \mathbf{b}^1 \mathbf{g}$ and \mathbf{b}^1 is unchanged) this implying that there is no improvement in the CSP manifold: $\mathcal{K}(0, 1) = \mathcal{K}(0, 1)$. Table 1.3 displays the results of further A_r -refinements. Note that for an infinite number of refinements $\mathbf{a}_1(0, \infty)$ tends to its ideal form, the off-diagonal block $\Lambda_1^2(0, \infty) = 0$ and the diagonal values tend to the ideal values.

(k, m)	$\mathbf{a}_1(k, m)$	Λ_1^1	Λ_2^1	Λ_1^2
$(0, 0)$	$(1, 1)^T$	$-\frac{1}{\varepsilon} + \mathcal{O}(\varepsilon^{-1})$	$-1 + \mathcal{O}(\varepsilon^{-1})$	$\mathcal{O}(\varepsilon^{-1})$
$(0, 1)$	$(1, \mathcal{O}(\varepsilon))^T$	$-\frac{1}{\varepsilon} + \mathcal{O}(1)$	$-1 + \mathcal{O}(1)$	$\mathcal{O}(1)$
$(0, 2)$	$(1, \mathcal{O}(\varepsilon^2))^T$	$-\frac{1}{\varepsilon} + \mathcal{O}(\varepsilon)$	$-1 + \mathcal{O}(\varepsilon)$	$\mathcal{O}(\varepsilon)$
\vdots	\vdots	\vdots	\vdots	\vdots
$(0, \infty)$	$(1, 0)^T$	$-\frac{1}{\varepsilon}$	-1	0

Table 1.3: Effect of A_r -refinements starting from basis (1.103)-(1.104).

The simplified system reads:

$$\frac{d}{dt} \begin{pmatrix} y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \left(\frac{1}{\Psi} \left(\left(-y + \frac{z}{1+z} \right) - \frac{\varepsilon z}{(1+z)^2} \right) - z \right). \quad (1.117)$$

The eigenvalues of J are this time 0 and $\mathcal{O}(1)$ which indicates that the system is slow (non stiff) on the manifold whilst still yielding an $\mathcal{O}(\varepsilon)$ error due to the manifold error.

Effect of B^r -refinements

We now return to our initial arbitrary basis Eqs. (1.103)-(1.104). Performing a B^r -refinement yields:

$$\mathbf{a}_1(1, 0) = \Psi(1, 1)^T ; \quad \mathbf{a}_2(1, 0) = -(\Psi)^{-1}(F_z, 1)^T \quad (1.118)$$

$$\mathbf{b}^1(1, 0) = (1, -F_z) ; \quad \mathbf{b}^2(1, 0) = (1, -1) \quad (1.119)$$

The resulting matrix $\Lambda(1, 0)$ reads

$$\Lambda(1, 0) = \begin{pmatrix} -\frac{1}{\varepsilon} \left(1 - \frac{\varepsilon F_z}{\Psi}\right) & -\frac{F_z - G_z}{\Psi^2} \\ 1 - \frac{\Psi}{\varepsilon} & -\frac{1}{\Psi} \end{pmatrix}. \quad (1.120)$$

Thus the upper-right fast to slow mode coupling block has decreased by an order: $\Lambda_2^1(1, 0) = \mathcal{O}(\varepsilon \Lambda_2^1(0, 0))$. The modal amplitudws are

$$f^1(1, 0) = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{\varepsilon z (1-z)}{(1+z)^3} \quad (1.121)$$

$$f^2(1, 0) = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2} + z \quad (1.122)$$

The CSP manifold $\mathcal{K}(1, 0)$ is the the locus $f^1(1, 0) = 0$ and is therefore

$$y = \psi(z) = \psi_0(z) + \varepsilon \psi_1(z) + \varepsilon^2 \psi_2(z) + \dots = \frac{z}{1+z} + \mathcal{O}(\varepsilon^2) \quad (1.123)$$

underlining an improvement of $\mathcal{O}(\varepsilon)$ compared to $\mathcal{K}(0, 0)$ or $\mathcal{K}(0, 1)$. This is entirely due to the decrease in coupling of the fast mode f^1 from the slow f^2 which makes it a "purer" fast mode.

Table 1.4 displays the results of further B^r -refinements. Note that for an infinite number of refinements $\mathbf{b}^1(\infty, 0)$ tends to its ideal form, the off-diagonal block $\Lambda_2^1(\infty, 0) = 0$ and the diagonal values tend to the ideal values.

The simplified system reads:

$$\frac{d}{dt} \begin{pmatrix} y \\ z \end{pmatrix} = -\Psi^{-1} \begin{pmatrix} F_z \\ 1 \end{pmatrix} \left(\frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2} - z \right). \quad (1.124)$$

(k, m)	$\mathbf{b}^1(k, m)$	Λ_1^1	Λ_2^2	Λ_2^1
$(0, 0)$	$(1, 0)$	$-\frac{1}{\varepsilon} + \mathcal{O}(\varepsilon^{-1})$	$-1 + \mathcal{O}(\varepsilon^{-1})$	$\mathcal{O}(\varepsilon^{-1})$
$(1, 0)$	$(1, -\frac{1}{(1+z)^2} + \mathcal{O}(1))$	$-\frac{1}{\varepsilon} + \mathcal{O}(1)$	$-1 + \mathcal{O}(1)$	$\mathcal{O}(1)$
$(2, 0)$	$(1, -\frac{1}{(1+z)^2} + \mathcal{O}(\varepsilon))$	$-\frac{1}{\varepsilon} + \mathcal{O}(\varepsilon)$	$-1 + \mathcal{O}(\varepsilon)$	$\mathcal{O}(\varepsilon)$
\vdots	\vdots	\vdots	\vdots	\vdots
$(\infty, 0)$	$(1, -\frac{1}{(1+z)^2})$	$-\frac{1}{\varepsilon}$	-1	0

Table 1.4: Effect of B^r -refinements starting from basis (1.103)-(1.104).

The associated eigenvalues are both $\mathcal{O}(\varepsilon^{-1})$, indicating that the B_r -refinement does not decrease the stiffness of the system.

Effect of time derivatives on CSP manifold accuracy

We have seen that only the B^r -refinement improves the accuracy of the CSP manifold. Therefore we now wish to examine the effect of B^r -refinements on the manifold accuracy in the cases where the time derivatives, appearing in the refinement formulae, are neglected or included. This time let us start from yet another initial arbitrary basis, the unit basis:

$$\mathbf{a}_1(0, 0) = (1, 0)^T ; \quad \mathbf{a}_2(0, 0) = (0, 1)^T \quad (1.125)$$

$$\mathbf{b}^1(0, 0) = (1, 0) ; \quad \mathbf{b}^2(0, 0) = (0, 1) \quad (1.126)$$

Tables 1.5,1.6 show the B^r -refinements results when time derivatives are neglected. We note from Table 1.5 that for an infinite number of refinements \mathbf{b}^1 tends to the 'fast' left row eigenvector of J (see Eq. (1.39)) which is the ILDM choice for basis vectors. This shows that when derivatives are neglected, the CSP refinements tend to the fast/slow eigenspaces of J . The same can be observed in Table 1.6 showing the various terms in the asymptotic expansion of the CSP manifold. We note that for an infinite number of refinements the CSP manifold coincides with the ILDM.

(k, m)	$\mathbf{b}^1(k, m)$	Λ_1^1	Λ_2^2	Λ_2^1
$(0, 0)$	$(1, 0)$	$-\frac{1}{\varepsilon}$	-1	$-F_z/\varepsilon$
$(1, 0)$	$(1, -F_z)$	$-\frac{1}{\varepsilon}$	-1	F_z
$(2, 0)$	$(1, -(1 + \varepsilon)F_z)$	$-\frac{1}{\varepsilon}$	-1	εF_z
\vdots	\vdots	\vdots	\vdots	\vdots
$(\infty, 0)$	$(1, -\frac{F_z}{1-\varepsilon})$	$-\frac{1}{\varepsilon}$	-1	0

Table 1.5: Effect of B^r -refinements without time derivatives, starting from basis (1.125)-(1.126).

$\psi^{(k,m)}(z)$	$\mathcal{O}(1)$	$\mathcal{O}(\varepsilon)$	$\mathcal{O}(\varepsilon^2)$	$\mathcal{O}(\varepsilon^3)$	$\mathcal{O}(\varepsilon^4)$
ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	
$\psi^{(0,0)}(z)$	$\frac{z}{1+z}$	$-\frac{z}{(1+z)^2}$	0	0	0
$\psi^{(1,0)}(z)$	$\frac{z}{1+z}$	0	$\frac{(z-1)z}{(1+z)^3}$	0	0
$\psi^{(2,0)}(z)$	$\frac{s}{1+z}$	0	$\frac{2z^2}{(1+z)^3}$	$\frac{(z-1)z}{(1+z)^3}$	0
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$\psi^{(\infty,0)}(z)$	$\frac{z}{1+z}$	0	$\frac{2z^2}{(1+z)^3}$	$\frac{2z^2}{(1+z)^3}$	$\frac{2z^2}{(1+z)^3}$
$\psi_{ILD M}(z)$	$\frac{z}{1+z}$	0	$\frac{2z^2}{(1+z)^3}$	$\frac{2z^2}{(1+z)^3}$	$\frac{2z^2}{(1+z)^3}$

Table 1.6: CSP manifolds obtained from successive B^r -refinements without time derivatives, starting from basis (1.125)-(1.126). The CSP manifold $\mathcal{K}(k, m)$ is expressed in terms of its asymptotic expansion $y = \psi^{(k,m)}(z) = \psi_0^{(k,m)}(z) + \varepsilon\psi_1^{(k,m)}(z) + \varepsilon^2\psi_2^{(k,m)}(z) + \dots$.

Let us now consider the case where B^r -refinements are carried out in their complete form without neglecting the time derivatives. It is immediate to see from Tables 1.7, 1.8 that \mathbf{b}^1 tends to its ideal form and that the SIM \mathcal{M}_ε of the Davis-Skodje system, $y = y/(1+z)$, is generated term by term by the CSP manifold $\mathcal{K}(k, m)$. The latter result was also obtained through the Fraser-Roussel method in Sect. 1.5.

(k, m)	$\mathbf{b}^1(k, m)$	Λ_1^1	Λ_2^2	Λ_2^1
$(0, 0)$	$(1, 0)$	$-\frac{1}{\varepsilon}$	-1	$\frac{1}{\varepsilon} \frac{1+z-\varepsilon+z\varepsilon}{(1+z)^3}$
$(1, 0)$	$(1, -\frac{1}{(1+z)^2} + \varepsilon \frac{1-z}{(1+z)^3})$	$-\frac{1}{\varepsilon}$	-1	$\frac{1+z-\varepsilon+z\varepsilon}{(1+z)^3}$
$(2, 0)$	$(1, -\frac{1}{(1+z)^2} + \varepsilon^2 \frac{1-4z+z^2}{(1+z)^4})$	$-\frac{1}{\varepsilon}$	-1	$-\varepsilon \frac{-1+\varepsilon-4z\varepsilon+z^2(1+\varepsilon)}{(1+z)^4}$
\vdots	\vdots	\vdots	\vdots	\vdots
$(\infty, 0)$	$(1, -\frac{1}{(1+z)^2})$	$-\frac{1}{\varepsilon}$	-1	0

Table 1.7: Effect of B^r -refinements including time derivatives, starting from basis (1.125)-(1.126).

A 3 Species Kinetics Problem

The model problem analyzed in the previous section allowed us to demonstrate the ability of CSP refinements to return the correct asymptotic expansion of the slow invariant manifold for a singularly perturbed system possessing an explicit, constant, small parameter ε and prescribed slow and fast

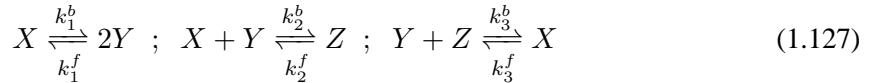
$\psi^{(k,m)}(z)$	$\mathcal{O}(1)$	$\mathcal{O}(\varepsilon)$	$\mathcal{O}(\varepsilon^2)$	$\mathcal{O}(\varepsilon^3)$	$\mathcal{O}(\varepsilon^4)$
ψ_0	ψ_1	ψ_2	ψ_3	ψ_4	
$\psi^{(0,0)}(z)$	$\frac{z}{1+z}$	$-\frac{z}{(1+z)^2}$	0	0	0
$\psi^{(1,0)}(z)$	$\frac{z}{1+z}$	0	$\frac{(z-1)z}{(1+z)^3}$	0	0
$\psi^{(2,0)}(z)$	$\frac{s}{1+z}$	0	0	$-\frac{z(1-4z+z^2)}{(1+z)^4}$	0
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
$\psi^{(\infty,0)}(z)$	$\frac{z}{1+z}$	0	0	0	0
$\psi_{ILD M}(z)$	$\frac{z}{1+z}$	0	$\frac{2z^2}{(1+z)^3}$	$\frac{2z^2}{(1+z)^3}$	$\frac{2z^2}{(1+z)^3}$

Table 1.8: CSP manifolds obtained from successive B^r -refinements including time derivatives, starting from basis (1.125)-(1.126). The CSP manifold $\mathcal{K}(k, m)$ is expressed in terms of its asymptotic expansion $y = \psi^{(k,m)}(z) = \psi_0^{(k,m)}(z) + \varepsilon\psi_1^{(k,m)}(z) + \varepsilon^2\psi_2^{(k,m)}(z) + \dots$.

variables. Moreover, the Davis-Skodjie model allowed us to evaluate the CSP basis vectors and their time derivatives analytically due to its low dimensionality ($N = 2$).

In this section, we introduce a second model problem, mimicking a stiff chemical kinetics mechanism, characterized by having a non-constant, non-explicitly defined small parameter and non-prescribed slow and fast variables. Moreover, the CSP basis vectors and their derivatives are sufficiently simple that they can still be evaluated analytically, but complicated enough so as to allow a non-trivial validation of the implementation procedures for the CSP refinements illustrated in Secs. 1.8.2 and 1.8.3.

Let us consider the following set of three symbolic reactions modelling the dynamics of 3 species X , Y and Z :



Let us also assume that the forward k^f and reverse k^b reaction rate constants of these three reactions are defined as:

$$k^f = (k_1^f, k_2^f, k_3^f) = \left(\frac{5}{\varepsilon}, \frac{1}{\varepsilon}, 1 \right) \quad (1.128)$$

$$k^b = (k_1^b, k_2^b, k_3^b) = \left(\frac{5}{\varepsilon}, \frac{1}{\varepsilon}, 1 \right) \quad (1.129)$$

where ε is a small parameter. The set of ODE's describing the evolution of Eqs. (1.127) is thus:

$$\frac{d\mathbf{W}}{dt} = \mathbf{g} \quad (1.130)$$

where $\mathbf{W} = (X, Y, Z)^T$ and the vector field is:

$$\mathbf{g} = \begin{bmatrix} -\frac{5X}{\varepsilon} - \frac{XY}{\varepsilon} + YZ + \frac{5Y^2}{\varepsilon} + \frac{Z}{\varepsilon} - X \\ 10\frac{X}{\varepsilon} - \frac{XY}{\varepsilon} - YZ - 10\frac{Y^2}{\varepsilon} + \frac{Z}{\varepsilon} + X \\ \frac{XY}{\varepsilon} - YZ - \frac{Z}{\varepsilon} + X \end{bmatrix} \quad (1.131)$$

System 1.130 is stiff, this showing from a large spectral gap in the eigenvalues of the Jacobian (see Fig. 1.6) which are real and negative and were computed numerically as the system was integrated with the LSODE package [41].

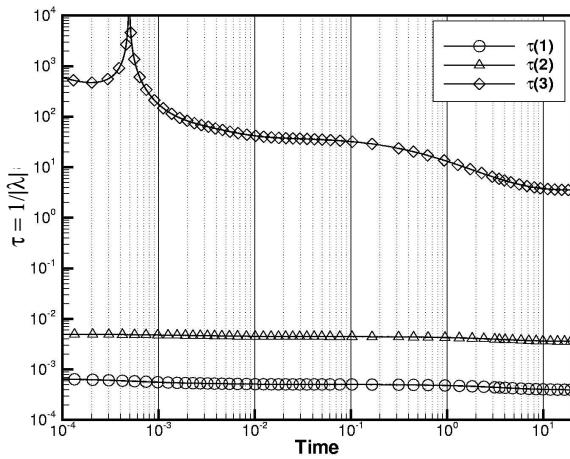


Fig. 1.6: Reciprocals of the absolute instantaneous value of the eigenvalues of J ; $\varepsilon = 0.01$, $X(0)=0.5$, $Y(0)=0.5$, $Z(0)=0.5$.

As shown in Fig. 1.7, any trajectory is attracted towards a one-dimensional manifold leading to the equilibrium point $\mathbf{W}^* = (1, 1, 1)^T$, after having experienced a fast initial transient during which the two (fast) mode amplitudes associated with the two fast time scales vanish.

The problem can be cast in CSP form as:

$$\frac{d\mathbf{W}}{dt} = \mathbf{a}_1 f^1 + \mathbf{a}_2 f^3 + \mathbf{a}_3 f^3 \quad (1.132)$$

where \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 are 3-dimensional column CSP vectors, and the f^i are the corresponding modal amplitudes, the usual CSP ordering notwithstanding of faster modes first. As the fastest amplitude vanishes, $f^1 \approx 0$, we set $M = 1$. Similarly, as the second fastest amplitude vanishes, $f^1 \approx 0$, we have $M = 2$. Correspondingly we can construct simplified problems exhibiting only the nonvanishing

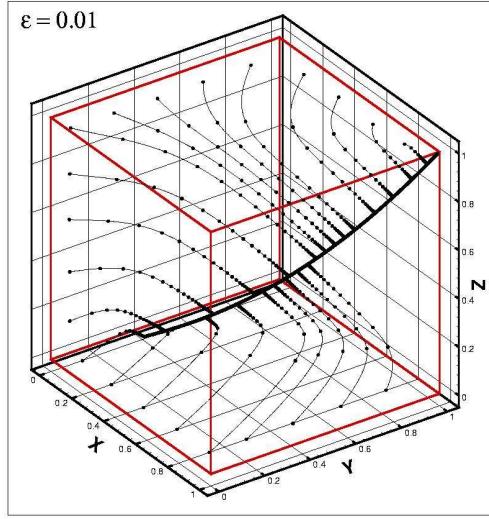


Fig. 1.7: All solution trajectories in the phase space are attracted by a 1-D manifold; markers are drawn at constant time intervals; $\epsilon = 0.01$.

modes. For the first two modes (or amplitudes) to be declared vanished (or "exhausted") within a pre-specified tolerance, the following conditions must be met:

$$\begin{aligned} \tau_2 \|\mathbf{a}_1 f^1\| &< \epsilon_{rel} \|\mathbf{W}\| + \epsilon_{abs} \\ \tau_3 \|\mathbf{a}_1 f^1 + \mathbf{a}_2 f^2\| &< \epsilon_{rel} \|\mathbf{W}\| + \epsilon_{abs} \end{aligned} \quad (1.133)$$

where τ_i denotes the timescale and corresponds to $1/\Lambda_i^i$, ϵ_{rel} is a relative error and ϵ_{abs} is an absolute error. In the numerical results that will be reported next, $\epsilon_{rel} = 1.e - 03$ and $\epsilon_{abs} = 1.e - 10$.

A leading order approximation of the one-dimensional manifold is the ILDM which can be obtained analytically by employing the left eigenvectors of J as the CSP covectors \mathbf{b}^i and by setting the two fastest modal amplitudes f^1, f^2 equal to zero:

$$f^1 = \mathbf{b}^1 \cdot \mathbf{g} = 0 \quad f^2 = \mathbf{b}^2 \cdot \mathbf{g} = 0 \quad (1.134)$$

This yields two algebraic relations which can be solved for Z for convenience, this not necessarily implying that Z is the fast variable:

$$Z = \frac{X}{Y} - \frac{5(-X + Y^2)(1 + 3X + 4Y + 6Y^2)}{2Y(X + 2Y + 2Y^2)\varepsilon} \quad (1.135)$$

$$Z = \frac{X}{Y} + \frac{-X - 3X^2 - 4XY - 5XY^2 + 3X^2Y^2 + 4XY^3 + 6XY^4}{Y(1 + 3X + 4Y + 6Y^2 + Y\varepsilon + XY\varepsilon + 4Y^2\varepsilon + 2Y^3\varepsilon)} \quad (1.136)$$

The solution of Eqs. (1.135)-(1.136), is the intersection of two surfaces in the phase space as shown in Fig. 1.8. It is also apparent how the trajectory in Fig. 1.8 progressively relaxes on the intersection of the two surfaces as soon as the two fastest modes, associated with the two fastest time scales, vanish one after the other.

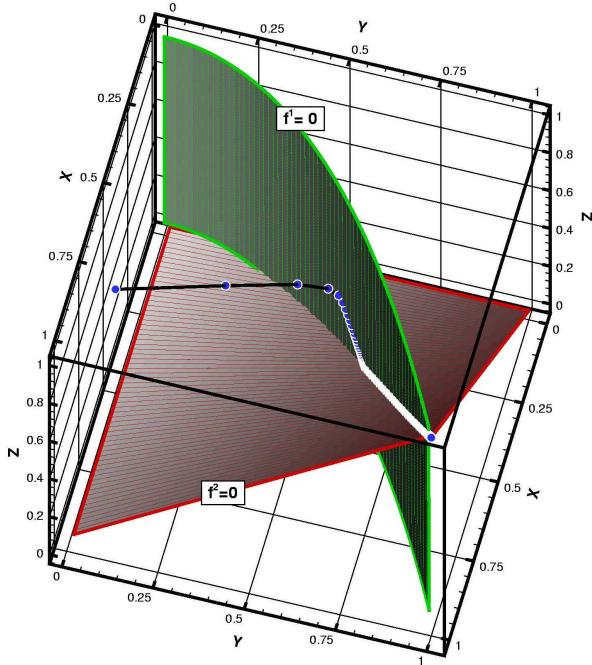


Fig. 1.8: One-dimensional manifold defined as the intersection of the two surfaces of Eq. (1.134)

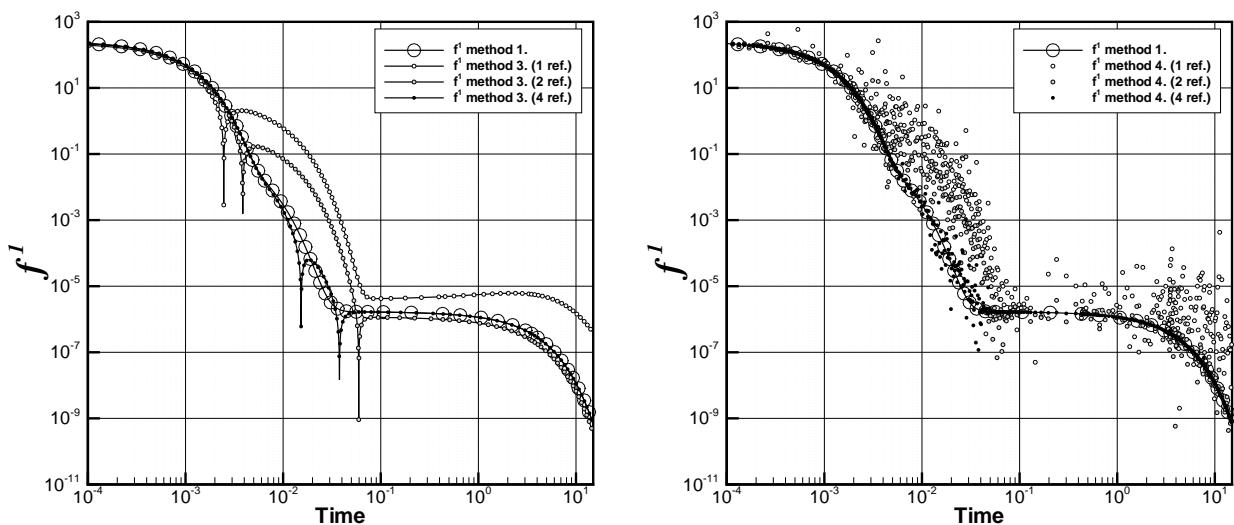
As with the Davis-Skodje problem, we will apply the CSP basis vector refinement procedure to this 3-species dynamical system. In this model problem, the CSP basis vectors can be found both algebraically or numerically by adopting the algorithms presented in Secs. 1.8.2 and 1.8.3. A number of different refinement strategies can be envisaged, each diversified on the basis of (*i*) the choice of the initial basis vectors, (*ii*) the number of initial refinements not including the basis vectors time derivatives, and (*iii*) whether a final refinement including time derivatives is carried out (recalling that a maximum of one such refinements is feasible using the formulae given in Secs. 1.8.2 and 1.8.3). Table 1.9 summarizes all the different strategies considered here.

Method	Initial Basis	No. of Refinements: $\frac{dJ}{dt} = 0$	No. of Refinements: $\frac{dJ}{dt} \neq 0$
1	Eigenvectors	0	0
2	Eigenvectors	0	1 (M fixed)
3	Identity	1, 2, 4	0
4	Random	1, 2, 4	0

Table 1.9: List of refinement strategies (methods) and of the associated algorithmic options.

No analytical expression of the one-dimensional SIM \mathcal{M}_ε is clearly available here. However we know that it corresponds to the locus $\{(X, Y, X)^T | f^1 = 0, f^2 = 0\}$ where the fast amplitudes f^1 and f^2 are evaluated using an ideal basis that fully decouples the modes. Thus a straight forward assessment of the effectiveness of the CSP basis vectors refinements is to monitor, along a trajectory, the evolution of f^1 and f^2 , these being evaluated through the refined CSP vectors, and verify whether the refinements induce such fast amplitudes to saturate to lower values on the CSP manifold. If this is the case the refinements will have resulted in a CSP manifold closer to the system's SIM.

Method 1 involves the eigenvectors of J as the projection basis vectors and hence coincides with the ILDM method. We adopted the results of this method (reported in Figs. 1.9 to 1.10) as a reference enabling us to compare all other methods described in Table 1.9.

Fig. 1.9: Methods 3 (left) and 4 (right). An initial identity or random basis converge to the eigenvectors of J (method 1) under the action of consecutive refinements.

First, we want to assess that the CSP refinements, when the contribution of the time derivatives of the basis vectors is not included (see phase (*i*) in Sec.1.8.2), converge to the results of Method 1 (i.e. an initial arbitrary basis converges to the eigenvectors of J). To this aim, Method 3 (initial unit basis) and Method 4 (initial random basis) involve a number (1,2, or 4) of CSP refinements wherein the contribution of the time derivatives of the basis vectors is neglected. The results shown in Fig. 1.9 confirm that the evolution of the fast amplitude f^1 as obtained by Method 3 and 4 converge to that obtained by Method 1. This result may be thought of as the numerical counterpart of the analytical results obtained for the Davis-Skodje problem displayed in Table 1.6, Sec.1.8.4.

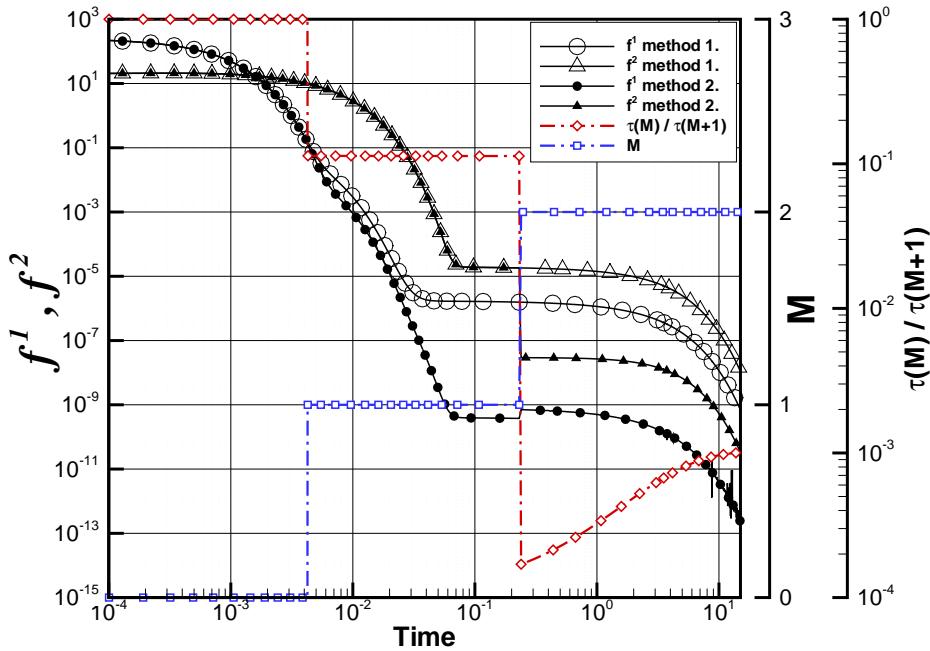


Fig. 1.10: Method 2. Single refinement of the eigenvector basis with f^1 fixed \tilde{M} .

Next, we want to assess the effect of the CSP refinements when the contribution of the time derivatives of the basis vectors is indeed included. To this aim, Method 2 involves starting with the eigenvector basis (as in Method 1), then finding the number of exhausted modes \tilde{M} with respect to this basis, and finally carrying out one CSP refinement wherein the contribution of the time derivatives of the basis vectors is included (see phase (*ii*) in Sec.1.8.2) and assuming $M = \tilde{M}$. Figure 1.10 shows the number \tilde{M} of exhausted modes found by Method 1, according to the criteria of Eqs. 1.133, and the evolution of the fast amplitudes f^1 and f^2 as obtained by both Method 1 and

Method 2. Since the B^r -refinement (the only relevant for the fast amplitudes) affects only the first M amplitudes, then when $\tilde{M} = 1$, only $f^1 = \mathbf{b}^1 \mathbf{g}$ is affected and f^2 is not, while when $\tilde{M} = 2$, both f^1 and f^2 are affected. It is clear from Figure 1.10 that by not neglecting time derivative terms, refinements cause the amplitudes f^1 and f^2 to drop to much lower values (the order of the drop is τ_M / τ_{M+1}) so that a more accurate description of the manifold is obtained.

Chapter 2

Chemical Kinetics Simplification by CSP

2.1 Introduction

In this chapter we will develop an algorithm, following a suggestion by Lam in [42], to automatically generate simplified (skeletal) chemical kinetics mechanisms for reactive systems on the basis of a Computational Singular Perturbation (CSP) database. We will primarily focus our attention on spatially homogeneous systems and then discuss the algorithm's applications to premixed laminar flames. The CSP database used comprises suitable indices computed on the basis of prior detailed chemical kinetics simulations. Differently from Lam's indications, we introduce and use fast importance indices, rather than mode participation indices, to discard unimportant reactions from the detailed mechanism among those concurring in the construction of the slow manifold. The simplified mechanism accurately replicates the kinetics of a kernel set of species specified by the user. The reduction in the overall dimension of the ODE system is obtained at the cost of a bounded loss in accuracy. The simplification algorithm was tested with the GRI-Mech 3.0 mechanism to produce NO (nitric oxide) accurate simplified mechanisms for the spatially homogeneous system. The diagnostic features of the algorithm will be addressed together with an accuracy assessment and speed-up analysis to establish the computational gain offered by the simplified mechanisms. The concluding section of the chapter will address the algorithm's applications to premixed laminar flames.

2.2 Theoretical Background

The set of ODE's describing the dynamics of a N -dimensional system defined in $C \subseteq R^N$ can be written in the form:

$$\frac{d\mathbf{u}(t)}{dt} = \mathbf{g}(\mathbf{u}) \quad ; \mathbf{u} \in C \quad ; \mathbf{g} \in TC\mathbf{u} \quad (2.1)$$

where $TC\mathbf{u}$ is the tangent bundle at $\mathbf{u} \in C$. Typically combustion processes are characterized by stiff systems possessing a slow invariant manifold associated with the equilibrium point. The existence of the manifold is granted by the presence of a significant temporal gap in the spectrum of the system's intrinsic timescales and can be defined as the locus of the phase space where no fast timescale is still active. In chapter 3 we will thoroughly discuss the definition of an intrinsic timescale spectrum and its relationship to the invariant geometric properties of the system but for the time being it will be sufficient to consider such spectrum as the set of eigenvalues of the Jacobian matrix of \mathbf{g} . The stiffness of Eq. (2.1) will be manifested in the eigenvalues of the Jacobian matrix of \mathbf{g} , a number of them being large in magnitude and with a dominant real and negative part.

For the analysis of a stiff problem involving N unknowns, the basis of the CSP method is the splitting of the N -dimensional space in two subdomains, in which the fast and the slow time scales act. An additional subdomain, the stoichiometric manifold, corresponds to the atomic species conservation laws. Here, a brief summary of the CSP method is presented and the related CSP tools will be defined.

For a set of equations such as Eq. (2.1), there exists a specific set of basis vectors $[\mathbf{a}_1(t), \dots, \mathbf{a}_N(t)]^T$ and its dual $[\mathbf{b}^1(t), \dots, \mathbf{b}^N(t)]$ which span R^N at time t , such that the N -dimensional domain of \mathbf{u} can be split in two subdomains:

- one (M)-dimensional (fast) subdomain E_r wherein the fast time scales act;
- one $N - (M+E)$ -dimensional (slow) subdomain E_s wherein the slow time scales act, defined as the slow manifold;
- one E -dimensional subdomain E_e , corresponding to the E atomic species conservation laws and defined as the stoichiometric manifold.

so that $R^N = E_r \oplus E_s \oplus E_e$. As the system evolves in time, the basis vectors might rotate and the subdomains change their dimensions. The criterion used to find the dimension M of the fast subdomain will be described in what follows.

CSP inspects the projection of \mathbf{g} into the fast subdomain. When this projection becomes exponentially small, CSP provides a simplified system of equations containing no fast time-scales. This process is done in such a way that the solution computed on the basis of the simplified system differs from the reference solution obtained by integrating the original system, Eq. (2.1), within the

desired accuracy.

Eq. (2.1) can be rewritten by expanding \mathbf{g} in terms of the CSP basis vectors as follows:

$$\frac{d\mathbf{u}}{dt} = \sum_{r=1}^M \mathbf{a}_r f^r + \sum_{s=M+1}^{N-E} \mathbf{a}_s f^s = \mathbf{g}_{fast} + \mathbf{g}_{slow} \quad (2.2)$$

where $\mathbf{a}_r \in E_r$, $\mathbf{a}_s \in E_s$ and their dual $\mathbf{b}^r \in E_r^*$ dual space of E_r , $\mathbf{b}^s \in E_s^*$ dual of E_s , and where

$$f^i = \mathbf{b}^i \cdot \mathbf{g} \quad i = 1, \dots, N \quad (2.3)$$

When the projection of \mathbf{g} into the fast subdomain is small, then the M non-linear algebraic equations (also referred to as equations of state):

$$f^r = \mathbf{b}^r \cdot \mathbf{g} \approx 0 \quad r = 1, \dots, M \quad (2.4)$$

are the equations for the $N - (M + E)$ -dimensional slow manifold hyper-surface embedded in R^N . The solution of Eq. (2.1) is constrained to lie close to the slow manifold (after an initial transient) and its evolution is governed by the non-stiff simplified system obtained by enforcing Eqs.(2.4) into Eq. (2.2):

$$\frac{d\mathbf{u}}{dt} \approx \mathbf{a}_{M+1} f^{M+1} + \dots + \mathbf{a}_N f^N = \mathbf{g}_{slow} \quad (2.5)$$

The M modes having vanishing amplitudes f^r are referred to as *fast exhausted* modes. The $N - M$ modes having non vanishing amplitudes f^s are referred to as *slow active* modes. The criterion to find the dimension M of the fast subdomain is the following. Let us first introduce an error vector \mathbf{u}_{err} built on the basis of the solution vector \mathbf{u} , as follows:

$$u_{err}^i = \varepsilon_{rel}^i |u^i| + \varepsilon_{abs}^i \quad (2.6)$$

where ε_{rel}^i and ε_{abs}^i are the maximum relative and absolute errors on the i -th element of the solution (state) vector \mathbf{u} respectively. The number M of fast modes which - within the limits of accuracy specified by the given error vector - are considered exhausted, is defined as the largest integer lying between 1 and N which satisfies the inequality for each $i = 1, \dots, N$:

$$|\tau^{M+1} \sum_{r=1}^M a_r^i f^r| < u_{err}^i \quad (2.7)$$

where τ^{M+1} denotes the fastest of the active (non-exhausted) timescale. If the eigenvectors of the Jacobian matrix J of \mathbf{g} are used as CSP basis vectors, then we assume $\tau^i = 1/|\lambda^i|$ where λ^i is the eigenvalue of J corresponding to the i -th eigenvector.

Throughout the present chapter the set of eigenvectors of the Jacobian matrix J of $\mathbf{g}(t)$ is taken as the set of CSP basis vectors $\{\mathbf{a}_i(t)\}_{i=1}^N$. While this choice may be reasonable for systems possessing an equilibrium point (a trivial limit set), it may bear no relationship to the spectrum of intrinsic timescales for systems possessing richer asymptotic behaviors such as persistent oscillations or chaotic oscillations (non-trivial limit sets). Provided the system has an invariant slow manifold associated to an equilibrium point, CSP theory can indeed provide a refinement procedure [10] of the eigenvectors capable of further increasing the decoupling between the dynamics of the slow and fast modal amplitudes thereby increasing the accuracy of the statements made in Eqs.(2.4) and (2.5). In this latter case, the timescales τ^i will coincide with the inverse absolute values of the diagonal elements of matrix Λ introduced in chapter 1 section 1.8.

CSP can also serve as a diagnostic tool to analyze the system evolution. To illustrate this feature, let us assume that the detailed kinetic mechanism consists of N_s species, E elements and N_r elementary reactions. For a constant pressure, spatially homogeneous combustion, let us write the conservation equations of the N_s species mass fractions Y^i and the equation for the time rate of change of the temperature, T , in the CSP canonic form:

$$\frac{d\mathbf{y}}{dt} = \mathbf{S}_1 R^1 + \mathbf{S}_2 R^2 + \dots + \mathbf{S}_{2N_r} R^{2N_r} = \mathbf{SR} \quad (2.8)$$

where, \mathbf{u} is the N -dimensional vector ($N = N_s + 1$) of the unknowns (Y^1, \dots, Y^{N_s}, T) .

Further, the vectors \mathbf{S}_1 to \mathbf{S}_{2N_r} define the $(N, 2N_r)$ -dimensional generalized stoichiometric matrix S as

$$S = [\mathbf{S}_1, \mathbf{S}_2, \dots, \mathbf{S}_{2N_r}] = [Q\tilde{S}, Q\tilde{S}] \quad (2.9)$$

where the (N_s, N_r) matrix of stoichiometric coefficients \tilde{S} is defined as:

$$\tilde{S}_{ik} = \nu_{ik}, \quad i = 1, \dots, N_s; \quad k = 1, \dots, N_r \quad (2.10)$$

where ν_{ik} is the net stoichiometric coefficient of species i in reaction k . Note that only $N_s - N_e$ stoichiometric vectors \mathbf{S}_j are linearly independent, since the kinetic system is constrained to evolve on a N_e -dimensional stoichiometric manifold, whose geometry is defined by the conservation laws

for the N_e atomic species. In turn, this circumstance prompts the occurrence of N_e zero eigenvalues in the Jacobian matrix of \mathbf{g} .

The (N, N_s) matrix Q in Eq. (2.9) is defined as:

$$Q = \begin{bmatrix} \frac{W_1}{\rho} & 0 & \dots & 0 & 0 \\ 0 & \frac{W_1}{\rho} & \dots & 0 & 0 \\ \dots & 0 & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & \frac{W_{N_s}}{\rho} \\ -\frac{W_1}{\rho c_p} h_1 & \dots & \dots & -\frac{W_{N_s-1}}{\rho c_p} h_{N_s-1} & -\frac{W_{N_s}}{\rho c_p} h_{N_s} \end{bmatrix} \quad (2.11)$$

where ρ is the mixture density, W_i is molecular weight of the i -th species, h_i is the specific enthalpy of the i -th species, and c_p is the average specific heat at constant pressure of the mixture.

Finally, the rates R^1 to R^{2N_r} in Eq. (2.8) define the generalized $(2N_r)$ -dimensional vector \mathbf{R} as:

$$\mathbf{R} = [r_{f,1}, r_{f,2}, \dots, r_{f,N_r}, -r_{b,1}, -r_{b,2}, \dots, -r_{b,N_r}] \quad (2.12)$$

where r_f and r_b are the third-body and pressure-fall-off corrected (if applicable) forward and reverse rates of progress of the N_r reversible reactions ¹.

The adoption of this form allows us to recast Eq. (2.4) and (2.5) as:

$$f^r = \sum_{k=1}^{2N_r} (\mathbf{b}^r \cdot \mathbf{S}_k) R^k = \sum_{k=1}^{2N_r} B_k^r R^k = 0 \quad r = 1, \dots, M, \quad (2.13)$$

$$\frac{du^i}{dt} = \sum_{s=M+1}^N a_s^i \sum_{k=1}^{2N_r} (\mathbf{b}^s \cdot \mathbf{S}_k) R^k = \sum_{k=1}^{2N_r} C_k^i R^k \quad i = 1, N. \quad (2.14)$$

where $B_k^r = \mathbf{b}^r \cdot \mathbf{S}_k$, $C_k^i = \sum_{s=M+1}^N a_s^i (\mathbf{b}^s \cdot \mathbf{S}_k)$, and k ranges from 1 to $2N_r$.

¹Note that the system $d\mathbf{y}/dt = S\mathbf{R}$ will then result into

$$\begin{aligned} \frac{dY^i}{dt} &= \frac{W_i}{\rho} \sum_{k=1}^{2N_r} \nu_{ik} R^k = \frac{W_i}{\rho} \dot{\omega}_i \\ \frac{dT}{dt} &= -\frac{1}{\rho c_p} \sum_{i=1}^{N_s} W_i h_i \sum_{k=1}^{2N_r} \nu_{ik} R^k = -\frac{1}{\rho c_p} \sum_{i=1}^{N_s} W_i h_i \dot{\omega}_i \end{aligned}$$

where $\dot{\omega}_i$ is the molar production rate of the i -th species and where we assumed $\nu_{i(2k)} = \nu_{ik}$ for $k = 1, N_r$. We should also point out that the above system is easily solvable by numerical integration coupled with CHEMKIN software. Moreover an analytical form (rather than a numerical estimate) of the Jacobian matrix of the chemical source term, whose elements are $\partial g^i / \partial y^j$, is available, and was indeed found to determine the CSP basis vectors as its eigenvectors.

2.3 Pointers to CSP Radicals

Following [12, 17], the concept of CSP radical was introduced, in the context of CSP theory, as an equivalent of quasy steady state species. As in the quasy-steady-state assumption, which yields approximate algebraic relations between the entries of \mathbf{u} , CSP yields the *M equations of state* (2.4), corresponding to the exhausted M fastest modes, and which define the slow manifold. The equations of state can therefore be used to solve for M unknowns in terms of the remaining $(N-M)$, thereby eliminating the need for M ODE's in system (2.1). Following Lam's definition (see also [14, 16]), CSP radicals are the species that should be solved for in the equations of state. The identification of CSP radicals is not arbitrary (and should neither rely on experience and intuition, potentially ambiguous criteria), but rather it is based on the concept of ad-hoc scalar quantities called 'pointers' to CSP radicals. Each species is associated to a fast mode and the CSP radical candidate is selected as that having the gratest pointer.

CSP identifies one species for each fast mode m as a CSP radical candidate as follows. First, the $N \times N$ fast mode projection matrix Q_m is constructed, formed by the dyadic product of the right eigenvector (or refined CSP vector) \mathbf{a}_m times the dual \mathbf{b}^m corresponding to the m -th fast mode:

$$Q_m = \mathbf{a}_m \mathbf{b}^m \quad (\text{no sum on } m) \quad (2.15)$$

Because of the biorthonormality of \mathbf{a}_i and \mathbf{b}^j , expressed by $\mathbf{b}^j \mathbf{a}_i = \delta_i^j$, the following holds with I being the identity matrix:

$$\sum_{m=1}^N Q_m = AB = BA = I \quad (2.16)$$

where $A = [\mathbf{a}_1(t), \dots, \mathbf{a}_N(t)]^T$ and $B = [\mathbf{b}^1(t), \dots, \mathbf{b}^N(t)]$. As already noted in Eq.(2.2) the system of Eq.(2.1) may be rewritten as

$$\frac{d\mathbf{u}}{dt} = \mathbf{g} = AB\mathbf{g} = \left(\sum_{m=1}^N Q_m \right) \mathbf{g} \quad (2.17)$$

By letting \mathbf{e}^k be a N -dimensional row vector with all zeros except the k -th element which contains unity and \mathbf{e}_k its dual coloumn vector, the evolution of the k -th component of \mathbf{u} , i.e. u^k , may be written as

$$\frac{du^k}{dt} = g^k = \mathbf{e}^k \mathbf{g} = \left(\sum_{m=1}^N \mathbf{e}^k Q_m \right) \mathbf{g} = \sum_{i=1}^N \sum_{m=1}^N \mathbf{e}^k Q_m \mathbf{e}_i g^i \quad (2.18)$$

From Eq.(2.16) we clearly have that $\sum_{m=1}^N \mathbf{e}^k Q_m \mathbf{e}_i = \delta_i^k$ and therefore we can formally write

$$\frac{du^k}{dt} = \sum_{m=1}^N \mathbf{e}^k Q_m \mathbf{e}_k g^k = (\mathbf{e}^k Q_1 \mathbf{e}_k) g^k + \dots + (\mathbf{e}^k Q_N \mathbf{e}_k) g^k \quad (2.19)$$

where the generic term $\mathbf{e}^k Q_m \mathbf{e}_i$ is the k -th diagonal entry of matrix Q_m .

Eq.(2.19) clearly shows the relative influence of each mode $m = 1, \dots, N$ on the evolution of the k -th species. We can therefore define the scalar quantity

$$\hat{Q}_m(k) = |\mathbf{e}^k Q_m \mathbf{e}_k| \quad (2.20)$$

as a *CSP radical pointer* of the m -th fast mode. Whenever $\hat{Q}_m(k)$ is non-small then the corresponding k -th component of \mathbf{u} is a candidate as a *CSP radical* and will be among those affected the most by the reactions involved in mode m . Geometrically $\hat{Q}_m(k) = |(\mathbf{e}^k \mathbf{a}_m) (\mathbf{b}^m \mathbf{e}_k)|$ is a measure of the mutual projection of the m -th mode \mathbf{a}_m onto the k -th coordinate axis representing the k -th species.

In fact, each fast mode can select more than one CSP radical, but not necessarily the one with the largest pointer will be a true CSP radical. To address this issue, one must consider the whole set of CSP radical candidates as selected by all the M fast modes. To this aim, one can associate M species to the M -dimensional fast subspace, as opposed to associating them to individual modes, by first constructing the fast subspace matrix Q_M formed by the sum of the M fast mode projection matrices:

$$Q(M) = \sum_{m=1,M} Q_m \quad (2.21)$$

and then by considering the vector $\hat{Q}(M)(k)$ defined as:

$$\hat{Q}(M)(k) = |\mathbf{e}^k Q(M) \mathbf{e}_k| \quad (2.22)$$

where $\hat{Q}(M)(k)$ is the k -th diagonal element of $Q(M)$. The elements of $\hat{Q}(M)(k)$ can then be sorted in descending order according to their magnitude:

$$\begin{aligned} \hat{Q}(M)(k_{csp}(1)) &> \dots > \hat{Q}(M)(k_{csp}(M)) >> \dots \\ \dots &>> \hat{Q}(M)(k_{csp}(M+1)) > \dots > \hat{Q}(M)(k_{csp}(N)) \end{aligned} \quad (2.23)$$

where the subscript k_{csp} is a vector of integers storing in its first M elements the indices pointing to the M CSP radicals, and in the last $N - M$ elements the indices pointing to the $N - M$ non-CSP radicals. This way the original state vector \mathbf{u} can be partitioned in two subsets: the CSP radicals and the remaining species.

2.4 Importance Indices

It is of interest to assess the relative importance of each reaction in affecting both the slow and the fast dynamics of a given species. As far as the slow, manifold-constrained dynamics is concerned this can be done by normalizing \mathbf{g}_{slow} in Eq.(2.5) with the introduction of the “slow subspace” importance index (Ref. [12]), defined as:

$$(I_k^i)_{slow} = \sum_{s=M+1}^N a_s^i (\mathbf{b}^s \cdot \mathbf{S}_k) R^k / \sum_{j=1}^{N_r} \left| \sum_{s=M+1}^N a_s^i (\mathbf{b}^s \cdot \mathbf{S}_j) R^j \right| \quad (2.24)$$

where no sum on k is assumed, $i = 1, N$ and where by definition the sum of the absolute values of the importance index is 1, and therefore each importance index is always smaller than 1. The sum, $(I^i)_{slow}$, over all reactions of the “slow subspace” importance index defined as:

$$(I^i)_{slow} = \sum_{j=1}^{2N_r} (I_j^i)_{slow} = \frac{\sum_{j=1}^{2N_r} \left[\sum_{s=M+1}^N a_s^i (\mathbf{b}^s \cdot \mathbf{S}_j) \right] R^j}{\sum_{j=1}^{2N_r} \left| \left[\sum_{s=M+1}^N a_s^i (\mathbf{b}^s \cdot \mathbf{S}_j) \right] R^j \right|} ; i = 1, \dots, N$$

is typically a non vanishing quantity, its value expressing the nondimensional time rate of change of a species u^i occurring at the slow time scales.

For a given species, “the elementary reactions with the largest $(I_k^i)_{slow}$ are the rate controlling reactions for that species” [42]. On the other hand, a further set of (fast) reactions, generating fast timescales, is responsible for the trajectory to stay on the manifold. If the state of the system were perturbed off the manifold, a subset of species, the CSP radicals [43], would respond according to the fast timescales so as to force the trajectory to relax back on the manifold. To address the issue of identifying such reactions, Lam proposed to use the participation indices to the fast modes, relating reactions to fast modes, in conjunction with CSP radical pointers, relating fast modes to CSP radical species. This indirect connection may however be bypassed by defining a fast subspace importance index,

$$(I_k^i)_{fast} = \sum_{r'=1}^M a_{r'}^i (\mathbf{b}^{r'} \cdot \mathbf{S}_k) R^k / \sum_{j=1}^{N_r} \left| \sum_{r'=1}^N a_{r'}^i (\mathbf{b}^{r'} \cdot \mathbf{S}_j) R^j \right| \quad (2.25)$$

The sum $(I^r)_{fast}$ of this “fast subspace” importance index, over all reactions, is always very small, consistently with the fact that the fast mode amplitudes are also very small, because of the cancellations occurring between competing reactions of large and opposite reaction rates magnitude.

To summarize, a reaction can affect the evolution of a CSP radical either when:

- its contribution to the currently active (slow) modes is non-small (non-small “slow subspace” importance index). In this case the CSP radical will evolve on the currently active (slow) time scale, following the constraint set by the manifold which expresses a functional dependency of the radical from the slow evolving species.
- when its contribution to the competing effects which produces the manifold is non-small (non-small “fast subspace” importance index). In this latter case, a perturbation of the rate constants of the reactions ranked as important will produce a sudden change in the manifold geometry, which will force the CSP radical to adjust over the fast (exhausted) time scale.

A reaction can affect the evolution of non-CSP-radicals when:

- its contribution to the currently active (slow) modes is non-small (non-small “slow subspace” importance index). The non-CSP-radicals will always evolve on the currently active (slow) time scale.

CSP indices and pointers represent the core of a CSP database and constitute a set of self-consistent quantities capable of revealing the influence of reactions on individual species be it in their evolution in the slow or fast subspaces.

2.4.1 Importance indices vs. conventional sensitivity coefficients

Conventional sensitivity analysis has no notion of any geometrical decomposition of the phase space into a fast subspace and a slow subspace or manifold. On the other hand the manifold, and all reactions which concur to its generation, is the structure determining the asymptotic behaviour of the kinetic system as it is asymptotically attractive for every possible trajectory, ultimately leading to the equilibrium point. This being said, if the system should locally exhibit a high sensitivity, in the traditional sense, to some reaction rate, then a change in such rate should lead an analyst to

expect a change in the asymptotic response of the system. This, however, may not be the case if the perturbation of the reaction rate, thought of as a phase space vector, should happen to be locally parallel to the manifold as such perturbation would merely affect the system's slow evolution within the manifold towards the equilibrium point. The risk of being misled by a traditional sensitivity analysis could be overcome by defining appropriate sensitivity coefficients that separately probe both the slow evolution of the system on the manifold, taking into account perturbations projected locally onto such manifold, and the fast approach towards the manifold, with orthogonal perturbations, ultimately determining the very shape of the manifold itself and hence the overall asymptotic behaviour of the system.

CSP importance indices play a different role with respect to sensitivity coefficients in that the former will not directly measure the perturbation of a given species in response to a variation in a reaction rate. Rather CSP importance indices will give a measure of the relative influence of a reaction in the expression of the i -th component of the fast or slow vector fields \mathbf{g}_{slow} and \mathbf{g}_{fast} which ultimately determine the evolution of the i -th species be it a slow species or a CSP-radical.

2.5 Simplification of chemical kinetics

This section will describe how the slow and fast importance indices together with the pointers to CSP radicals can be used to identify the most important reactions and species at a given instant of time of the system dynamics.

The concept of “being important” is not an inherent property of a reaction or species; each of them might or might not be important only with respect to a specific target set by the analyst. Typically, what matters to an analyst is to have a selection procedure able to identify the important reactions for a subset of the dependent variables only, say a state variable like temperature plus a selected subset of species, whereas the reactions mostly affecting the remaining, non interesting, components might be discarded as well as the reactions identified as non important with respect to the selected dependent variables. A procedure to accomplish this task can be organized as follows. First, a time-accurate solution of the system Eq. (2.1) involving the detailed kinetics mechanism is found by means of a suitable numerical technique of integration. The time-accurate solution state vector $\mathbf{u}(t_n)$ is saved as a database at a number N_{sol} of time instants t_n , with $0 \leq t_n \leq t_{eq}$, $n = 1, N_{sol}$, with t_{eq} being the time at which the system reaches a stationary (equilibrium) state.

Next, an off-line diagnostic tool is run to analyze this database. Suppose that the analyst is interested in studying only a portion $D_j = \{t : t \in (t_{begin,j}, t_{end,j})\}$ of the whole time period $D = \{t : t \in$

$(0, t_{eq})\}$. Clearly D_j might as well coincide with D .

At each time $t_n \in D_j$, the diagnostic tool will compute $\mathbf{g}(\mathbf{u}(t_n))$, the corresponding Jacobian matrix $J(t_n)$, the number of exhausted modes $M(t_n)$ together with the appropriate CSP basis vectors $\mathbf{a}_r(t_n)$ and $\mathbf{b}^r(t_n)$ with $r = 1, M(t_n)$. These data will allow for the pointers to CSP radicals, the slow and fast importance indices to be determined using (2.23), (2.24) and (2.25).

Such indices are made available at every time step t_n and can be visualized as two distinct matrices

- The slow importance index matrix defined for all species:

$$(I_k^i(t_n))_{slow} \quad i = 1, N \quad k = 1, N_r$$

- The fast importance index matrix defined for the M CSP radical species alone:

$$(I_k^i(t_n))_{fast} \quad i = jcsp(1), jcsp(M) \quad k = 1, N_r$$

The underlying concept for the simplification of the kinetic mechanism at time t_n is the elimination of all the reactions whose importance indices relative to some relevant set of species are smaller than a prescribed user specified tolerance. The identification of such set of species, referred to as '*active*', on which to base such elimination may be carried out iteratively starting from a user-specified kernel set² $S_{user} = S_{[0]} \subseteq [1, N] \in \mathbb{N}$ (where subscript $[0]$ indicates the 0-th order of iteration) which may include any number of species of interest. A set of active reactions $R_{[0]}(t_n, tol) \subseteq [1, N_r] \in \mathbb{N}$ may then be defined by retaining those reactions with importance indices relative to the set $S_{[0]}$ above the prescribed tolerance tol . Slow indices are examined for all species in $S_{[0]}$, while fast indices are inspected for the subset of CSP radicals $S_{[0]}^{rad}$ only, where $S_{[0]} = S_{[0]}^{maj} \cup S_{[0]}^{rad}$ and $S_{[0]}^{maj}$ is the set of non-radical or major species. A new set of active species $S_{[1]}$ is identified by collecting all species participating in $R_{[0]}$, that is any species $i \in S_{[0]}$ such that $\nu_{ik} \neq 0$ for some $k \in R_{[0]}$, where ν_{ik} is the stoichiometric coefficient of the i -th specie in the k -th reaction as in Eq.(2.10). The procedure can thus be repeated iteratively this giving rise to a nested sequence of sets, $S_{user} = S_{[0]} \subset S_{[1]} \subset \dots \subset S_{[h]}$, until convergence is reached, that is $S_{[h]} = S_{[h-1]}$. The corresponding set of active reactions $R_{[h]}$ will represent the simplified kinetic mechanism at time t_n and, dropping the iteration index h , will be addressed as $R(t_n, tol)$ or $R(t_n, tol, S_{user})$ if the

²Subsets of species or reactions will be identified as sets of integer numbers each identifying a particular species or reaction. If say H₂O is identified with no. 6 then a set composed of H₂O alone will be written as [6] (alternatively, for sake of clarity, the same set may be addressed as [H₂O])

dependence on the initial species kernel set is to be pointed out. Fig.2.1 displays the number of active species and reactions upon convergence within the sets $S_h(t_n, tol)$ and $R_h(t_n, tol)$ respectively, for a stoichiometric methane/air combustion example that will extensively be examined in the next section.

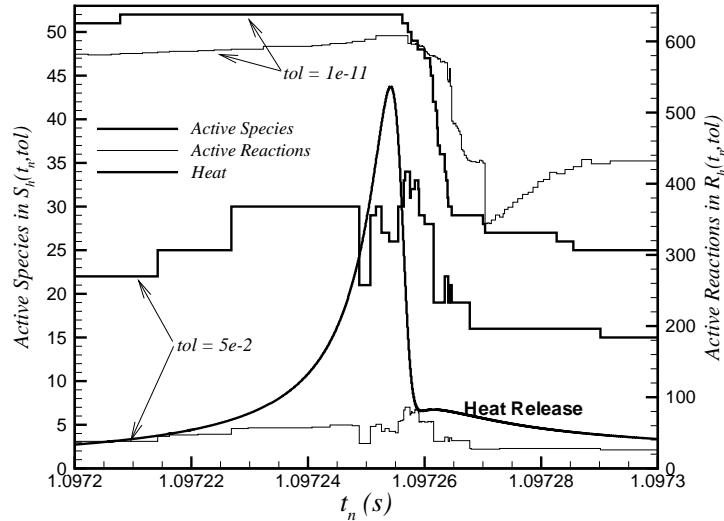


Fig. 2.1: Number of active species and reactions in sets $S_h(t_n, tol)$ and $R_h(t_n, tol)$.

The iterative procedure will be performed at every time instant $t_n \in D$ (or in a partial time domain $t_n \in D_j$) thus yielding a set of simplified kinetic mechanisms $R(t_n, tol, S_0)$ with $n = 1, N_{sol}$. A global, time independent, simplified mechanism $R(tol, S_0) \subseteq [1, N_r]$, relative to the whole time domain D (or D_j), may thus be defined as the union of all such mechanisms:

$$R(tol, S_{user}) = \bigcup_{n=1}^{N_{sol}} R(t_n, tol, S_{user}) \quad (2.26)$$

The logical structure of the algorithmic procedure just described can be found in Tab. 2.1, where the, user specified, input parameters are the kernel set of initial species S_{user} and the value of the tolerance threshold tol . Note that the simplified mechanism $R(tol, S_{user})$, being related to a solution of Eq.(2.1), is tailored to such solution and especially to the imposed pressure value and the initial temperature and composition. More comprehensive mechanisms (in the meaning of [44]) may be obtained as the union of several simplified mechanisms, each related to different initial conditions for Eq. (2.1) and pressures.

A non-iterative version of the algorithm may be achieved by not performing the innermost loop

For all $t_n, n \in [1, N_{sol}]$

{ **Start with kernel set of active species :** $S_{[h=0]} = S_{[0]} = S_{user}$;

While $[S_{[h]}(t_n, tol, S_{user}) = S_{[h-1]}(t_n, tol, S_{user})]$

{ $h = h + 1$;

Define new active reactions set :

$$R_h(t_n, tol, S_{user}) = \{k : (I_k^i(t_n))_{slow} > tol; i \in S_{[h-1]}(t_n, tol, S_{user})\} \cup \\ \{k : (I_k^i(t_n))_{fast} > tol; i \in S_{[h-1]}^{rad}(t_n, tol, S_{user})\};$$

Define new active species set :

$$S_{[h]}(t_n, tol, S_{user}) = \{i : \nu_{ik} \neq 0; k \in R_h(t_n, tol, S_{user})\};$$

};

};

Define global simplified mechanism $R(tol, S_{user})$ as in Eq.(2.26)

Table 2.1: Structure of the simplification algorithm

more than once. This clearly implies an inefficient use of the CSP data available since only importance indices relative to S_0 are eventually used. Nevertheless the resulting instantaneous simplified mechanisms $R_1(t_n, tol, S_0)$, involving only a small set of species $S_1(t_n, tol, S_0)$, have quite a significant diagnostic value as they embed the significant processes for the set of species S_0 alone, which may be constituted by a single species, at each time instant thus unveiling the basic production/destruction pathways for such species. Iteration, on the other hand, is capable of highlighting the indirect dependencies between reactions and species which will ultimately result in more accurate, and certainly more populated, simplified mechanisms.

A further important aspect to be considered is that of the initial conditions of the reactive process. Each CSP database relates to a given solution and therefore to an initial composition, temperature and pressure. Thus the simplified mechanisms obtained from a particular database will themselves be related to such initial conditions. In other words, they will be designed for such initial conditions. A comprehensive simplified mechanism however should be suitable for a wide range of conditions not merely one. This may be overcome by using more than one CSP database at a time, each relating to different initial conditions, so that a reaction that might have been discarded for some initial condition may very well be reinserted for some other thus giving the simplified mechanisms a greater comprehensiveness.

2.6 Results

The 325 reactions and 53 species GRI Mech 3.0 model [45] for methane combustion in air is used to test the simplification procedure. Solutions are obtained for a lean case (equivalence ratio $\Phi = 0.3$) and for a stoichiometric case ($\Phi = 1.0$) both at $p = 1$ atm. The initial temperature was set at

$T = 1000$ K for the stoichiometric case, and was slightly reduced to $T = 968$ K for the lean case so that an ignition time of $\sim 1s$, similar to the stoichiometric case, is achieved. Note that ignition time was defined as the time needed for the temperature to reach 5% of the overall peak temperature, and may be indicated by t_5 . The reaction time was defined as $t_{reac} = t_{95} - t_5$, i.e the time interval needed for the temperature to rise from 5% to 95% of the peak temperature. Fig.2.2 shows ignition and reaction times, final temperatures and peak heat release values as a function of initial temperature, for equivalence ratios in the range $\Phi \in [0.1, 2.0]$. Fig.2.2 shows that, for all the cases considered, ignition occurs for initial temperatures $T > 795$ K confirming that the initial temperatures chosen above fall well within the explosion limits for both mixtures, allowing for similar ignition times.

Based on the solutions, a CSP analysis is carried out so as to obtain a CSP database comprised of the slow/fast importance indices and pointers to CSP radicals. Simplified mechanisms $R(tol, S_0)$ are then computed through the algorithm described above, each mechanism corresponding to a value of the tolerance threshold tol imposed on importance indices and on the initial kernel set of species adopted S_0 .

The particular starting kernel set S_0 identifies the species whose indices are always taken into consideration within the simplification algorithm. The choice of such set is therefore ultimately crucial in the determination of the simplified mechanisms. A major factor intervening in the choice of S_0 is the accuracy with which the simplified mechanisms are required to reproduce the dynamics of a given species so that if such species is excluded from S_0 , the simplified mechanism might fail at such task. Complex dependancies among species, however, generally allow for species not included in S_0 , but related to them, to be equally summoned within the iterative process. A further aspect concerns the time domain of existence of the importance indices $I_k^i(t_n)$. If at time t_n a given specie is present in traces and its rate of production/depletion is also negligible, its index is not tabulated³. For this reason S_0 should include at least a major reactant whose indices are tabulated throughout the ignition and explosion time domains - domains which would otherwise be neglected. Species of interest such as nitric oxides, and more generally nitrogen compounds, which are present in non-trace quantities only within and after the explosion, should therefore never be chosen alone as kernel species but in conjunction with some major reactant.

In the present study both methane and nitric oxide were chosen as kernel species ($S_0 = [\text{CH}_4, \text{NO}]$), this ensuring that the whole time domain, from the initial condition to equilibrium, is uniformly

³For a species to be declared a trace species the criterion used was either an absolute threshold of 10^{-12} on their mass fraction or a threshold of 10^{-6} times the peak mass fraction.

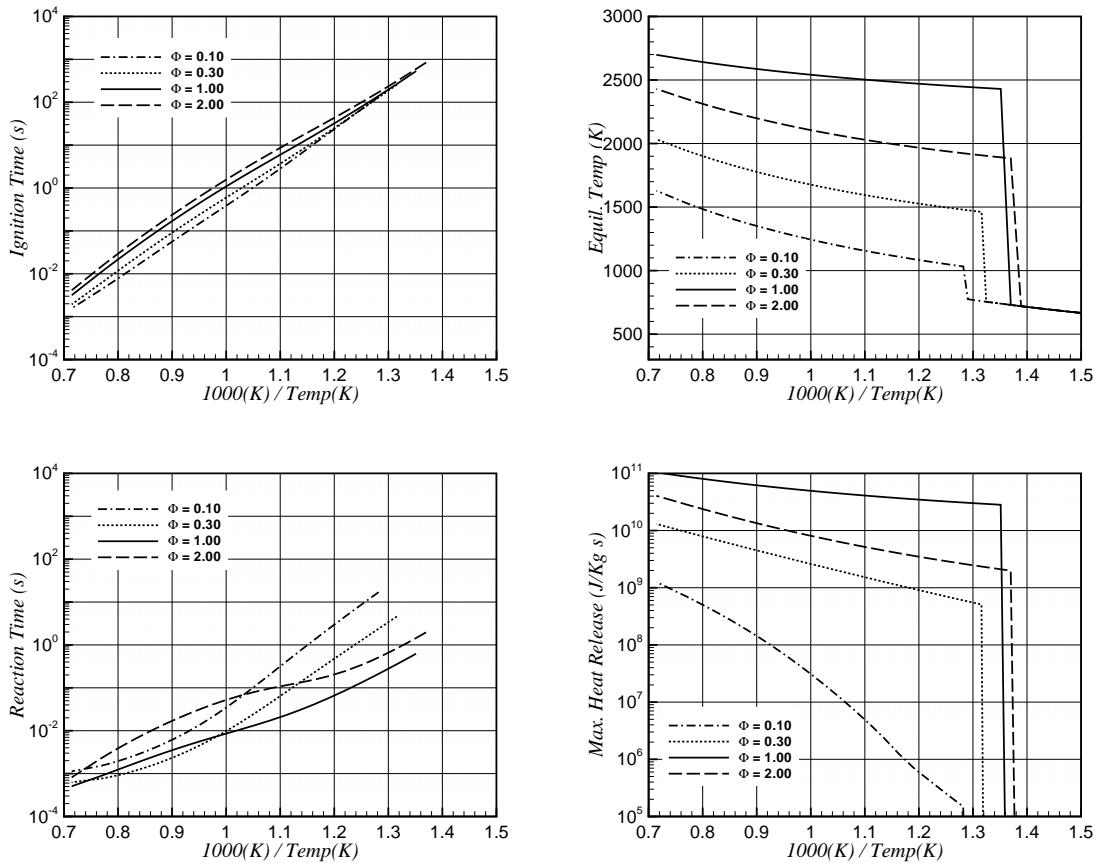


Fig. 2.2: Left two figures: Explosion limits in terms of ignition and reaction times as a function of initial temperature for a methane/air mixture at $p = 1 \text{ atm}$. Right two figures: Equilibrium temperatures and maximum heat release.

scanned. The choice of NO is motivated by the interest in examining nitric oxide formation pathways at different equivalence ratios and in testing the capability of the simplification algorithm to fully capture and reproduce such pathways.

Fig.2.3 shows the percentage of the original 325 reactions and 53 species that were found active in each simplified mechanism as a function of the tolerance threshold tol used. Comparison of iterative and non-iterative versions of the algorithm (the latter identifiable by subscript 1 in active species and reactions sets S and R) show that mechanisms resulting from the iterative algorithm are clearly more populated in terms of species and reactions for a given tolerance value as they also include processes that affect both CH_4 and NO indirectly. Note that species begin to be eliminated only after more than 50% of original reactions are dropped.

Table(2.2) is an example of a simplified mechanism obtained through the iterative algorithm for the stoichiometric and lean mixtures. Backward and forward reactions are individually listed in

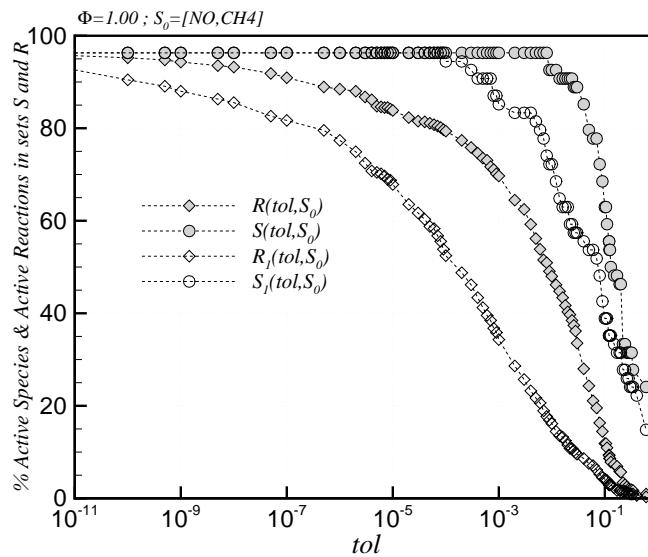


Fig. 2.3: Percentage of original reactions found active in simplified mechanisms $R(tol, S_0)$, $R_1(tol, S_0)$ and corresponding percentage of active species in $S(tol, S_0)$, $S_1(tol, S_0)$ as a function of the tolerance threshold tol imposed on importance indices.

order of decreasing importance index as found by the algorithm ⁴. Nitrogen chemistry is shown separately for the sake of clarity. The tolerance threshold was set at $tol = 0.15$ for the stoichiometric case and at $tol = 0.13$ for the lean case, this leading for both to 48 active reactions in the set $R(tol, [CH_4, NO])$. This level of simplification is possibly the highest achievable before errors on the equilibrium state become unacceptable.

In both cases fuel depletion basically occurs via the process $CH_4 \rightarrow CH_3 \rightarrow CO \rightarrow CO_2$ and a close analysis of more detailed mechanisms reveals that such process occurs through three pathways:

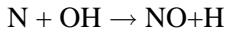
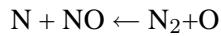
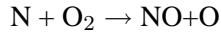
- (i) $CH_3 \rightarrow C_2H_6 \rightarrow \dots C_2H_x \dots \rightarrow CO \rightarrow CO_2$
- (ii) $CH_3 \rightarrow CH_2(S) \rightarrow CH_2 \rightarrow CH \rightarrow HCO \rightarrow CO \rightarrow CO_2$
- (iii) $CH_3 \rightarrow CH_2O \rightarrow HCO \rightarrow CO \rightarrow CO_2$

Fig.2.4 is a schematic of the methane breakup pathway to CO_2 for the stoichiometric simplified mechanism of Table(2.2) which clearly shows that only pathways (ii) and (iii) are fully retained while (i), shown as dotted arrows for the more detailed $tol = 0.9$ case, is cut off from the mechanism. Importance indices further reveal that path (iii) is the most responsible for CH_4 conversion to CO_2 this path being recognizable higher up in the reaction list. The latter statement holds its validity for the lean mechanism as well where path (ii) is dropped altogether leaving path (iii) as the main carrier of carbon from fuel to carbon dioxide for stoichiometric to lean mixtures.

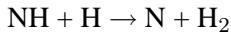
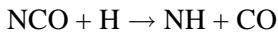
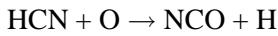
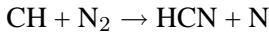
Following the same approach we can examine the nitric oxide (NO) formation pathways for both

⁴The ChemKin output contains both forward and backward reactions

the stoichiometric and lean case. The former case suggests unambiguously that thermal NO is the main production pathway through the following three elementary reactions

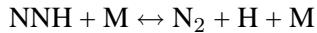
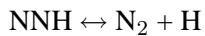
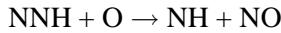


The prompt NO pathway is also recognizable, at least partially in the sequence $\text{HCN} \rightarrow \text{NCO} \rightarrow \text{NH} \rightarrow \text{N} \rightarrow \text{NO}$, through the reactions



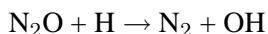
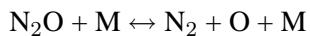
where the subsequent oxidation of N to NO occurs via reaction $\text{N} + \text{O}_2 \rightarrow \text{NO} + \text{O}$.

A third mechanism for NO formation can be singled out involving the oxidation of the NNH radical through the following path



The latter three reactions were included in the simplified mechanism by virtue of their fast importance indeces. This may suggest a role of such reactions limited to the evolution of the combustion process within the fast subspace of the phase space spanned by the trajectory within the explosion time frame.

Let us now examine NO formation for the lean combustion mechanism. With reference to Table(2.2) it appears that the thermal NO mechanism plays a lesser role in lean conditions, whereas a mechanism involving nitrous oxide (N_2O) as an intermediate species seems to be dominating via the following three reactions:



Prompt NO does not seem to play a role whereas the NNH path seems to be still active as in the stoichiometric case.

The above considerations stem from the mere analysis of the automatically generated simplified mechanisms and seem to be in good accordance with the literature [46, 47, 48, 49, 50, 51, 52, 53, 54]. This suggests that the iterative algorithm is capable of selecting important processes not just merely for the kernel species but also for the species directly and indirectly connected to them. The algorithm, therefore, can be thought of as working backwards, from the most relevant process regarding a kernel species to all the relevant processes connected to it. This way, pathways leading to the formation of a species such as the N₂O mechanism for NO, are entirely identified even when, as in such case, they involve pre-processes such as the formation of the intermediate N₂O from atmospheric nitrogen which does not directly involve NO.

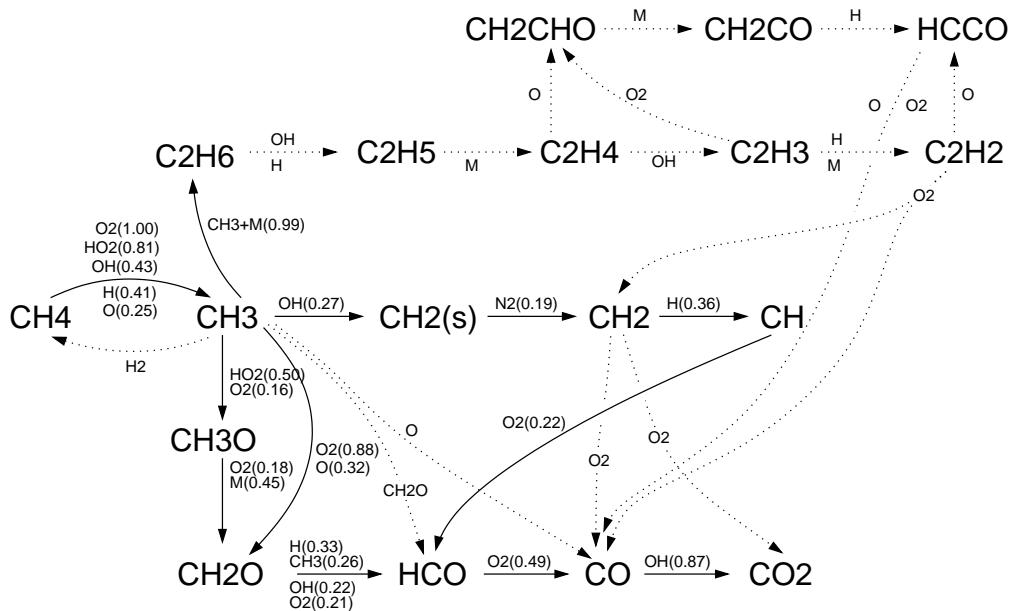


Fig. 2.4: Shown as continuous arrows is the CH₄ depletion pathway for the $\Phi = 1$ simplified mechanism $R(tol = 0.15, [\text{CH}_4, \text{NO}])$. Superimposed as dotted arrows is the same pathway for the $R(tol = 0.09, [\text{CH}_4, \text{NO}])$ which is clearly more complete.

Fig. 2.5 is a representation of the active species set $S_h(t_n, tol)$, upon convergence of the iterative algorithm, as it varies in time for the $\Phi = 1.0$ case with $S_0 = [\text{CH}_4, \text{NO}]$ and $tol = 0.024$. The tolerance level was kept lower in this case so as to have more populated active species sets and highlight their domain of relevance more clearly. Note that both kernel species CH₄ and NO are always active. The figure shows quite clearly how carbon radicals cease to be relevant (in the sense explained in the previous section) immediately after the maximum heat release peak whereas, starting from the same instant, nitrogen species begin to be included within the relevant species set, indicating the initiation of the NO formation process. The latter may be analyzed in Fig. 2.6 where a subset of the active reactions $R_h(t_n, tol)$ containing only nitrogen chemistry are displayed

$R(tol, [\text{CH}_4, \text{NO}])$	$R(tol, [\text{CH}_4, \text{NO}])$
tol=0.150 ; $\Phi = 1.0$	tol=0.130 ; $\Phi = 0.3$
1.00(O_2, s) $H_2O + CH_3 \leftarrow O_2 + CH_4$ $2CH_3(+M) \rightarrow C_2H_6(+M)$ $CH_3 + O_2 \rightarrow OH + CH_2O$ $OH + CO \rightarrow H + CO_2$ 0.87(CO_2, s) 0.81(H_2O_2, s) 0.67(O_2, s) 0.50(CH_3O, f) 0.49(HCO, f) 0.48(O_2, s) 0.47(H_2O, s) 0.45(H, f) 0.43(OH, f) 0.43(H_2O_2, s) 0.41(CH_4, s) 0.41(H_2, f) 0.36(CH, f) 0.33(CH_2O, s) 0.32(CH_2O, s) 0.31(O, f) 0.31(O, f) 0.27($CH_2(S), f$) 0.26(CO, s) 0.26(CO, s) 0.25(O, f) 0.24(HCO, f) 0.22(HCO, f) 0.22(CH, f) 0.21(O, f) 0.21(C_2H_6, f) 0.21(O, f) 0.19($CH_2(S), f$) 0.19(HCO, f) 0.18(CH_3O, f) 0.16(CH_3, f)	1.00(O_2, s) $HO_2 + CH_3 \leftarrow O_2 + CH_4$ $2CH_3(+M) \rightarrow C_2H_6(+M)$ $OH + CO \rightarrow H + CO_2$ $CH_3 + O_2 \rightarrow OH + CH_2O$ $OH + CH_4 \rightarrow CH_3 + H_2O$ $2OH(+M) \leftarrow H_2O_2(+M)$ $H + CH_4 \rightarrow CH_3 + H_2$ $OH + H_2 \rightarrow H + H_2O$ $CH + H_2 \leftarrow H + CH_2$ $H + CH_2O \rightarrow HCO + H_2$ $O + CH_3 \rightarrow H + CH_2O$ $2OH \rightarrow O + H_2O$ $2OH \leftarrow O + H_2O$ $OH + CH_3 \rightarrow CH_2(S) + H_2O$ $OH + CO \leftarrow H + CO_2$ $CH_3 + CH_2O \rightarrow HCO + CH_4$ $O + CH_4 \rightarrow OH + CH_3$ $HCO + H_2O \rightarrow H + CO + H_2O$ $OH + CH_2O \rightarrow HCO + H_2O$ $CH + O_2 \rightarrow O + HCO$ $O + H_2 \rightarrow H + OH$ $O_2 + CH_2O \rightarrow HO_2 + HCO$ $O + H_2 \leftarrow H + OH$ $CH_2(S) + N_2 \rightarrow CH_2 + N_2$ $HCO + M \leftarrow H + CO + M$ $CH_3O + O_2 \rightarrow HO_2 + CH_2O$ $CH_3 + O_2 \rightarrow O + CH_3O$
0.61(NO, s) 0.44(NH, f) 0.42(HCN, s) 0.41(N, f) 0.37(N, f) 0.34(NCO, s) 0.33(NNH, f) 0.32(NNH, f) 0.31(N_2, s) 0.27(N, f) 0.22(N, s) 0.17(NNH, f) 0.16(NNH, f) 0.15(H_2O, s)	0.92(N_2O, s) $N_2O(+M) \leftarrow N_2 + O(+M)$ $N_2O(+M) \rightarrow N_2 + O(+M)$ 0.50(H_2, s) 0.46(NH, f) 0.38(NNH, f) 0.37(NO_2, s) 0.37(NNH, f) 0.32(NO_2, s) 0.28(NO, s) 0.23(NH, f) 0.21(NO, s) 0.18(NH, f) 0.17(NO_2, s) 0.15(N_2O, s) 0.15(NO, s) 0.13(NNH, f)

Table 2.2: Example of simplified mechanisms $R(tol, [\text{CH}_4, \text{NO}])$ obtained for stoichiometric (left) and lean (right) combustion at atmospheric pressure. Reactions are listed in order of decreasing peak importance index as found by the iterative algorithm. Such peak indices are shown on the left of the relative reaction and, in brackets, the species they refer to together with the specification of whether it is a (s)low or (f)ast index. Tolerance levels were calibrated to yield the same number of active reactions in both cases. Nitrogen chemistry was regrouped at the bottom for clarity. The stoichiometric case highlights the importance of thermal NO and partially of prompt NO, while for the lean case the N_2O -intermediate mechanism is predominant.

as a function of time for the two cases $\Phi = 1.0$ and $\Phi = 0.3$. The different mechanisms for the formation of NO are visible, thermal and prompt NO appearing for the stoichiometric case and N_2O intermediate mechanism for the lean case.

Let us now examine the individual importance indices as a function of time. The reactions present in the simplified mechanism will give us a guideline as to which indices to examine. Supposing we wish to investigate the relative importance of the thermal and N_2O mechanisms involved in the formation of NO, we will recover the importance indices relative to NO. We will restrict our observations to slow indices (I_k^{NO})_{slow} as NO is never found to be a CSP radical to a fast mode, in accordance to Fig.2.5. Fig.2.7 displays such indices in the explosion region (where the number of

exhausted modes M is found to attain its minimum value). Forward and backward reactions indices are shown only if their sum is substantially different from zero thereby indicating a net production of nitric oxide. The stoichiometric and lean cases can be observed to differ extensively: the N_2O mechanism is extremely modest in the former case. On the contrary the N_2O mechanism (namely reaction $\text{N}_2\text{O} + \text{O} \rightarrow 2\text{NO}$) in the lean explosion can be observed to be considerably more influential than the thermal mechanism in affecting nitric oxide's source term g_{slow}^{NO} , i.e. the slow dynamics of NO. Further observation of the stoichiometric case reveals that the oxidation of atomic nitrogen is only briefly relevant and occurs within the heat release region only. Most of the subsequent NO is generated by the oxidation of molecular nitrogen in a far slower fashion through reaction $\text{N} + \text{NO} \leftarrow \text{N}_2 + \text{O}$. Finally, in a similar manner for both cases, a considerable influence of the reaction $\text{NNH} + \text{O} \rightarrow \text{NH} + \text{NO}$ is observed within the heat release region only.

Let us now examine the near-equilibrium region depicted in Fig.2.8. Such region is characterized by having a number of exhausted modes $M = 47$ uniformly throughout for both the stoichiometric and lean cases. The conservation laws for the $E = 5$ atomic species present in the system ($\text{C}, \text{H}, \text{O}, \text{N}, \text{Ar}$), however, constitute additional constraints between the molecular compounds, which can be thought of as generating a 'stoichiometric manifold', and which add to the M constraints represented by the M exhausted modes. An additional constraint is represented by the conservation of absolute (or mean) mixture enthalpy. The number of effectively active modes is thereby given by $N - (M + E + 1)$ where N is the overall number of unknowns including temperature. In the case of the quasi-equilibrium region we have $N - (M + E + 1) = 54 - (47 + 5 + 1) = 1$ indicating that only one mode is active. The asymptotic behavior of the system may therefore be viewed, within the phase space, by a trajectory evolving on a 1-dimensional manifold. The N -dim original system will reduce to a 1-dim system for the only slow species left while the remaining species will be determined by the $(M + E + 1)$ constraints as a function of the independently evolving slow species. Such slow species will be identified by the CSP radical pointer to the only non-exhausted mode, i.e. mode 48. Table(2.3) is a summary of CSP data extracted at a representative instant of the near-equilibrium region. The analysis of CSP radical pointer values shows that nitric oxide, as expected, was found to be the specie most affected by the only slow mode. A similar analysis of the participation indices to the slow mode reveals the reactions participating the most to such mode. Again, thermal NO reactions are predominant in the stoichiometric case while the N_2O mechanism aquires a non negligible role in the lean case. Also shown in Table(2.3) is the CSP equivalent (and un-physical) stoichiometry as it results from Eq.(2.5) after projection of the source term onto the CSP basis vectors, - which replaces the physical canonic form of Eq.(2.8) - where \mathbf{a}_{48} is the

equivalent CSP stoichiometric vector and f^{48} an equivalent rate.

Similar conclusions can be drawn from Fig.2.8 displaying the near-equilibrium importance indices.

The most evident aspect for the stoichiometric case is the non-equilibrium of reaction $\text{N} + \text{NO} \leftrightarrow \text{N}_2 + \text{O}$ which makes it the major rate limiting step for NO, as widely established in the literature. For the lean case similar conclusions may be drawn in addition to which the non-equilibrium of $\text{N}_2\text{O} + \text{O} \leftrightarrow 2\text{NO}$ is also observed, underlining a substantial role of the N_2O -mechanism - almost half of that of thermal-NO in this near-equilibrium region - in determining the rate of NO.

Φ	t_n	M	Active Modes	Driving Mode #	Slow species	CSP Equivalent stoichiometry	Contributing Reactions for driving mode
1.0	1.1077s	47	1	48	NO(0.88)	$0.02\text{H}+0.05\text{O}+0.38\text{O}_2 + 0.14\text{OH}+0.07\text{CO}_2+0.5\text{N}_2 \leftrightarrow 0.05\text{H}_2+0.03\text{H}_2\text{O} + 0.07\text{CO}+\text{NO}$	$(+0.185)\text{N}+\text{NO} \leftarrow \text{N}_2+\text{O}$ $(+0.086)\text{N}+\text{OH} \leftarrow \text{NO}+\text{H}$ $(+0.022)\text{N}+\text{O}_2 \rightarrow \text{NO}+\text{O}$ $(-0.117)\text{N}+\text{NO} \leftarrow \text{N}_2+\text{O}$ $(-0.074)\text{N}+\text{OH} \leftarrow \text{NO}+\text{H}$ $(-0.019)\text{N}+\text{O}_2 \leftarrow \text{NO}+\text{O}$
0.3	3649s	47	1	48	NO(0.77)	$0.5\text{O}_2+0.5\text{N}_2 \leftrightarrow \text{NO}$	$(+0.114)\text{N}+\text{NO} \leftarrow \text{N}_2+\text{O}$ $(+0.044)\text{N}_2\text{O}+\text{O} \rightarrow 2\text{NO}$ $(-0.113)\text{N}+\text{NO} \leftarrow \text{N}_2+\text{O}$ $(-0.043)\text{N}_2\text{O}+\text{O} \leftarrow 2\text{NO}$

Table 2.3: Summary of the main features of the asymptotic behaviour of methane combustion. Number in brackets in the slow species column is the value of the CSP pointer. Numbers in the last column are the values of the participation indices to the driving slow mode of the specified reaction

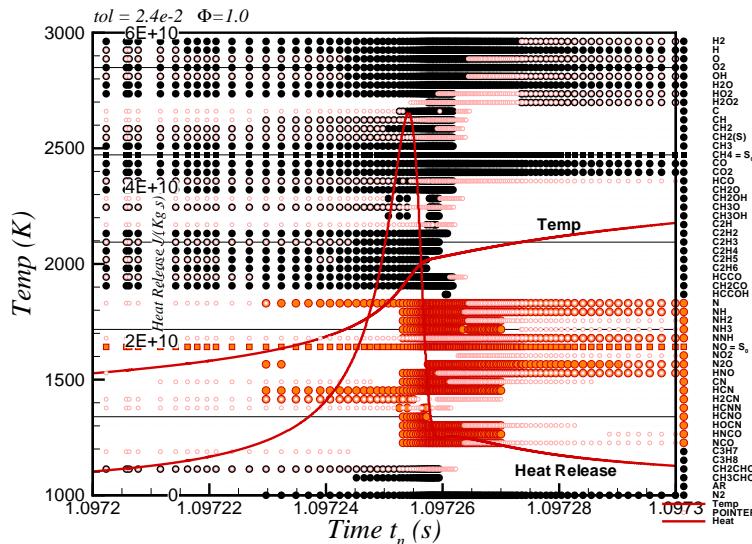


Fig. 2.5: Active species sets $S_h(t_n, tol)$ obtained upon convergence for the $\Phi = 1.0$ case with $S_0 = [\text{CH}_4, \text{NO}]$ and $tol = 0.024$. The time interval shown is 0.1ms corresponding to the maximum heat release, or explosion, region. Each filled dot indicates an active species. Also shown, with smaller empty dots, the pointers to CSP radicals.

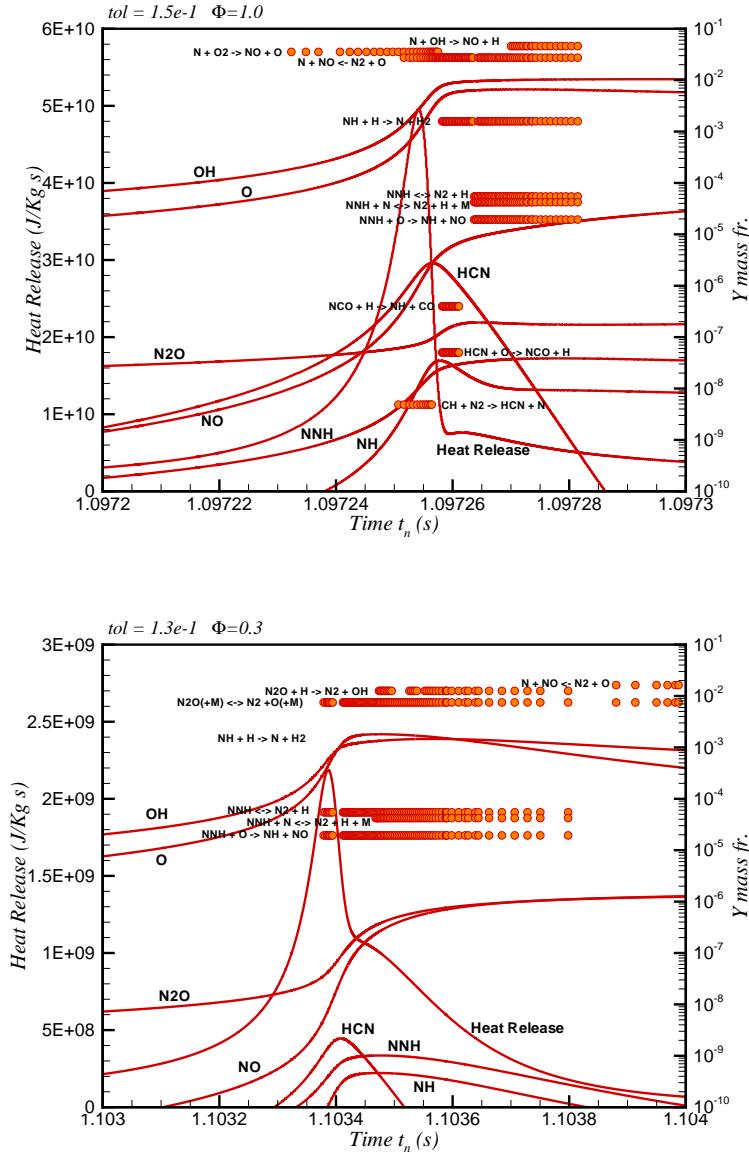


Fig. 2.6: Active reactions sets $R_h(t_n, tol)$ for each time record t_n obtained upon convergence for the $\Phi = 1.0$ case (top figure) (with $S_0 = [\text{CH}_4, \text{NO}]$ and $tol = 0.15$) and for the $\Phi = 0.3$ case (bottom figure) (with $S_0 = [\text{CH}_4, \text{NO}]$ and $tol = 0.13$). The explosion region is 0.1ms long for the $\Phi = 1.0$ case and 1.0ms for the $\Phi = 0.3$ case.

2.6.1 Comprehensive mechanism

As mentioned previously, simplified mechanisms with a higher degree of comprehensiveness [44] may be achieved through the iterative algorithm by accessing more than one CSP database at a time, each relative to particular initial conditions. The databases relative to the stoichiometric and lean cases examined previously may be therefore merged to yield simplified mechanisms having a wider range of applicability in terms of equivalence ratio. Table(2.4) is one of such comprehensive mechanisms originated by cross-referencing the $\Phi = 1.0$ and $\Phi = 0.3$ CSP databases. Note that

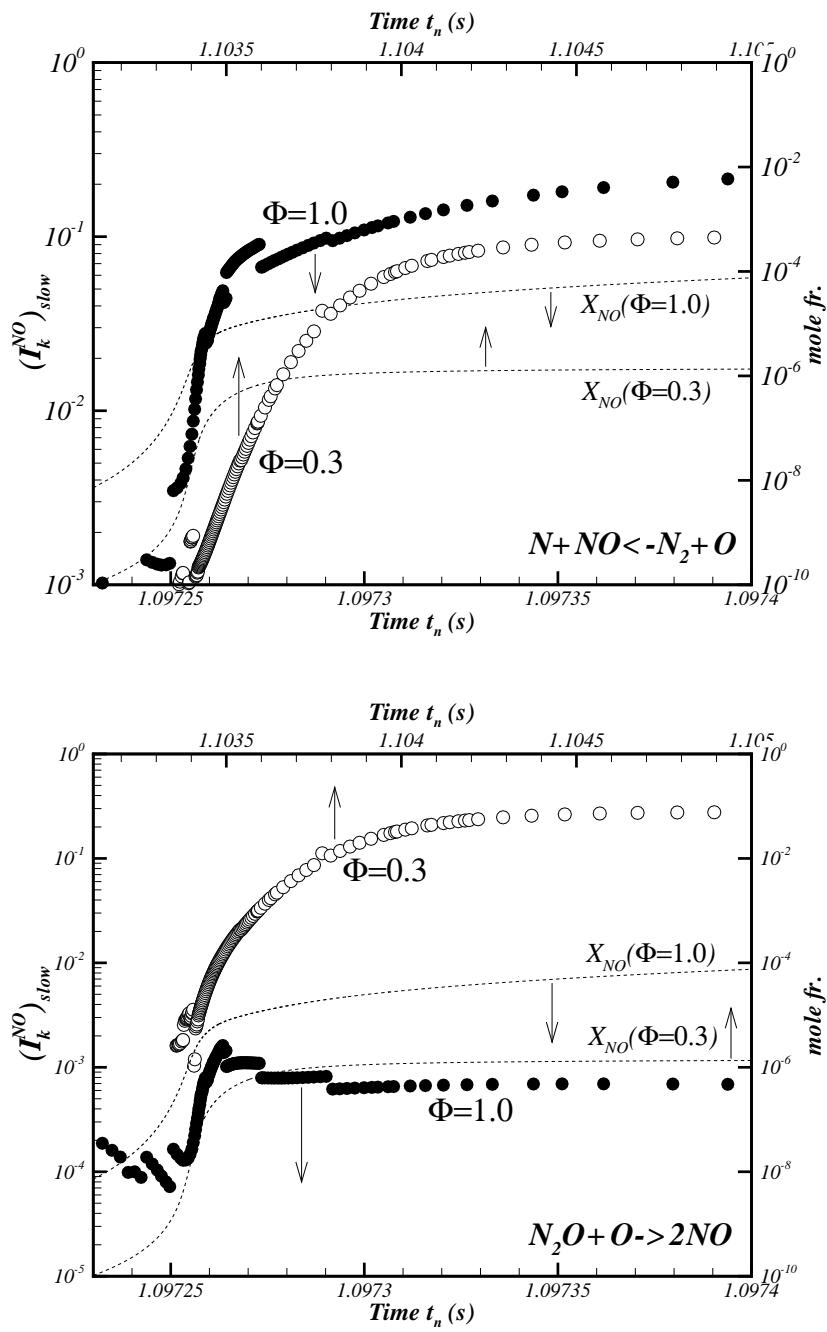


Fig. 2.7: Slow importance indices $(I_k^{NO})_{slow}$ for nitric oxide relative to reactions in thermal and N_2O nitric oxide mechanisms. The upper figure refers to the stoichiometric case, the lower to the lean case. Both refer to the explosion region.

pathways (i) and (ii) of the CH_4 breakup process to CO_2 are fully reproduced whereas for the NO formation both the thermal and N_2O mechanisms are included. The same technique may be utilized to produce valid mechanisms for given pressure ranges or initial temperatures.

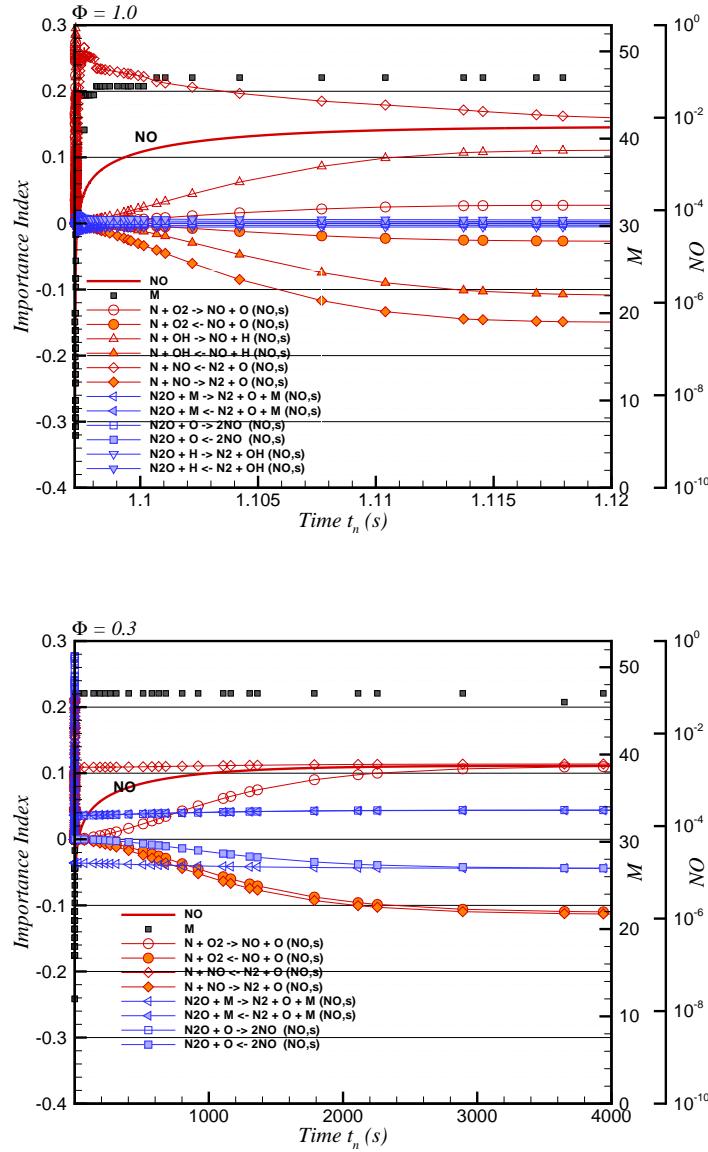


Fig. 2.8: Slow importance indices (I_k^{NO})_{slow} for nitric oxide relative to reactions in thermal and N_2O nitric oxide mechanisms. The upper figure refers to the stoichiometric case, the lower to the lean case. Both refer to the near-equilibrium region.

2.6.2 Accuracy assessment

The accuracy of each simplified mechanism is verified by re-integrating the set of ODE's with the same initial conditions, having replaced the complete (detailed) mechanism with the simplified ones. The resulting solutions can thus be compared with the detailed GRI Mech 3.0 reference solutions and an accuracy assessment be carried out. The $\Phi = 1.0$ and $\Phi = 0.3$ cases seen previously were considered and re-runs were performed using more than 80 simplified mechanisms for each case, with $S_0 = [\text{CH}_4, \text{NO}]$ and tolerance thresholds in the range $\text{tol} \in [10^{-11}, 6 \times 10^{-1}]$. The accuracy

$R(tol, [CH_4, NO])$	
1.00(O_2 , s)	$tol=0.19 \Phi = 0.3$ and $\Phi = 0.3$ databases
0.99(C_2H_6 , s)	$HO_2 + CH_3 \leftarrow O_2 + CH_4$
0.91(CO_2 , s)	$2CH_3(+M) \rightarrow C_2H_6(+M)$
0.91(CH_2O , s)	$OH + CO \rightarrow H + CO_2$
0.81(H_2O_2 , s)	$CH_3 + O_2 \rightarrow OH + CH_2O$
0.70(O_2 , s)	$CH_3 + H_2O_2 \leftarrow HO_2 + CH_4$
0.58(CH_4 , s)	$H + O_2 \rightarrow O + OH$
0.50(CH_3 , f)	$OH + CH_4 \rightarrow CH_3 + H_2O$
0.49(HCO , f)	$HO_2 + CH_3 \rightarrow OH + CH_3O$
0.49(H_2 , s)	$HCO + O_2 \rightarrow HO_2 + CO$
0.49(H , f)	$OH + H_2 \rightarrow H + H_2O$
0.48(O_2 , s)	$H + CH_2O(+M) \leftarrow CH_3O(+M)$
0.45(OH , f)	$H + O_2 \leftarrow O + OH$
0.43(H_2O_2 , s)	$2OH \leftrightarrow O + H_2O$
0.41(CH_4 , s)	$2OH(+M) \leftarrow H_2O_2(+M)$
0.41(H_2 , f)	$H + CH_4 \rightarrow CH_3 + H_2$
0.39(CH_4 , s)	$OH + H_2 \leftarrow H + H_2O$
0.38(CH_2O , s)	$O + CH_4 \rightarrow OH + CH_3$
0.38(CH_2O , s)	$O + CH_3 \rightarrow H + CH_2O$
0.37(CH , f)	$CH + H_2 \leftarrow H + CH_2$
0.30(CH_2O , s)	$H + CH_2O \rightarrow HCO + H_2$
0.32(HCO , f)	$OH + CH_2O \rightarrow HCO + H_2O$
0.27($CH_2(S)$, f)	$OH + CH_3 \rightarrow CH_2(S) + H_2O$
0.26(CH_3O , s)	$CH_3 + O_2 \rightarrow O + CH_3O$
0.26(CO , s)	$OH + CO \leftarrow H + CO_2$
0.26(CO , s)	$CH_3 + CH_2O \rightarrow HCO + CH_4$
0.24(HCO , f)	$HCO + H_2O \leftarrow H + CO + H_2O$
0.22(CH , f)	$CH + O_2 \rightarrow O + HCO$
0.21(O , f)	$O + H_2 \rightarrow H + OH$
0.21(C_2H_6 , s)	$O_2 + CH_2O \rightarrow HO_2 + HCO$
0.21(O , f)	$O + H_2 \leftarrow H + OH$
0.20(CH_3O , f)	$CH_3O + O_2 \rightarrow HO_2 + CH_2O$
0.19($CH_2(S)$, f)	$CH_2(S) + N_2 \rightarrow CH_2 + N_2$
0.92(N_2O , s)	$N_2O(+M) \leftarrow N_2 + O(+M)$
0.61(NO , s)	$N + O_2 \rightarrow NO + O$
0.50(N_2O , s)	$N_2O(+M) \rightarrow N_2 + O(+M)$
0.46(NH , f)	$NNH + O \rightarrow NH + NO$
0.42(HCN , s)	$HCN + O \rightarrow NCO + H$
0.41(N , f)	$N + NO \leftarrow N_2 + O$
0.38(NNH , f)	$NNH \leftarrow N_2 + H$
0.37(N , f)	$N + OH \rightarrow NO + H$
0.37(NNH , f)	$NNH \rightarrow N_2 + H$
0.34(NCO , s)	$NCO + H \rightarrow NH + CO$
0.30(N_2 , s)	$CH + N_2 \rightarrow HCN + N$
0.28(NO , s)	$N_2O + O \rightarrow 2NO$
0.26(N , f)	$N + OH \leftarrow NO + H$
0.23(NH , f)	$NH + O \rightarrow NO + H$

Table 2.4: Comprehensive simplified mechanism obtained by cross-referencing the CSP databases of the stoichiometric and lean cases.

assessment includes the percentage error on all equilibrium values (such as temperature and molar composition), the percentage error on ignition time and a r.m.s. of the instantaneous percentage error on molar composition (and of other instantaneous quantities such as temperature and heat release) computed along the reaction time. The r.m.s. error was defined by the following equation:

$$E(X) = 10^2 \left(\frac{1}{t_{reac}} \int_{t_5}^{t_{95}} \left(\frac{X_{det}(t) - X_{simpl}(t)}{X_{det}(t)} \right)^2 dt \right)^{\frac{1}{2}} \quad (2.27)$$

where $X_{det}(t)$ and $X_{simpl}(t)$ are the time dependent variables relative to the detailed and simplified mechanisms, respectively. Such variables can represent either the instantaneous values of the mole fraction of a generic species, the temperature or heat release.

A first consequence of the simplification is a shift in the ignition time noticeable between the detailed and the simplified mechanism solutions. Such ignition time shift is an unavoidable consequence of having dropped chain propagating or chain terminating reactions that can anticipate or postpone the onset of ignition. In Fig.2.9, showing the evolution of temperature and molar fraction for some

species of interest, a shift is only noticeable starting from mechanisms comprised of 37 species 81 reactions or less, and is maintained within a 20% difference - with respect to the complete GRI mech. - for mechanisms as small as 26 species 39 reactions. This can be verified from the summary of percentage errors in ignition time shown in Fig.2.11 where data is expressed as a function of the number of active reactions present in the simplified mechanisms. Also shown in the same figure are the errors on reaction time.

With this in mind, rms values obtained through Eq.(2.27) were computed only after having suitably eliminated the shift in ignition time. The criterion adopted to compensate for this temporal bias is to perform a time shift on the simplified solution by forcing the correspondence of its heat release peak with that of the detailed solution. Fig.2.10 portrays the shifted time evolution of nitric oxide mole fraction for the GRI full mechanism and for mechanisms ultimately yielding rms errors $E(X_{NO})$ up to 43%.]

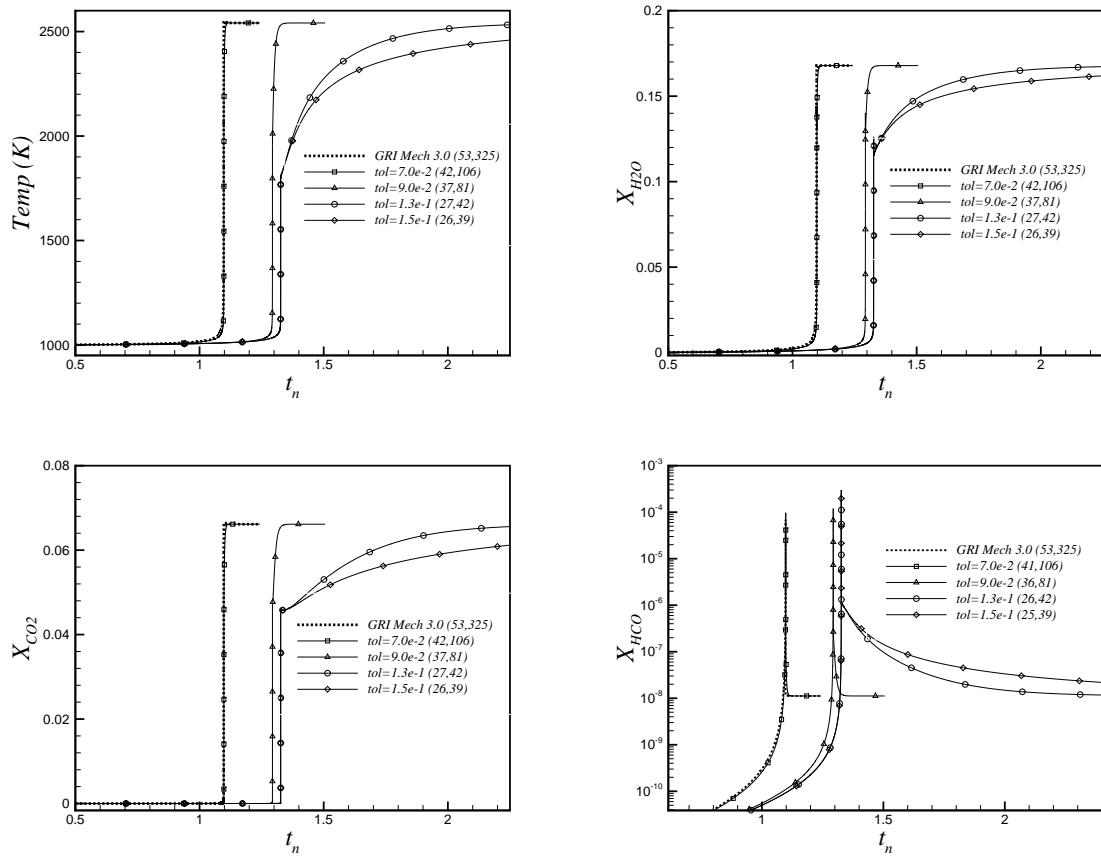


Fig. 2.9: Temperature and mole fractions for H_2O, CO_2, HCO for various simplified mechanisms (stoichiometric case). Mechanisms are indicated by their tolerance value. Shown in brackets are the number of active species and reactions in each mechanism

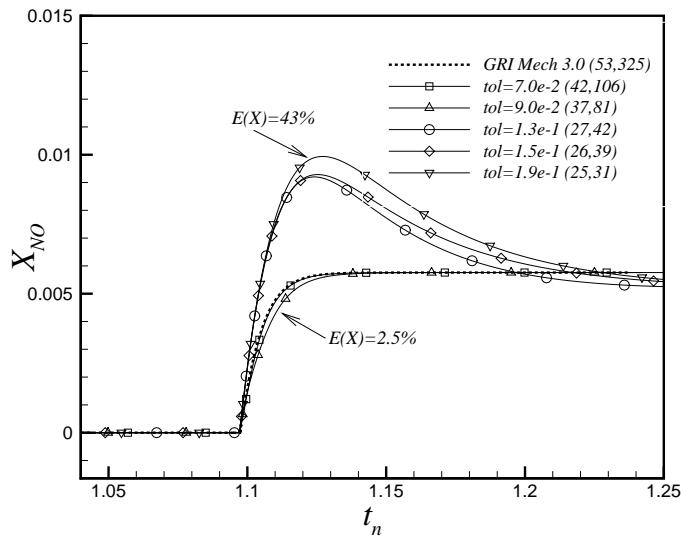


Fig. 2.10: Time shifted evolution of X_{NO} for mechanisms (stoichiometric case) yielding up to 42% $E(X_{NO})$ rms errors.

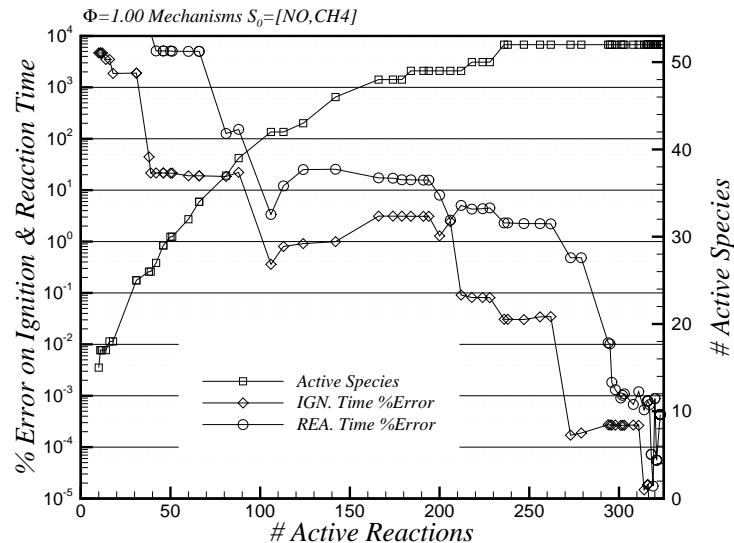


Fig. 2.11: Ignition and reaction times % errors wrt detailed mech. for stoichiometric mechanisms $R(tol, [CH_4, NO])$. Ignition and reaction times are defined as the time for the temperature to rise 5% and 95% of the peak temperature respectively.

Fig. 2.12 (top) shows the rms errors⁵ for nitric oxide, methane and water mole fractions as a function of the number of active reactions in each simplified mechanism. Such error obviously decreases for an increasing number of active reactions, the error range [1%, 40%] corresponding to active reactions

⁵Note that the error $E(X)$ fails to converge to zero as the number of active reactions tends to the GRI Mech. 3.0 value of 325. This is ultimately due to the accuracy of the simplified solution which, as we mentioned, is preliminarily shifted so as to compensate for the difference in ignition time, and such shift will be affected by a maximum accuracy of the order of the smallest time increment Δt used.

in the range [200, 31] (corresponding to a 62% to 10% of active reactions with respect to the GRI mech.) and active species in the range [49, 25] (corresponding to a 92% to 47% of active species). Also shown on the same figure is the speed-up factor obtained by comparing the complete GRI Mech 3.0, for which we define a CPU running time τ_c , with the simplified mechanisms, the speedup factor being estimated by dividing their running times by τ_c . Considering rms errors $E(X)$, for the 3 species shown, in the range [1%, 40%] to be acceptable, translates into speed-up factors in the range 1.6 to 6.4. On a far more conservative side, imposing the constraint $E(X) < 10\%$ on a wider pool of species such as CO₂, H₂O, O₂, HCO, and $E(X) < 1\%$ on temperature and on the kernel species NO, CH₄, would yield a speed-up factor of 2.4 with a 124 reactions 43 species simplified mechanism.

To conclude the accuracy assessment Fig.2.13 summarizes the percentage error of some equilibrium values found through the simplified mechs. with respect to the corresponding GRI Mech. 3.0 values. We observe that errors stay within the 0.01% limit for mechanisms with more than 39 active reactions and 25 active species or more. Less populated mechanisms fail to reproduce the equilibrium state essentially because the number of active reactions is less than the number of independent reactions⁶ for the given number of active species and the consequent elemental composition. Two main transitions in the equilibrium state can be essentially observed: the first is noted for mechanisms having less than 50 active species (< 218 active reactions and $tol < 0.012$) and occurs in correspondence to the removal of NO₂ from the mechanisms. The second transition occurs for mechanisms with less than 25 active species (< 31 active reactions and $tol < 0.2$) and was observed to occur when the following species were removed: H₂O₂, CH, CH₂, NH, NNH, HCN, NCO. Table(2.5) shows the contents of the removed species set $S_{rem}(tol, [\text{NO}, \text{CH}_4])$ as the (stoichiometric) mechanisms' tol increases. Also on Table(2.5) is the set $S_{rem}(tol, [\text{CO}_2, \text{CH}_4])$ relative to mechanisms for which NO was replaced by CO₂ as a kernel species. Note in this latter case the early elimination of several nitrogen compounds that were found to be non-relevant for this particular choice of starting kernel, where nitrogen is absent.

The top figure in Fig.2.14 is a log-log plot of the CPU speed-up factor as a function of the percentage of active reactions and species in mechanisms $R(tol, [\text{NO}, \text{CH}_4])$ (with respect to the 325-reactions 53-species full GRI mechanism). Straight lines represent fits of the kind $\log y = A \log x + B$. The dependency of the speed-up from the percentage of active species is found to be circa inverse

⁶Given a system of N_s species comprising E elements, we can construct an $E \times N_s$ formula matrix A where each integer element A_{ij}^i indicates the number of atoms of element i in the j -th species. The number of independent reactions \mathcal{R} is given by the dimension of the nullspace of A (i.e. $\dim(\text{Ker } A) = N_s - \text{Rank } A$) i.e. by the number of independent $N_s \times 1$ stoichiometric vectors \mathbf{S}_k that satisfy the elemental balance of reaction k i.e. $A\mathbf{S}_k = \mathbf{0}$.

quadratic as $A \approx -2.45$ ($y \sim 1/x^{2.45}$) while the dependency from active reactions is circa inverse with $A \approx -0.8$.

Rms errors $E(X)$ may also be viewed as functions of the speed-up factor which is a combined measure of the number of species and reactions present in a simplified mechanism and therefore a less ambiguous one. Fig.2.14 shows such relationships in a log-lin plot and can be viewed as errors per unit workload induced by the simplified mechanism. The fit in this case reveals a dependency $y \sim x^6$ of $E(X_{NO})$ from the speed-up, this meaning that a doubling of the speed-up is roughly achieved at a cost of a 64-fold increase in the NO rms errors.

tol	Active Spec.	Active Reac.	S_{rem}	tol	Active Spec.	Active Reac.	S_{rem}
$1.00e - 11$	52	323	AR	$1.00e - 11$	52	323	AR
$9.00e - 3$	50	228	C_3H_7	$9.00e - 3$	48	216	C_3H_7
			C_3H_8				C_3H_8
$1.40e - 2$	49	212	NO_2				H_2CN
$2.60e - 2$	48	179	$HCCOH$				NO_2
$4.00e - 2$	46	142	CH_2OH				CN
			CH_3OH				HCN
$5.00e - 2$	43	124	NH_3				$HCNN$
			$HCNO$				$HCNO$
			CH_3CHO				$HOCN$
$1.10e - 1$	32	60	C				$HNCO$
			C_2H				NCO
			C_2H_3				$HCCOH$
			C_2H_4				NH_2
			C_2H_5				NH_3
			NH_2				NNH
			H_2CN				CH_2OH
			$HCNN$				CH_3OH
			$HOCN$				NH
			$HNCO$				N_2O
			CH_2CHO				HNO
$1.22e - 1$	29	46	CH_2CO				C
			HNO				C_2H
			CN				C_2H_3
$1.50e - 1$	26	39	C_2H_2				C_2H_4
			$HCCO$				C_2H_5
			N_2O				CH_2CHO
$2.20e - 1$	18	18	H_2O_2				CH_3CHO
			CH				CH
			CH_2				C_2H_2
			$CH_2(S)$				$HCCO$
			NH				CH_2CO
			NNH				H_2O_2
			HCN				CH_2
			NCO				$CH_2(S)$
							N
							NO

Table 2.5: Left: $S_{rem}(tol, [\text{NO}, \text{CH}_4])$ set containing species progressively removed from simplified mechanisms as tol is increased. Right: $S_{rem}(tol, [\text{CO}_2, \text{CH}_4])$

2.6.3 Extraction of an optimal mechanism

An optimal mechanism is clearly one that, while guaranteeing an accurate equilibrium state, maximizes the speed-up factor whilst minimizing both the species mole fraction rms errors and the absolute error on ignition and reaction times. A simple procedure was implemented to extract optimal mechanisms each selected among those exhibiting overall average percentage errors within

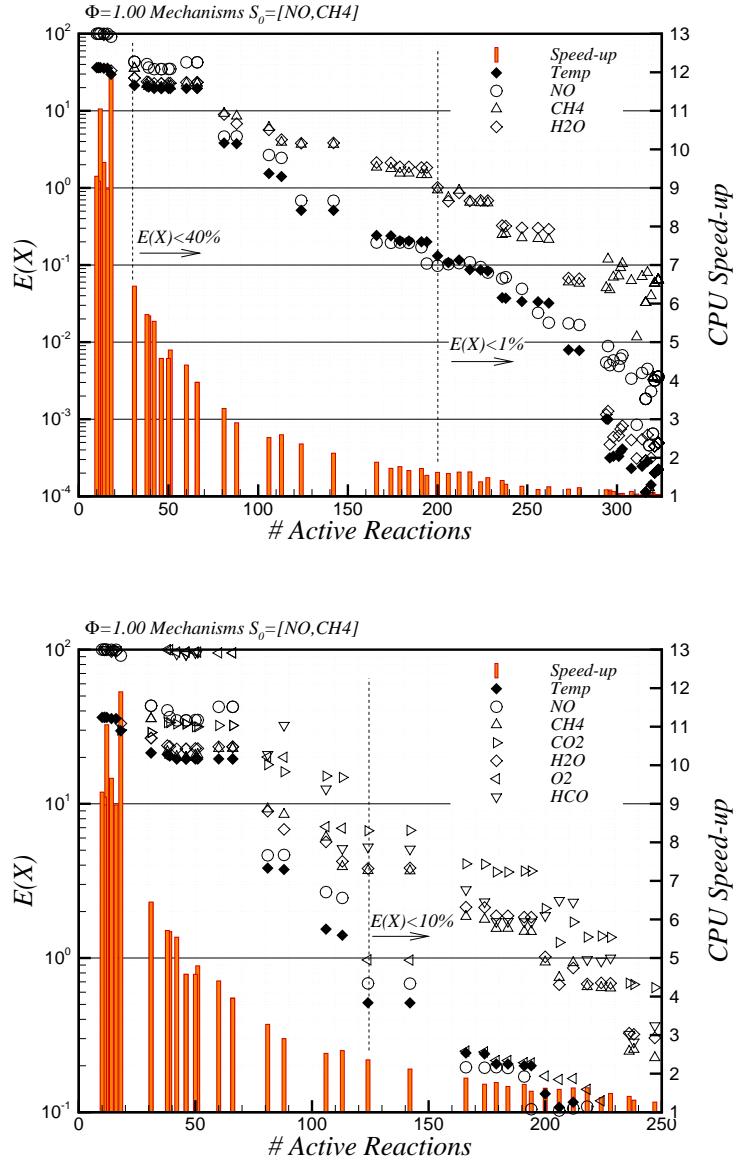


Fig. 2.12: Top: Rms % errors $E(X_{CH_4})$ and $E(X_{NO})$ relative to CH₄ and NO instantaneous mole fractions and $E(T)$ as a function of active reactions in the stoichiometric simplified mechanisms. Also shown is the speed-up factor of each mechanism. Bottom: similar analysis on a larger pool of species.

a prescribed range $[10^n\%, 10^{(n+1)}\%]$ with n an integer between -2 and 2 . Such overall error was defined as the equally weighed average of the ignition time error, the reaction time error and the average of the rms errors relative to the kernel species. Mechanisms exhibiting errors on equilibrium values greater than 1% were not considered in this analysis. After having filtered out those mechanisms not falling within a given error range, a cost function was constructed as the product of the overall average error and the speed-up factor, both suitably normalized. The optimal mechanism is clearly that which minimizes the cost function.

The search for optimal mechanisms was performed on two families of mechanisms having $S_0 =$

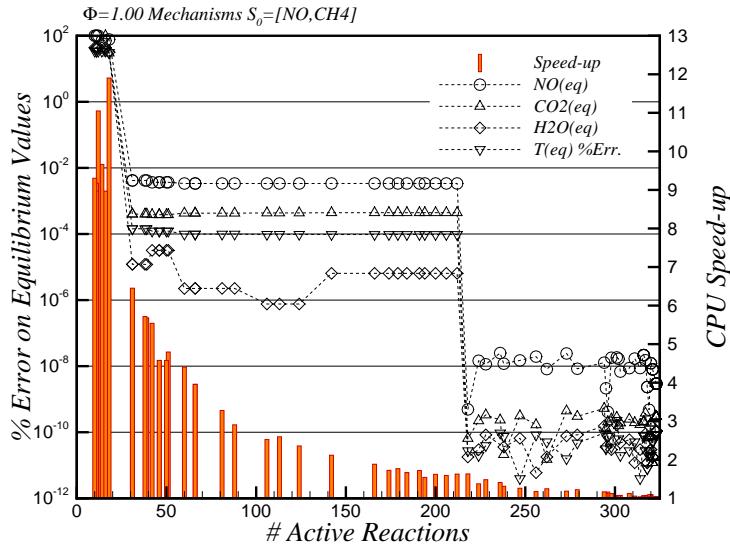


Fig. 2.13: Percentage error $10^2 |X_{simp} - X_{GRI3.0}| / X_{GRI3.0}$ in equilibrium values for temperature and $\text{NO}, \text{CO}_2, \text{H}_2\text{O}$ as a function of number of active reactions in simplified mechanism (stoichiometric case).

Opt. mech. data	[1, 10]	[10, 100]
S_0	[NO,CH ₄]	[NO,CH ₄]
tol	$1.2e - 2$	$9.0e - 2$
# Act. Species	50	37
# Act. Reactions	218	81
Speed-up	1.57	3.14
RMS % err. (CH ₄)	0.65	9.25
RMS % err. (NO)	0.11	4.63
RMS % err. (CO ₂)	1.36	17.93
Equil. Temp % err.	$2.75e - 11$	$9.89e - 5$
React. time % err.	4.27	126.53
Ignit. time % err.	$8.15e - 2$	18.46
S_0	[CO ₂ ,CH ₄]	[CO ₂ ,CH ₄]
tol	$1.6e - 2$	$4.0e - 2$
# Act. Species	39	32
# Act. Reactions	160	98
Speed-up	2.13	3.39
RMS % err. (CH ₄)	0.75	3.31
RMS % err. (NO)	1.09	4.52
RMS % err. (CO ₂)	1.23	6.52
Equil. Temp % err.	$9.90e - 5$	$1.54e - 4$
React. time % err.	2.61	24.12
Ignit. time % err.	2.45	0.94

Table 2.6: Performance summary for optimal mechanisms in the overall average error ranges [1%,10%] and [10%,100%].

[NO,CH₄] and $S_0 = [\text{CO}_2, \text{CH}_4]$ as kernel. For such two families Table.(2.6) reproduces the performance of the two optimal mechanisms in the overall error ranges [1%,10%] and [10%,100%]. The same results for additional error ranges are displayed in Fig.2.15. Note that the $R(tol, [\text{NO}, \text{CH}_4])$

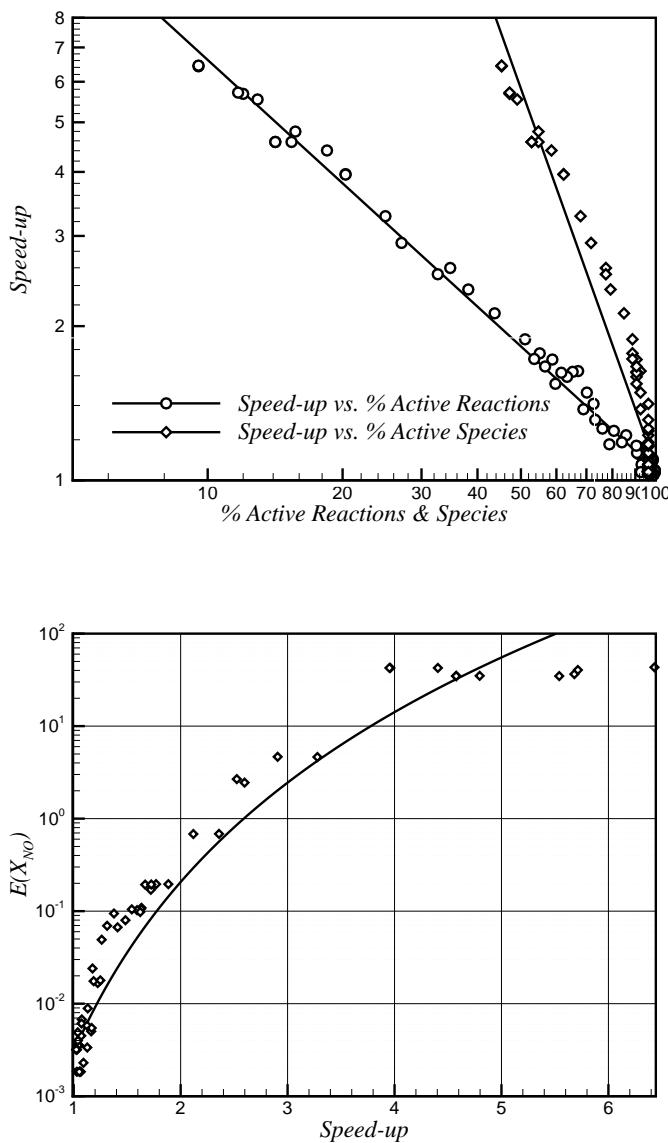


Fig. 2.14: Top: Log-Log plot of CPU speed-up as a function of the percentage of active reactions and species in simplified (stoichiometric) mechanisms. Bottom: rms error on NO mole fract. $E(X_{NO})$ as a function of the speed-up factor.

mechanisms are more populated in terms of active species as they contain more nitrogen chemistry than $R(tol, [CO_2, CH_4])$ mechanisms. This yields smaller speed-up factors for $R(tol, [NO, CH_4])$ mechanisms whilst exhibiting comparable accuracy with $R(tol, [CO_2, CH_4])$ mechanisms. The most promising result is possibly the 3.39 speed-up factor for the $R(tol = 0.04, [CO_2, CH_4])$ which exhibits rms errors of the order of 10%, an ignition time error of 1% and a 24% error on reaction time. Solutions for this optimal mechanism are shown in Fig. 2.16 after being shifted to compensate for the 1% in ignition time error.

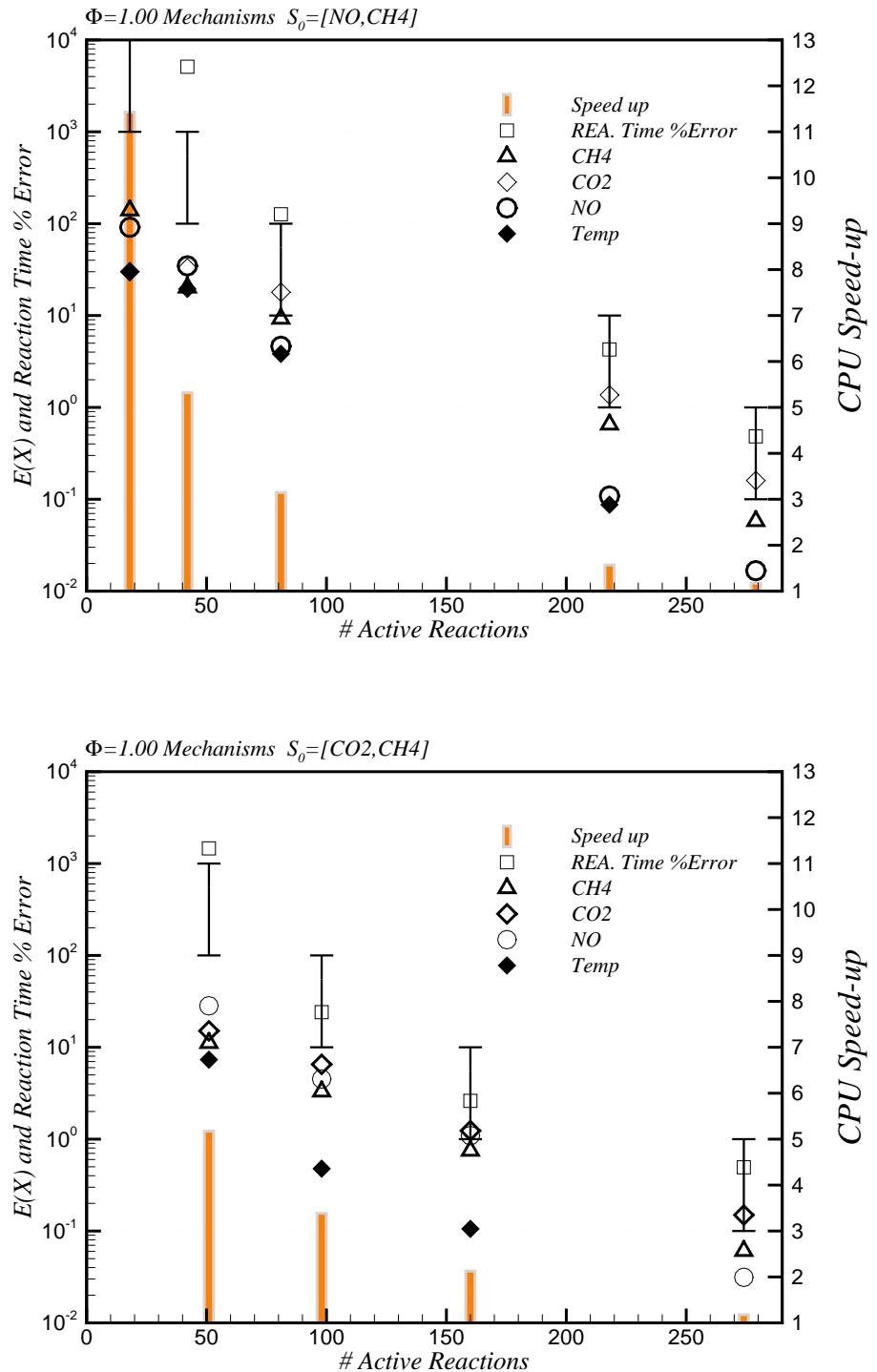


Fig. 2.15: rms % errors $E(X)$, % error on reaction times and speed-up factors for optimal mechanisms belonging to the $R[tol,NO,CH_4]$ and $R[tol,CO_2,CH_4]$ families. Also shown are the error ranges within which mechanisms are optimal.

2.7 Some intermediate conclusions

Let us now give some concluding remarks on the application of the CSP-based simplification algorithm to spatially homogeneous systems. The simplified mechanisms have been compared extensively to the detailed one and optimal mechanisms were extracted in appropriate error ranges.

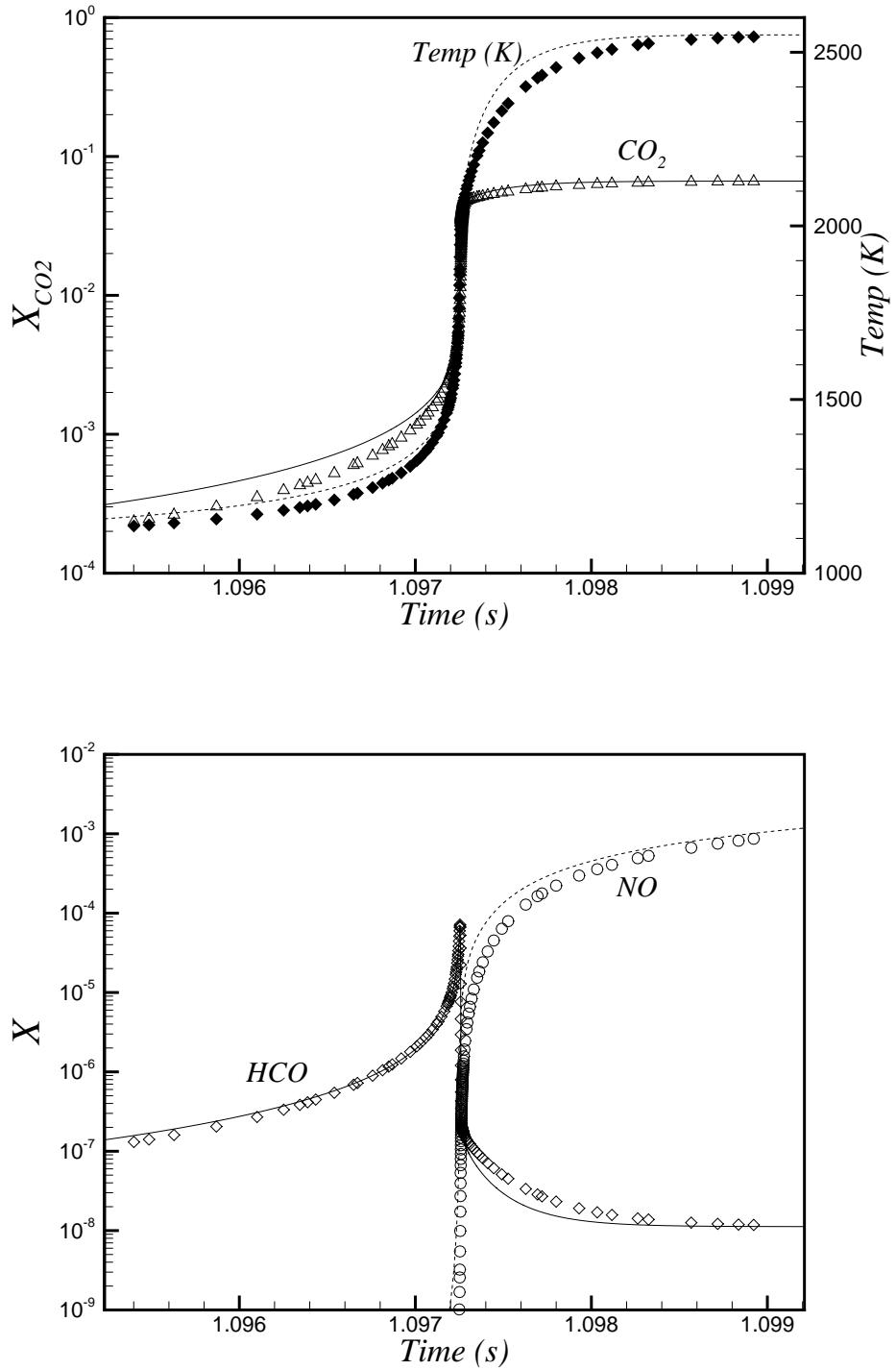


Fig. 2.16: Shifted solutions for the optimal mechanism $R(\text{tol} = 0.04, [\text{CO}_2, \text{CH}_4])$. Solid lines refer to GRI 3.0.

As a result, speed-up factors in CPU running times of up to 3.39 were observed for mechanisms exhibiting rms errors of the order of 10%.

The CSP database used in conjunction with the simplified systems has also proven to be an effective

diagnostic tool, providing ample insight into the kinetic system and allowing for the identification of kinetic sub-processes and pathways at various operative conditions of pressure, initial temperature and equivalence ratio. An analysis of NO-formation pathways was performed for lean and stoichiometric methane/air mixtures. The identification and assessment of the relative relevancy of various pathways such as thermal, N₂O, prompt and NNH was possible without an a-priori knowledge of the nitrogen chemistry.

The algorithm is highly automated, directly providing a Chemkin format output, this enabling a possible application of this approach in the context of adaptive kinetics methodologies where the simplified mechanisms are tailored to different regions of the spatio-temporal domain of integration of a reactive flow.

2.8 Extension of the CSP method to reactive flows.

The considerations of Chapter 1 and of the present Chapter are all based on a common ground: finite-dimensional spatially homogeneous systems governed by a set of ODE's. The presence of widely disparate intrinsic fast and slow timescales determine the existence of a slow invariant manifold (SIM) for the system, embedded within a finite-dimensional phase space. The knowledge of the SIM can both greatly facilitate the solution of the governing ODE's and provide significant physical understanding of the process at hand. Chemically reactive flows, on the other hand, are processes intrinsically non homogeneous in space, governed by systems of PDE's. Solutions can be viewed as trajectories in an adequate infinite dimensional functional space and thus the question arises as to if a slow manifold exists and how it can be represented. Recently, a number of studies appeared which discuss how and when the presence of diffusion can affect the structure of the slow manifold, either by considering its effects as a higher order correction in the finite-dimensional composition space [13, 55, 56, 57] or by identifying a manifold in an augmented space encompassing both compositions and mass fluxes [58, 59].

A mathematically rigorous way to treat PDE's in the context of fast-slow systems consists in retaining their nature of infinite dimensional systems, projecting them onto a suitable basis of orthogonal functions - enforcing the system's boundary conditions - so as to obtain a theoretically infinite set of ODE's. A tractable finite dimensional system, representing the original PDE, is clearly obtained after truncation. Any reduction method such as CSP or ILDM is then be applicable 'as is' to the ensuing finite dimensional system. It is to be noted that the resulting spectrum of timescales (obtainable through projection onto local eigenvectors, refined CSP vectors or ideal basis vectors as seen in Chapt. 1) describes an intrinsically coupled reactive-diffusive-convective process with no distinction between chemical and transport timescales, this being, after all, the process originally described by the PDE. We further note that any reactive-diffusive-convective manifold would be a geometric structure (of a high and finite-dimensional space) subjected to the problem's boundary conditions, as these are taken into account by the preliminary projection onto the basis of orthogonal functions. Because the original spatially inhomogeneous system is transformed into a homogeneous one described by a very large number of ODE's defined in a high and finite-dimensional space, the manifold is indeed a structure of the latter space. The latter approach may then be termed 'global' to indicate the independence of the manifold from the spatial coordinate. Further discussion on these issues can be found in [11].

Most of the efforts to define slow manifolds for reactive flows, mentioned earlier, are however based

on a 'local' approach, founded on the original PDE's, whereby the manifold, as in spatially homogeneous systems, is still a structure of the N -dimensional composition space, N being the number of species. The term 'local' is due to the fact that the manifold is not unique but rather a space-varying structure depending on the local spatial coordinate. The central ambiguity of such an approach is the choice of an appropriate set of N -dimensional basis vectors, related to the system's timescales and defining a slow/fast decomposition of the composition phase space. An immediate choice is to consider such basis vectors as the eigenvectors (possibly refined) of the local Jacobian matrix of the chemical source term alone. This approach, based on an analogy with spatially homogeneous systems, suffers from having neglected a possible influence of convective and diffusive timescales on the insurgence of a slow manifold defined in the composition space. Within the regions of physical space where fast transport phenomena balance fast chemistry, such an approach is expected to fail and the ensuing 'chemical' manifold to lose its applicability. In reactive-diffusive systems, such regions are characterized by high concentration gradients. In other regions, where diffusion is far slower than the fastest chemistry, a chemical manifold can indeed develop. Let us now briefly illustrate the local approach for a reactive-diffusive system involving N species and one physical space dimension x . The governing system of PDE's and the relative boundary conditions may be written as

$$\frac{d\mathbf{u}}{dt} = \mathbf{L}_d(\mathbf{u}) + \mathbf{f}(\mathbf{u}) ; \quad \mathbf{u}(0, x) = \mathbf{u}_0(x) \quad (2.28)$$

where $\mathbf{u}(t, x)$ is the N -dimensional composition vector, $\mathbf{f}(\mathbf{u})$ is the chemical source term, $\mathbf{L}_d(\mathbf{u}) = K\nabla^2\mathbf{u}$ is the diffusion operator with K an $N \times N$ diagonal matrix of diffusivities. The local approach defines a set of basis vectors $\{\mathbf{a}_i(\mathbf{u})\}_{i=1}^N$ and their dual $\{\mathbf{b}^i(\mathbf{u})\}_{i=1}^N$ as the local right and left eigenvectors of the Jacobian matrix of $\mathbf{f}(\mathbf{u})$. Projection on such basis yields:

$$\frac{d\mathbf{u}}{dt} = \mathbf{a}_1 h_1 + \cdots + \mathbf{a}_N h_N \quad (2.29)$$

$$h_i(t, x) = \mathbf{b}^i [\mathbf{L}_d + \mathbf{f}] . \quad (2.30)$$

where the usual CSP ordering of modes, from fastest to slowest, is enforced on the basis of chemical timescales only. In the presence of M fast chemical timescales, by analogy with spatially homogeneous system, we may define the slow manifold as the following set of relations:

$$h_i(t, x) = \mathbf{b}^i [\mathbf{L}_d + \mathbf{f}] = 0 \quad i = 1, \dots, M. \quad (2.31)$$

Singh et. al. [56] refer to such manifold as the Approximate Slow Invariant Manifold (ASIM) and

a similar definition can be found in [11]. The ASIM differs from the ILDM, expressed simply by $\mathbf{b}^i \mathbf{f} = 0$, $i = 1, \dots, M$, for the presence of the diffusion operator. Both [56] and [11] illustrate the accuracy of the ILDM and ASIM. Eq. (2.31) for the ASIM was justified in [56] as follows. Introducing a new variable $\mathbf{z} = B\mathbf{u}$ where $B = (\mathbf{b}^1, \dots, \mathbf{b}^N)^T$, so that $\mathbf{u} = Az$, with $A = (\mathbf{a}_1, \dots, \mathbf{a}_N)$, and expressing $\mathbf{f} = J\mathbf{u} + \mathbf{g}$, \mathbf{g} being the nonlinear part of \mathbf{f} , Eq.(2.28) becomes

$$\frac{1}{\lambda_i} \left(\frac{dz^i}{dt} + \mathbf{b}^i \sum_{j=1}^N z^j \frac{d\mathbf{a}_i}{dt} \right) = z^i + \frac{1}{\lambda_i} (\mathbf{b}^i [\mathbf{L}_d + \mathbf{g}]), \quad (2.32)$$

where λ_i are the eigenvalues of J , the Jacobian of \mathbf{f} . Eq. (2.32) is cast in a singularly perturbed form with $1/|\lambda_1|, \dots, 1/|\lambda_M|$, which are a measure of the local chemical timescales, as small parameters. Only in the case where the diffusive timescales, acting in the directions $\{\mathbf{a}_i\}$, $i = 1, \dots, M$, are slower than the corresponding chemical timescales, then Eq. (2.32) admits a slow manifold which can be obtained by neglecting the right hand side in Eq. (2.32) for $i = 1, \dots, M$:

$$z^i + \frac{1}{\lambda_i} (\mathbf{b}^i [\mathbf{L}_d + \mathbf{g}]) = 0 \quad i = 1, \dots, M. \quad (2.33)$$

Substituting in the above $z^i = \mathbf{b}^i \mathbf{u}$ and $\mathbf{g} = \mathbf{f} - J\mathbf{u}$ the expression of Eq.(2.31) is then recovered. The limit of applicability of Eq.2.31 as a reactive-diffusive manifold is therefore that diffusive timescales be slower than reactive timescales along the fast directions, defined by chemistry alone, $\{\mathbf{a}_i\}$, $i = 1, \dots, M$. If on the other hand this fails to happen and fast diffusive processes balance fast reactive processes, the system fails to be a fast-slow system and cannot be cast in singularly perturbed form, this negating the existence of a manifold altogether.

2.9 Applications to premixed laminar flames

In this section we will extend the local approach just described and the notions of model simplification based on the CSP analysis of a solution to the case where such a solution is represented by an quasi-one-dimensional flame propagation in a premixed gaseous mixture under a low Mach number conditions. For a CSP analysis to be carried out as described in the preceding sections a new canonic form of the governing system of equations will be needed to account for the spatial diffusive and convective operators.

The system of PDE's governing a generic reacting flow is composed of mass, momentum, energy and species conservation equations and may be represented formally by the following system:

$$\frac{d\mathbf{u}}{dt} = \mathbf{L}_c + \mathbf{L}_d + \mathbf{f}(\mathbf{u}) = \mathbf{g}(\mathbf{u}) \quad (2.34)$$

where \mathbf{L}_c and \mathbf{L}_d formally denote the convective and diffusion spatial operator respectively, whereas $\mathbf{f}(\mathbf{u})$ is a source term describing the production and consumption of species by finite rate chemical kinetics. Vector $\mathbf{u}(t)$ represents the vector and scalar unknowns such as density, velocity, species mass fractions and temperature.

Before illustrating the CSP canonic form for the governing equations a brief outline of the general procedure followed to perform the CSP analysis will be given. A steady isobaric quasi-one-dimensional solution for the premixed flame is first obtained using the PREMIX software package. The solution is then used as the initial condition for a reactive unsteady low Mach number⁷ [43, 60] solver capable of supplying a CSP output. In the low-Mach assumption, acoustic wave propagation is ignored, and the pressure field is decomposed into a spatially uniform component $P_0(t)$ (time-independent in open domains) and a hydrodynamic component $p(x, t)$. The relative non-dimensional 1-D governing equations in the low Mach, zero bulk viscosity and negligible body forces read:

$$\frac{\partial \rho}{\partial t} + \frac{\partial(\rho u)}{\partial x} = 0 \quad (2.35)$$

$$\frac{\partial(\rho u)}{\partial t} + \frac{\partial(\rho u^2)}{\partial x} = -\frac{\partial p}{\partial x} + \frac{1}{Re} \left(2 + \frac{\lambda}{\mu} \right) \frac{\partial^2 u}{\partial x^2} \quad (2.36)$$

$$\frac{\partial Y_i}{\partial t} = -u \frac{\partial Y_i}{\partial x} + \frac{1}{Re Sc} \frac{1}{\rho} \frac{\partial}{\partial x} (\rho V_i Y_i) + \frac{Da}{\rho} \dot{\omega}_i \quad (2.37)$$

$$\frac{\partial T}{\partial t} = -u \frac{\partial T}{\partial x} + \frac{1}{Re Pr} \frac{1}{\rho c_p} \frac{\partial}{\partial x} \left(k \frac{\partial T}{\partial x} \right) - \frac{1}{Re Sc} \frac{1}{c_p} \sum_{i=1}^{N_s} c_{pi} V_i Y_i \frac{\partial T}{\partial x} - \frac{Da}{\rho c_p} \sum_{i=1}^{N_s} W_i h_i \dot{\omega}_i \quad (2.38)$$

where ρ is the density, u the axial velocity, T the temperature, p the pressure (all non-dimensional)

⁷For a non reactive single species fwb the low Mach number ($M < 0.3$) assumption implies a negligible dependency of density from both temperature and pressure (incompressibility). Also, the velocity is high enough to neglect buoyancy effects ($Fr \rightarrow \infty$) (not so in natural convection). Moreover the Eckert number (measuring the relative importance of temperature rise due to compression to the overall temperature rise) can be assumed to be small so that both the dissipation and compression terms may be neglected in the energy equation. The non-reactive low Mach equations are generally written in non-dimensional form as:

$$\begin{aligned} \nabla \cdot \mathbf{V} &= 0 \\ \frac{D\mathbf{V}}{Dt} + \nabla p &= \frac{1}{Re} \nabla^2 \mathbf{V} \\ \frac{DT}{Dt} &= \frac{1}{Pr Re} \nabla^2 T \end{aligned}$$

In a reactive low Mach number fwb, the variation in density is therefore only due to the temperature rise in the reaction zone, and incompressibility does not hold.

and Y_i the i -th species mass fraction. Re , Da , Pr and Sc are the Reynolds, Damkohler, Prandtl and Schmidt numbers respectively. In the above equations $\dot{\omega}_i$ is the molar production rate of the i -th species where $\dot{\omega}_i = \sum_{i=1}^{N_r} \nu_{ik} (r_{f,i} - r_{b,i})$ being ν_{ik} the stoichiometric coefficient of species i (of the N_s present) in the k -th of the N_r reversible reactions and $r_{f,i}, r_{b,i}$ are the forward and backward reaction rates; W_i and \bar{W} are respectively the molecular weight of the i -th species and the mean molecular weight of the mixture; c_{pi} and c_p are respectively the constant pressure heat capacity of the i -th species and of the mixture while h_i the specific enthalpy of the i -th species. k is the thermal conductivity of the mixture, μ is the dynamic viscosity, λ the second viscosity coefficient, and V_i is the diffusion velocity of the i -th species which can be approximated using Fick's law $V_i = -D_{iN_s} Y_i dY_i / dx$ and where it was assumed that the N_s -th specie, namely molecular nitrogen (N_2), is dominant and D_{iN_s} is the binary mass diffusion coefficient of species i into the N_s -th species. V_{N_s} is found from the identity $\sum_{i=1}^{N_s} Y_i V_i = 0$. Note that in the species conservation Eq. (2.37) $i = 1, \dots, N_s - 1$ where Y_{N_s} is found from $\sum_{i=1}^{N_s} Y_i = 1$. The perfect gas state equation is: $P_0 = \rho T / \bar{W}$ where P_0 is assumed constant in the energy Eq. (2.38). Soret and Dufour effects, radiant heat transfer, and soot formation are neglected.

For the CSP analysis, focused only on the scalar conservation equations, we will formulate the PDE model of Eq.(2.34) assuming the N -dimensional unknowns vector to be $\mathbf{u} = (Y_1, \dots, Y_{N_s-1}, T)^T$ and the spatial convection and diffusion operators to be respectively:

$$\mathbf{L}_c = \begin{cases} -u \frac{\partial Y_i}{\partial x}, & \text{if } i = 1, 2, \dots, N-1; \\ -u \frac{\partial T}{\partial x}, & \text{if } i = N; \end{cases} \quad (2.39)$$

and,

$$\mathbf{L}_d = \begin{cases} \frac{1}{Re Sc} \frac{1}{\rho} \frac{\partial}{\partial x} (\rho V_i Y_i), & \text{if } i = 1, 2, \dots, N-1; \\ \frac{1}{Re Pr} \frac{1}{c_p} \frac{\partial}{\partial x} (k \frac{\partial T}{\partial x}), & \text{if } i = N. \end{cases} \quad (2.40)$$

and the chemical source term to be:

$$\mathbf{f}(\mathbf{u}) = \begin{bmatrix} \frac{Da}{\rho} \dot{\omega}_1 \\ \frac{Da}{\rho} \dot{\omega}_2 \\ \vdots \\ \frac{Da}{\rho} \dot{\omega}_{N_s-1} \\ -\frac{Da}{\rho c_p} \sum_{i=1}^{N_s} W_i h_i \dot{\omega}_i \end{bmatrix} \quad (2.41)$$

For the CSP indices to be defined and the CSP analysis carried out, Eq.(2.34) must be recast in a

CSP canonic form of the kind

$$\frac{d\mathbf{u}}{dt} = S\mathbf{R} = \mathbf{g}. \quad (2.42)$$

To this end we define the $(2N_r + 4)$ -dimensional vector \mathbf{R} as

$$\mathbf{R} = [r_{f,1}, r_{f,2}, \dots, r_{f,N_r}, 1, 1, -r_{r,1}, -r_{r,2}, \dots, -r_{r,N_r}, 0, 0]^T \quad (2.43)$$

and the (N_s, N_r) matrix of stoichiometric coefficients $\tilde{S} = \nu_{ik}$ as in Eq.(2.10). Recalling Eq.(2.11) we further define the (N, N_s) matrix Q

$$Q = \begin{bmatrix} \frac{Da W_1}{\rho} & 0 & \dots & 0 & 0 \\ 0 & \frac{Da W_1}{\rho} & \dots & 0 & 0 \\ \dots & 0 & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & \frac{Da W_{N_s}}{\rho} \\ -\frac{Da W_1}{\rho c_p} h_1 & \dots & \dots & -\frac{Da W_{N_s-1}}{\rho c_p} h_{N_s-1} & -\frac{Da W_{N_s}}{\rho c_p} h_{N_s} \end{bmatrix}$$

so that the $(N, 2N_r + 4)$ -dimensional generalized stoichiometric matrix S takes the form

$$S = [Q\tilde{S}, \mathbf{L}_c, \mathbf{L}_d, Q\tilde{S}, \mathbf{L}_c, \mathbf{L}_d] = \quad (2.44)$$

$$= [\mathbf{S}_1, \dots, \mathbf{S}_{N_r}, \mathbf{L}_c, \mathbf{L}_d, \mathbf{S}_1, \dots, \mathbf{S}_{N_r}, \mathbf{L}_c, \mathbf{L}_d]. \quad (2.45)$$

where the stoichiometric vectors \mathbf{S}_i are those of Eq.(2.9). Note that $S\mathbf{R} = \mathbf{L}_c + \mathbf{L}_d + \mathbf{f}(\mathbf{u}) = \mathbf{g}(\mathbf{u})$ this confirming that the canonical form Eq.(2.42) coincides with Eq.(2.34).

The CSP theoretical formulation is similar to the one illustrated at the beginning of the chapter. The comprehensive chemical-convective-diffusive source term can be projected onto the CSP basis as

$$\frac{d\mathbf{u}}{dt} = \mathbf{a}_1 h^1 + \dots + \mathbf{a}_N h^N = \sum_{r=1}^M \mathbf{a}_r h^r + \sum_{s=M+1}^N \mathbf{a}_s h^s \quad (2.46)$$

where the projections are

$$h^i = \mathbf{b}^i \cdot \mathbf{g} = \mathbf{b}^i \cdot [\mathbf{L}_c + \mathbf{L}_d + \mathbf{f}(\mathbf{u})] \quad i = 1, \dots, N \quad (2.47)$$

The canonical form just described can be used to recast Eq.(2.3) and Eq.(2.46) (the latter constrained on the slow manifold, i.e. placing $h^r \approx 0$) obtaining identical expressions to Eq.(2.13) and

Eq.(2.14) respectively. This allows us to still adopt the importance index definitions of Eq.(2.24) and Eq.(2.25). The only visible difference is in the total number of processes which account for convection and diffusion of each species as well as the N_r reversible reactions. So by considering Eq.(2.24) with $k = N_r + 1$ or $k = N_r + 2$, we can define the slow importance index respectively of convection or diffusion on the i -th species or on temperature:

$$(I_k^i)_{slow} = \sum_{s=M+1}^N a_s^i (\mathbf{b}^s \cdot \mathbf{L}_{c,d}) / \sum_{j=1}^{2N_r+4} \left| \sum_{s=M+1}^N a_s^i (\mathbf{b}^s \cdot \mathbf{S}_j) R_j^j \right| \quad (2.48)$$

A similar definition clearly applies for fast indices. Note, however, that the present formulation holds its validity only within the limits of applicability of the 'local' approach as discussed in section 2.8. The phase space fast/slow decomposition performed by means of 'chemical basis vectors' (e.g. eigenvectors of the Jacobian of the chemical source term) is indeed only feasible when diffusive and convective timescales acting along the fast directions are far slower than the chemical timescales, this being the only circumstance where a singularly perturbed form of the equations is meaningful and a manifold does indeed develop.

2.9.1 Model simplification for premixed laminar flames.

The 1-D premixed laminar flame problem was run using the low Mach number model with the 177 reactions and 32 species GRI Mech 1.2 [61] model for methane combustion in air. Solutions were obtained for several equivalence ratios Φ ranging from mixtures as lean as $\Phi = 0.7$ to as rich as $\Phi = 1.3$. A CSP database composed of importance indices and radical pointers was then constructed on the basis of the obtained solutions and the model simplification was performed in a similar way to that described for the perfectly stirred reactor.

The simplification algorithm was slightly modified with respect to the 'standard' version illustrated in Table 2.5. It was previously noticed that highly simplified mechanisms may have a number of active reactions which comes short of the number of independent reactions for the given number of active species and elements present. To avoid this, all the reactions that were not originally included in the global simplified mechanisms by virtue of their low importance indices, but which do however involve active species, are recovered and appended to the global simplified mechanism as shown in Table 2.9.1. The end result is a more populated global simplified mechanism in terms of number of reactions while the number of active species is left unchanged. This latter circumstance implies that the cost of the modification, in terms of computer time, is limited because, as we have seen, the number of species is a far more discriminating factor than the number of reactions.

```

For all  $t_n, n \in [1, N_{sol}]$ 
  { Start with kernel set of active species :  $S_{[h=0]} = S_{[0]} = S_{user}$ ;}
  While  $[S_{[h]}(t_n, tol, S_{user}) = S_{[h-1]}(t_n, tol, S_{user})]$ 
    {  $h = h + 1$ ;}
    Define new active reactions set :
       $R_h(t_n, tol, S_{user}) = \{k : (I_k^i(t_n))_{slow} > tol; i \in S_{[h-1]}(t_n, tol, S_{user})\} \cup$ 
        {  $k : (I_k^i(t_n))_{fast} > tol; i \in S_{[h-1]}^{rad}(t_n, tol, S_{user})\}$ };
    Define new active species set :
       $S_{[h]}(t_n, tol, S_{user}) = \{i : \nu_{ik} \neq 0; k \in R_h(t_n, tol, S_{user})\};$ 
    };
  };
Define global simplified mechanism  $R(tol, S_{user})$  as in Eq.(2.26)
Active species in global simplified mechanism:
 $S(tol, S_{user}) = \{i : \nu_{ik} \neq 0; k \in R(tol, S_{user})\}$ 
Recovered reactions:
 $R_{rec}(tol, S_{user}) = \{k \notin R(tol, S_{user}) : \nu_{ik} \neq 0; i \in S(tol, S_{user})\}$ 
Append Recovered reactions:
 $R(tol, S_{user}) = R(tol, S_{user}) \cup R_{rec}(tol, S_{user})$ 

```

Table 2.7: Modified simplification algorithm

The effect of the algorithm modification on the number of reactions is visible in Fig.(2.17), which refers to a parametric simplification campaign, with $tol \in [10^{-6}, 1]$, relative to a CSP database extracted from a stoichiometric flame ($\Phi = 1.0$) solution with $S_{user} = [\text{CH}_4, \text{CO}_2]$. The figure displays the number of reactions present in each of the simplified mechanisms, with and without the algorithm modification, as a function of the tolerance value used, the number of active species being the same in both cases. Note how the modification reintroduces a number of reactions that had originally been discarded.

2.9.2 Results

As mentioned earlier, several premixed flame solutions were computed using GRI Mech. 1.2, each corresponding to a pressure of $p = 1\text{Atm}$ and equivalence ratios $\Phi = 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3$. For each solution, a CSP database was constructed composed of importance indices and CSP radical pointers. The modified simplification algorithm was then used to generate a great number of simplified mechanisms, each corresponding to a different value of the tolerance threshold tol . The kernel species set was chosen as $S_{user} = [\text{CH}_4, \text{CO}_2]$, so as to include a major reactant and a major product. Comprehensive mechanisms were also generated by the modified algorithm, this being achieved by using all the available CSP databases simultaneously.

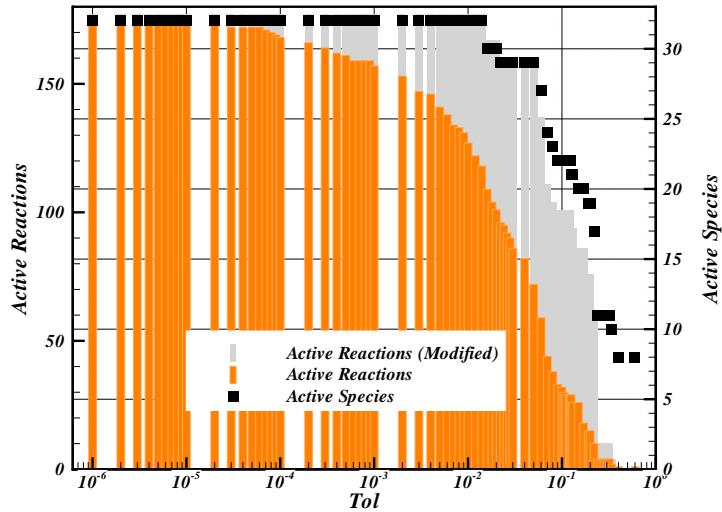


Fig. 2.17: Number of active reactions in sets $R(tol, S_{user})$ (simplified mechanisms without reaction recovery) and in $R(tol, S_{user}) \cup R_{rec}(tol, S_{user})$ (simplified mechanisms with reaction recovery). Also shown is the no. of active species which is the same in both cases.

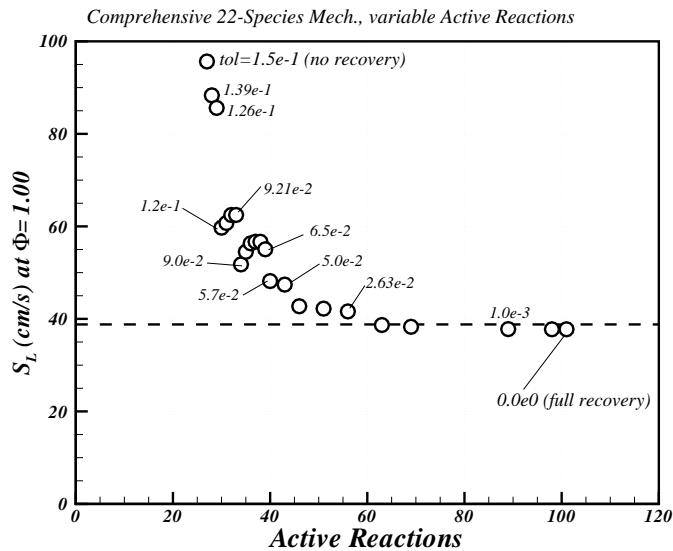


Fig. 2.18: Premixed stoichiometric ($\Phi = 1.00$) laminar flame speed S_L resulting from a 22-species comprehensive mechanism with variable number of active reactions. Dashed horizontal line represents the flame speed $S_L = 38.8\text{cm/s}$ resulting from GRI-Mech 1.2.

In particular, for a value of the threshold on the importance indices $tol = 0.15$ and the prescribed S_{user} , the unmodified simplification procedure yielded a simplified mechanism comprising 22 (out of 32) active species and 27 (out of 177) active reactions. We then considered the effect of progressively recovering, in order of importance, the originally discarded elementary reactions solely involving the 22 selected species, by gradually lowering the threshold value as shown in

Fig. 2.18, where the numbers beside the symbols indicate the threshold value, and where $tol = 0.0$ indicates a 22 species mechanism with full recovery i.e. comprising all the reactions (101) present in GRI-Mech 1.2 which involve the 22 species. Fig. 2.18 shows the resulting laminar flame velocities at $\Phi = 1.0$ which, as the number of recovered reactions increases, exhibits an exponential convergence towards the reference value of 38.8 cm/s obtained by PREMIX with GRI-Mech 1.2. In particular we observe that a 46 and a 63 reactions mechanism yields, respectively, an 8% and a 2% error in flame speed.

A sample of 8 comprehensive mechanisms were then chosen for validation and is summarized in Table 2.8. Also shown in Table 2.8 are the species that were progressively removed from the mechanisms as the tolerance was increased.

The validity and accuracy of the comprehensive mechanisms was tested by re-running PREMIX at each Φ with GRI Mech.1.2 being replaced by each of the simplified mechanisms in turn, and by estimating the resulting premixed laminar flame speeds S_L . The results of this parametric analysis are summarized in Fig.(2.19), where S_L is displayed as a function of the equivalence ratio Φ . The results of Fig.(2.19) show that the flame speed is reproduced extremely accurately up to and not including the 14-species 33-reactions mechanism. The degree of comprehensiveness of each mechanism is demonstrated by the low errors found at all Φ 's. The reduction in the number of species leads to a speed-up in CPU time which can reach a factor of 3.15 (with respect to the detailed mechanism) for the 19-species, 77-reactions mechanism.

2.9.3 CSP data analysis.

As with the perfectly stirred reactor, CSP data, such as importance indices, can serve as an effective diagnostic tool to analyse the premixed flame. In this respect, it is interesting to use such data to probe the individual role of each elementary reaction along the axial coordinate x from the inlet to the preheat zone and through the reaction zone. In particular we can analyse the pathway of the C-containing species, $\text{CH}_4 \rightarrow \text{CH}_3 \rightarrow \text{CH}_2\text{O} \rightarrow \text{CO} \rightarrow \text{CO}_2$, from the fuel to the intermediate species to the burnout of CO to form CO_2 . For each of these species, we can single out the reactions exhibiting the greatest importance indices and plot the latter along x so as to highlight the actual pathway structure.

Stoichiometric flame

Fig.2.21 is a collection of importance indices plots together with the mass fractions and temperature along $x(\text{cm})$ relative to CSP data computed for the stoichiometric ($\Phi = 1$) flame. Fig.2.21(h) is a

<i>tol</i>	Active Spec.	Active Reac.	S_{rem}
3.0×10^{-2}	30	159	H ₂ O ₂ C ₂ H
6.0×10^{-2}	29	157	HCCOH
7.0×10^{-2}	26	128	CH ₂ OH CH ₃ OH CH ₂ CO
1.0×10^{-1}	25	123	C
1.2×10^{-1}	23	107	C ₂ H ₂ HCCO
1.5×10^{-1}	22	101	C ₂ H ₃
1.9×10^{-1}	19	77	CH ₃ O C ₂ H ₄ C ₂ H ₅
2.4×10^{-1}	14	33	CH CH ₂ CH ₂ (S) C ₂ H ₆ HO ₂

Table 2.8: Summary of comprehensive mechanisms used for validation. Also shown, set S_{rem} containing species progressively removed from the simplified mechanisms as *tol* is increased.

plot of the slow subspace importance indices relative to temperature, which is one of the unknowns. Note the presence of the importance index for the heat diffusion on temperature along with the indices relative to four selected reactions. The interval $x \in [1.06, 1.12]$ may be considered as the reaction zone where the bulk of chemical energy is released as heat. Temperature is 'diffused' away from this region as can be seen from the negative value of the corresponding index. On the other hand recombination reactions such as $2\text{CH}_3(+\text{M}) \rightarrow \text{C}_2\text{H}_6(+\text{M})$ and $\text{H}+\text{CH}_3(+\text{M}) \rightarrow \text{CH}_4(+\text{M})$ tend to increase the temperature due to their exothermic nature. Note also that heat diffusion is the only responsible for the temperature increase in the region $x < 1.06\text{cm}$.

The preheat zone may be identified as developing from $x < 1.06$. The only noticeable reaction that extends as far upstream as the preheat zone is $\text{OH}+\text{CH}_4 \rightarrow \text{CH}_3+\text{H}_2\text{O}$ involving the production of the CH₃ radical. This can be seen in Fig.2.21(d), involving the CH₃ indices. Note how the CH₃ diffusion index - highlighting the upstream diffusion of such radical - is different(smaller) from that of other species because of the presence of the mentioned reaction far upstream, which also contributes to the formation of CH₃ (recall that slow indices should sum up to unity in modulus). Also far upstream are the CH₃-depleting reactions $\text{HO}_2+\text{CH}_3 \rightarrow \text{OH}+\text{CH}_3\text{O}$, $2\text{CH}_3+\text{M} \rightarrow \text{C}_2\text{H}_6+\text{M}$ and $\text{O}+\text{CH}_3 \rightarrow \text{H}+\text{CH}_2\text{O}$ which propagate radicals CH₂O, CH₃O and C₂H₆. The bulk of these last

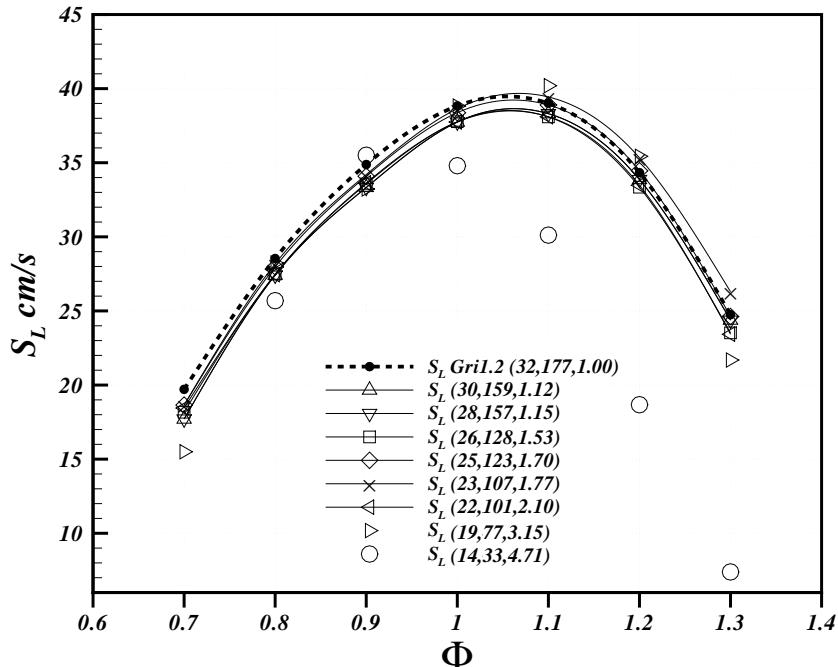


Fig. 2.19: Premixed laminar flame speeds S_L (cm/s) resulting from the detailed GRI Mech. 1.2 mechanism and from the comprehensive simplified mechanisms. The first number in brackets is the number of active species in the mechanism tested, the second is the number of reactions and the third number is the average speed-up factor with respect to the detailed mechanism.

three reactions may then be considered as a reaction propagation process.

We can now examine intermediate species such as CH_2O which, as shown in Fig.2.4, acts as a 'bridge' from CH_3 to the HCO radical which will ultimately yield CO . Fig.2.21(e) shows that CH_2O is produced in a sequence of three reactions which in order of increasing x are $\text{CH}_3\text{O}+\text{O}_2 \rightarrow \text{HO}_2+\text{CH}_2\text{O}$, $\text{H}+\text{CH}_2\text{O}+\text{M} \leftarrow \text{CH}_3\text{O}+\text{M}$ and $\text{O}+\text{CH}_3 \rightarrow \text{H}+\text{CH}_2\text{O}$. Before exiting the preheat zone, CH_2O is depleted to yield HCO through these reactions: $\text{OH}+\text{CH}_2\text{O} \rightarrow \text{HCO}+\text{H}_2\text{O}$ and $\text{H}+\text{CH}_2\text{O} \rightarrow \text{HCO}+\text{H}_2$.

The final part of the reaction zone involves the production of CO from HCO and its subsequent burnout into CO_2 . As shown in Fig.2.21(f), the CO indices indicate that CO is produced via $\text{HCO}+\text{O}_2 \rightarrow \text{HO}_2+\text{CO}$ and $\text{HCO}+\text{H}_2\text{O} \rightarrow \text{H}+\text{CO}+\text{H}_2\text{O}$. Subsequently CO is oxidized into CO_2 through $\text{OH}+\text{CO} = \text{H}+\text{CO}_2$ which eventually reaches equilibrium farther downstream (the backward and forward indices balance). This latter equilibrium may also be viewed in Fig. 2.21(g) involving the CO_2 indices.

Figs. 2.23-2.23 are a collection of velocity, temperature and mass fraction plots along the longitudinal coordinate for the detailed (Gri 1.2) stoichiometric PREMIX solution and the simplified PREMIX solution using the 19 species, 77 reactions mechanism. The temperature and velocity

profiles are both very accurately reproduced by the simplified mechanism, while some radical concentrations may differ slightly. In particular C_2H_6 concentration seems to be overestimated due to the elimination of the C_2H_X ($X < 6$) species from the simplified mechanism and therefore the truncation of the $\text{C}_2\text{H}_6 \rightarrow \text{C}_2\text{H}_X \rightarrow \text{C}_2\text{H}$ pathway.

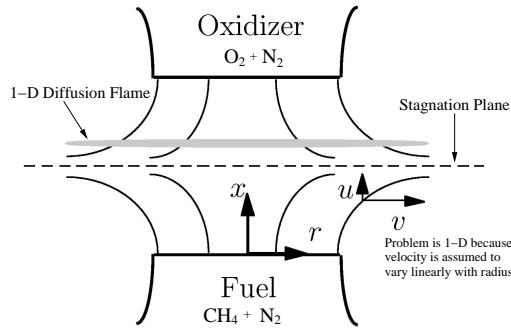


Fig. 2.20: Geometric configuration of the counterflow diffusion flame.

Lean flame

Fig. 2.22 is a collection of the same plots seen previously relative to the lean $\Phi = 0.7$ flame. The most noticeable difference is a wider reaction zone and a greater flame thickness⁸. This is clear from the temperature profile and from the temperature indices. Note that reaction $2\text{CH}_3(+\text{M}) \rightarrow \text{C}_2\text{H}_6(+\text{M})$ plays a lesser role in the temperature rise.

Also evident for the lean flame is that the CH_3 -producing reaction $\text{OH}+\text{CH}_4 \rightarrow \text{CH}_3+\text{H}_2\text{O}$ extends farther upstream in the preheat zone than in the stoichiometric case.

2.10 Applications to counterflow diffusion flames

Using a similar procedure to that described for premixed flames, a simplification study of the GRI-Mech 1.2 detailed mechanism was carried out for counterflow diffusion flames. The OPPDIFF [62] package was used to generate initial conditions for the unsteady low Mach number solver at different average normal strain rates a_n resulting from various inflow velocities $u_F = u_O$ where u_F and u_O are the fuel and oxidizer nozzle inflow velocities respectively, separated by 10 cm. The geometric configuration is 1-D owing to the assumption of a linear variation of the radial velocity with the radial distance from the axis as shown in Fig. 2.20. Both fuel and oxidizer mixtures are injected

⁸The definition of flame thickness may vary. We could adopt it as the interval where the temperature is between 5% and 95% of the inlet temperature.

at 300 K and 1 atm and are diluted with N₂ (60% N₂ molar composition in the fuel mixture and 79% in the oxidizer mixture) whereas inlet velocities are varied between 35cm/s to 80 cm/s. The solutions were then analyzed and the related CSP databases constructed.

<i>tol</i>	Active Spec.	Active Reac.	<i>S_{rem}</i>
1.05×10^{-1}	26	131	H ₂ O ₂
			C
			HCCOH
			CH ₂ OH
			CH ₃ OH
			CH ₂ CO
1.50×10^{-1}	24	112	HCCO
			CH ₃ O
1.70×10^{-1}	23	106	C ₂ H
2.00×10^{-1}	20	66	HO ₂
			C ₂ H ₂
			CH ₂ (S)
2.20×10^{-1}	17	43	CH ₂ O
			C ₂ H ₃
			C ₂ H ₆
2.60×10^{-1}	15	37	C ₂ H ₄
			C ₂ H ₅

Table 2.9: Summary of comprehensive mechanisms used for the counterfbw diffusion flame validation. Also shown, set *S_{rem}* of species progressively removed from the simplified mechanisms as *tol* is increased.

A number of comprehensive simplified mechanisms were generated from data- bases relative to different fuel (oxidizer) jet velocities ($u_F = u_O = 35, 40, 50, 60, 70, 80$ cm/s), where the complete recovery of non-active reactions was enforced as in the premixed flame case. As kernel set we chose $S_0=[\text{CH}_4,\text{O}_2]$, since in a counter-flow flames the simplified mechanism must be able to replicate accurately the decomposition of both fuel and oxidizer in approaching the flame region. However, since an accurate prediction of the rates of reactant consumption necessarily involves predicting the rates of formation of the products with similar accuracy, it is expected that the iterative algorithm will be able to identify the reaction pathways most relevant for the major products (CO₂, H₂O) and temperature, although the latter were not included in the kernel set.

A subset of 6 comprehensive mechanisms was selected for validation (Tab. 2.9), and employed in OPPDIFF re-runs using the same set of boundary conditions. Results are summarized in Fig. 2.25. Panel (a) reports the peak values of temperature (T_{\max}) as a function of the normal strain rate a_n (averaged over the spatial domain), whereas panels (c) and (f) report the reactant CH₄, O₂ and the products CO₂, H₂O fields, respectively, for a fuel (oxidizer) jet velocity of $u_F = u_O = 70$ cm/s. The temperature as well as reactants and products fields are reproduced within a maximum 2% error

with respect to the GRI-Mech 1.2 profiles even for the most simplified 15-species mechanism.

Although for none of the 7 simplified mechanisms the peak temperature as a function of the normal strain rate a_n , (Fig. 2.25-a) ever crosses the reference profile obtained with GRI-Mech 1.2, we find that some mechanisms over-predict the correct temperature while others under-predict it. The relation between this behavior and the mechanism simplifications has not yet been investigated.

In Fig. 2.25-c, it is apparent that the degree of leakage of CH_4 towards the oxidizer side and of O_2 towards the fuel side depends on the specific simplified kinetics, since the leakage of reactants is a characteristic effect of the kinetics proceeding at finite rates. Even with the most simplified mechanism, the largest departures from the reference values occur for mass fractions lower than 10^{-3} . The error on the fuel is monotonically increasing, whereas the error on the oxidizer has a non-monotonic dependence, with an increasing degree of simplification. It is noteworthy that the error in the products is even smaller than that in the reactants, even at very low mass fraction values (Fig. 2.25-f).

The CH_3 peak mass fraction dependence on the normal strain rate a_n , (Fig. 2.25-b), and the CH_3 (Fig. 2.25-d) and HCO (Fig. 2.25-e) mass fraction fields for $u_F = u_O = 70 \text{ cm/s}$ indicate that the error on intermediate species becomes significant for simplified mechanisms with a number of species smaller than 23. As expected, the accuracy on intermediate species is lower than that achieved for reactants and products, this loss of accuracy being the price to pay for achieving the mechanism simplification.

Overall, the errors are roughly constant and bounded throughout the strain rate range, indicating the applicability of comprehensive mechanisms over a range of boundary conditions.

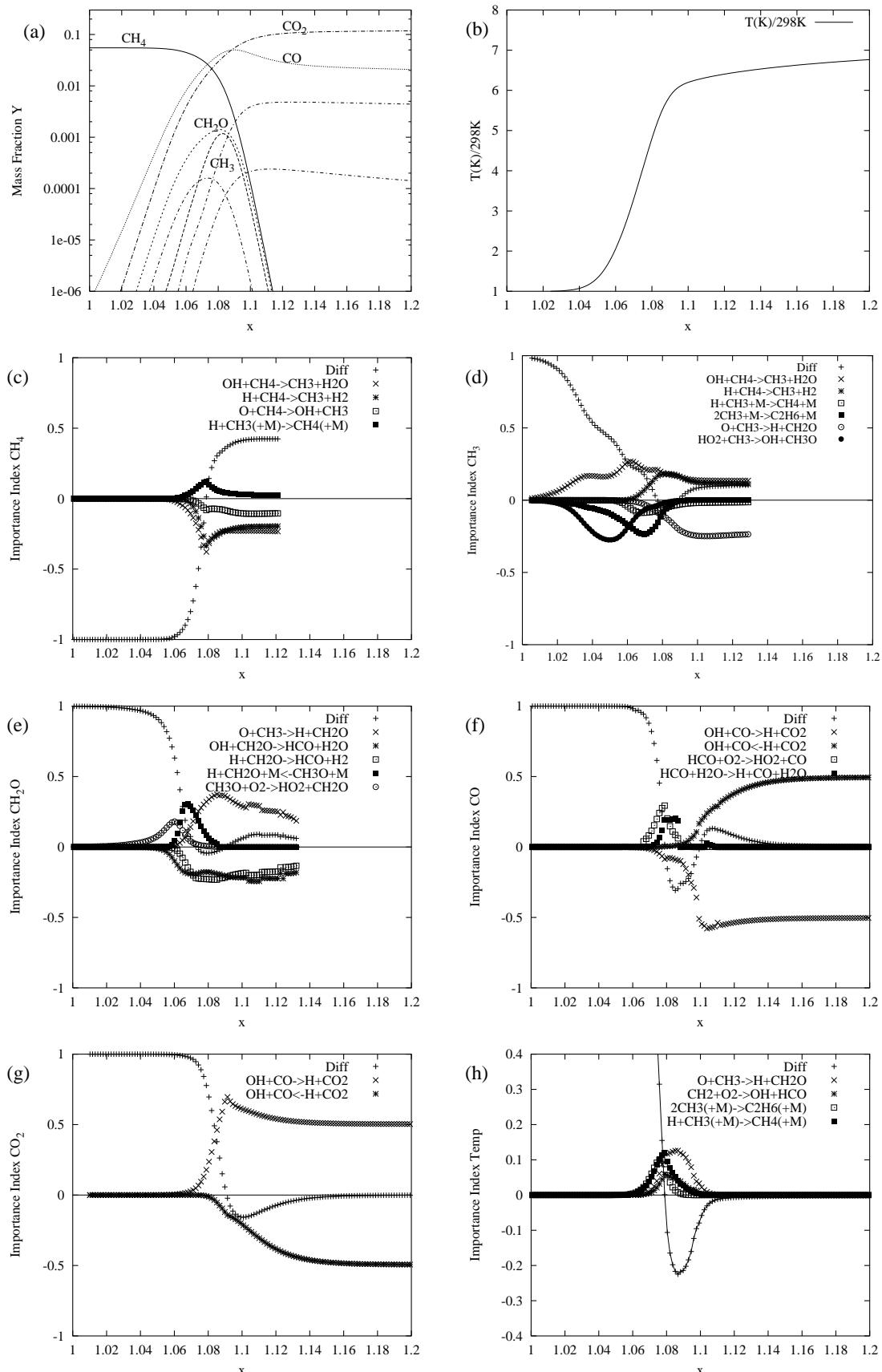


Fig. 2.21: Solution and importance indices for the $\Phi = 1.0$ laminar premixed methane-air flame.

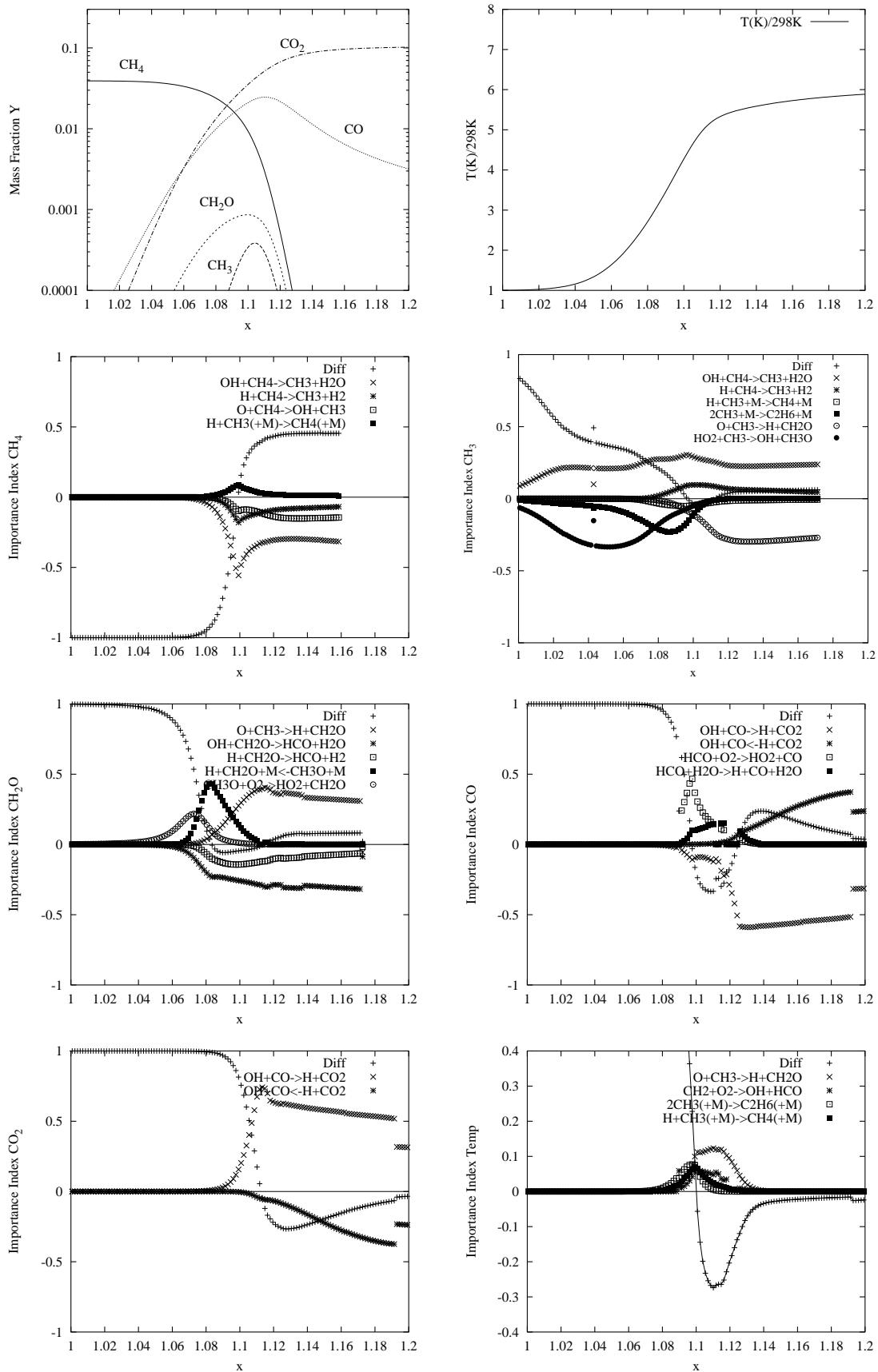


Fig. 2.22: Solution and importance indices for the $\Phi = 0.7$ laminar premixed methane-air flame.

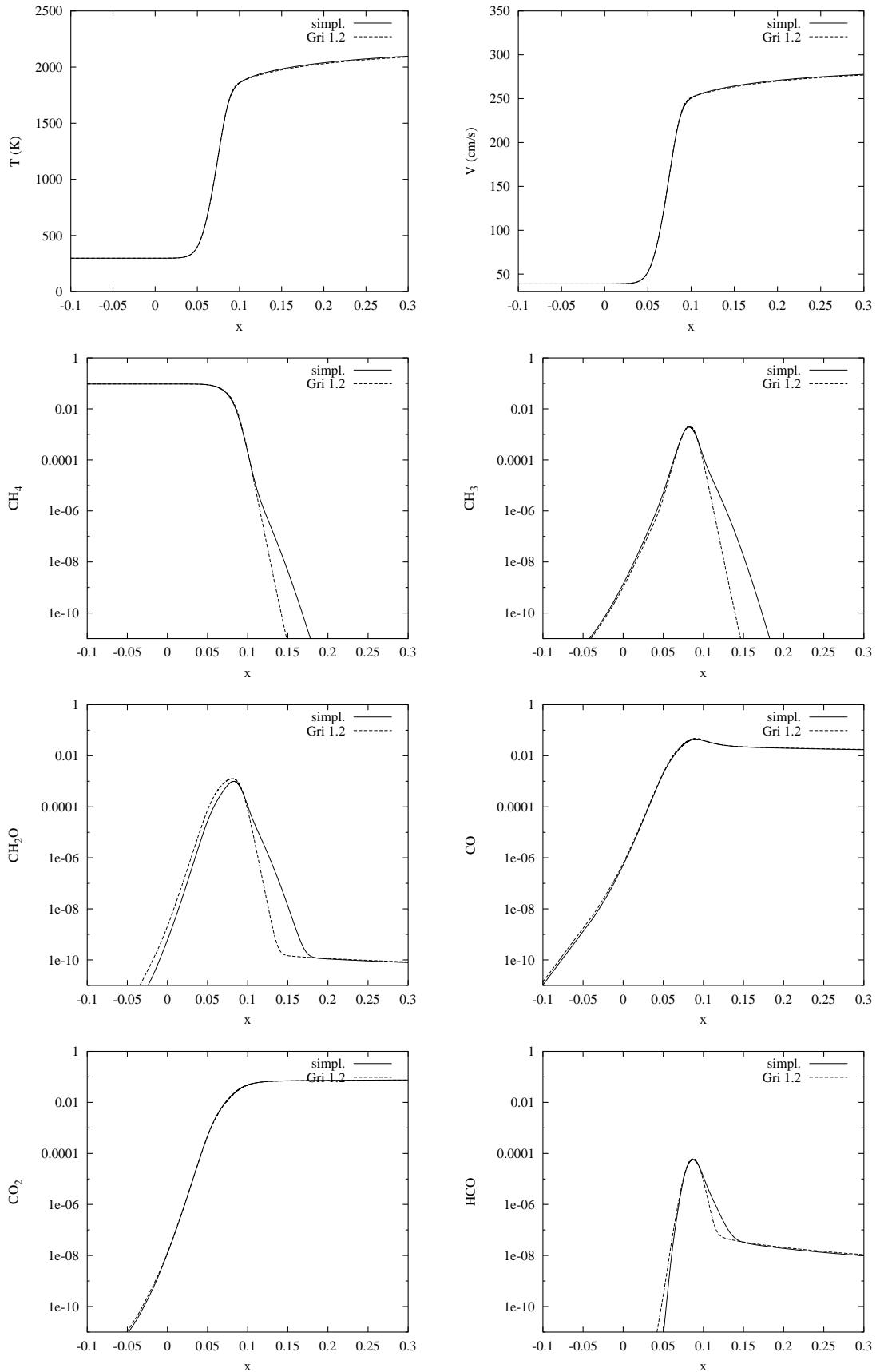


Fig. 2.23: Detailed Gri 1.2 and simplified solutions ($tol = 1.9 \times 10^{-1}$, 19 species, 77 reactions) for the $\Phi = 1.0$ laminar premixed methane-air flame.

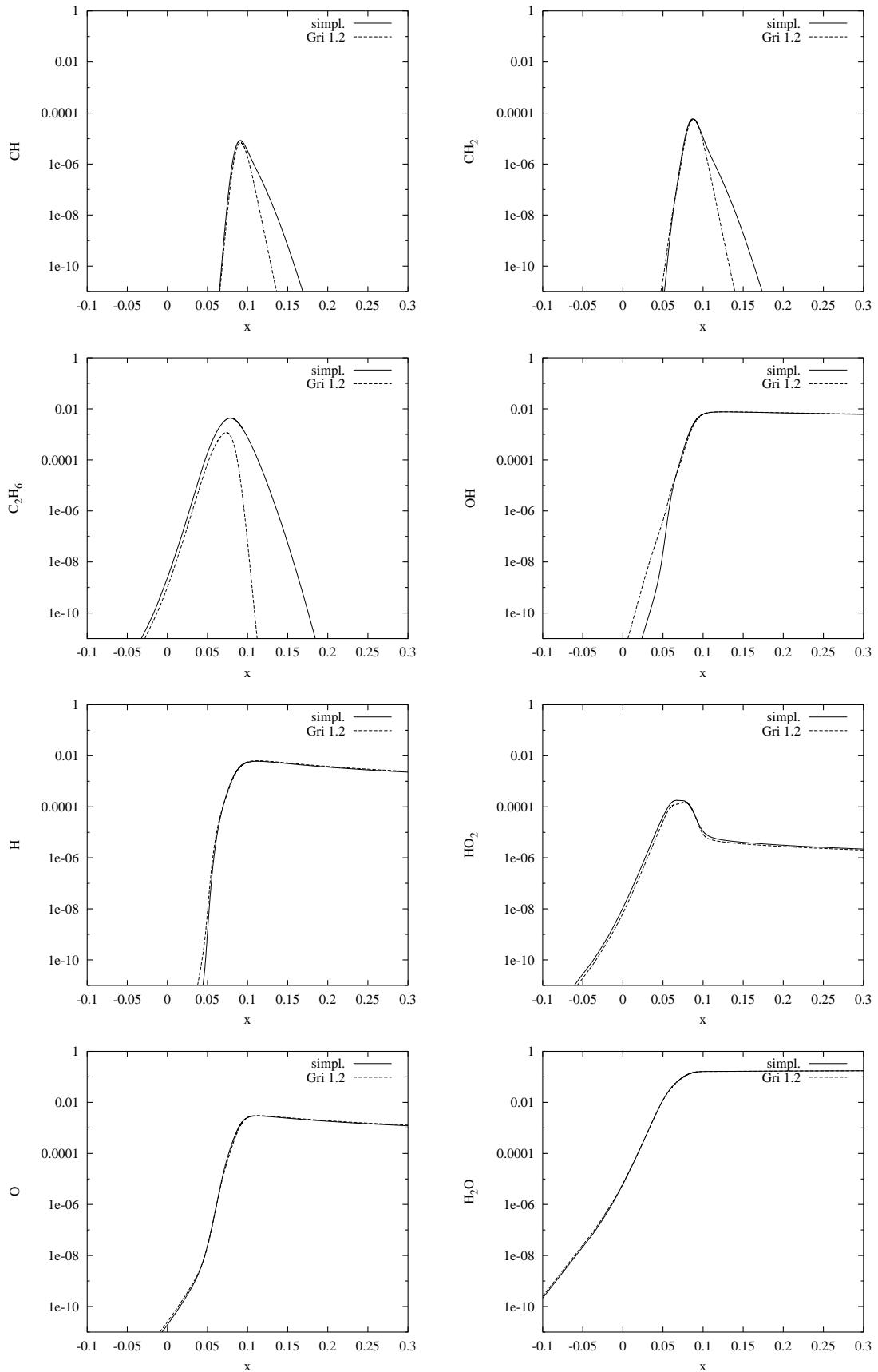


Fig. 2.24: (cont'd) Detailed Gri 1.2 and simplified solutions ($\text{tol} = 1.9 \times 10^{-1}$, 19 species, 77 reactions) for the $\Phi = 1.0$ laminar premixed methane-air flame.

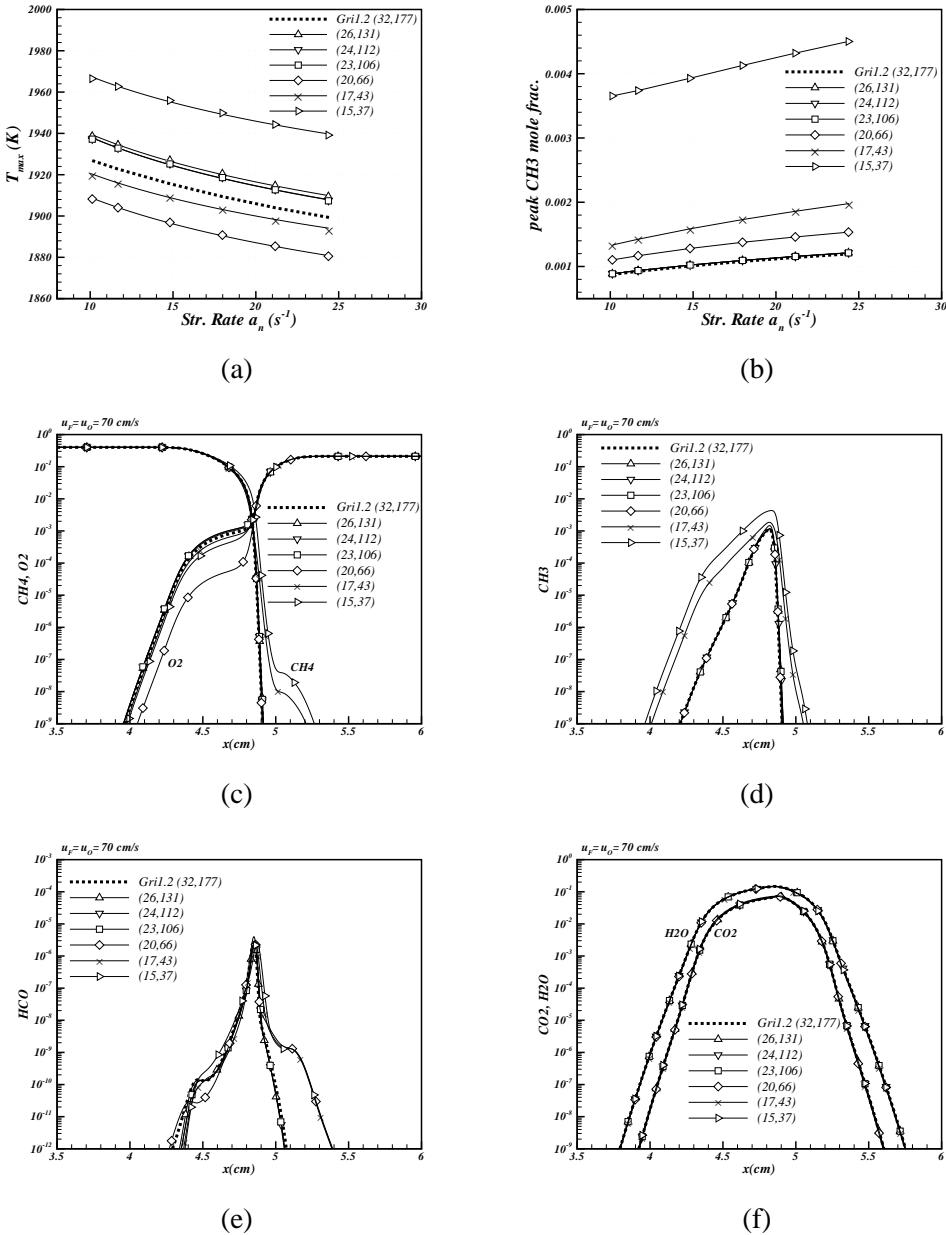


Fig. 2.25: (a),(b) Peak temperature T_{max} [K] and peak CH_3 mole fraction for various simplified mechs. at different average normal strain rate values a_n (s^{-1}) resulting from infbw fuel/oxydizer velocities $u_F = u_O \in [35 \text{ cm/s}, 80 \text{ cm/s}]$; numbers in brackets are active species and reactions in simplified mechanisms; (c)-(f) molar composition along x (cm) for the $u_F = u_O = 70 \text{ cm/s}$ case; (c) CH_4 and O_2 , (d) CH_3 , (e) HCO , (f) H_2O and CO_2 .

Chapter 3

Natural Tangent Dynamics with Recurrent Biorthonormalizations

3.1 Introduction

In this chapter we will attempt to restructure our approach to dynamical systems by broadening our view to systems exhibiting asymptotic behaviors that are less trivial than mere equilibrium points (i.e. limit cycles, chaotic attractors and in general behaviors characterized by persistent oscillations) as they are indeed frequent in phenomena ranging from combustion kinetics to enzymatic reactions [63, 64, 65, 66]. Such an attempt will heavily rely on the analytical closed theory of linear autonomous dynamical systems at first and subsequently will draw a parallel with non linear autonomous systems in terms of invariant global geometrical properties. As with linear systems, the definition and identification of an invariant geometrical structure for non linear systems will provide a natural setting in which the dynamics can be viewed allowing for the complete decoupling of each of the stable and unstable components of the dynamics. The final part of the chapter will be dedicated to the *Natural Tangent Dynamics with Recurrent Biorthonormalizations* technique (NTDRB for short) for the determination of the invariant geometrical structure of a dynamical system independently of its possible singularly perturbed nature and not necessarily possessing a trivial limit set.

Let us first recall some basic definitions and notations. Consider the system of ordinary differential equations

$$\frac{d\mathbf{x}}{dt} = \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \quad (3.1)$$

where $\mathbf{x} = \mathbf{x}(t) \in C \subseteq R^n$ is a vector valued function of time and $\mathbf{f} : C \rightarrow TC$ is a vector field

defined on the tangent bundle TC of the smooth manifold C . The vector field \mathbf{f} is said to generate a flow $\phi_t : C \rightarrow C$ where $\phi_t(\mathbf{x}) = \phi(\mathbf{x}, t) : C \times R \rightarrow R^n$ is a smooth function defined for all $\mathbf{x} \in C$ and $t \in I = (a, b) \subseteq R$ which satisfies 3.1 so that

$$\frac{d\phi(\mathbf{x}, t)}{dt} \Big|_{t=\tau} = \mathbf{f}(\phi(\mathbf{x}, \tau))$$

In its domain of definition ϕ_t satisfies the group properties (i) $\phi_0 = Id$ and (ii) $\phi_{t+s} = \phi_t \circ \phi_s$.

Given an initial condition $\mathbf{x}(0) = \mathbf{x}_0 \in C$ then the flow defines the solution of the Cauchy problem associated with 3.1, i.e. the unique trajectory $\mathbf{x}(t) = \phi(\mathbf{x}_0, t) = \phi_t(\mathbf{x}_0)$ based at \mathbf{x}_0 . This means that $\phi_t(\mathbf{x}_0)$ maps the initial condition to the condition (state of the system) at time t . $\phi_t(\mathbf{x})$ with $\mathbf{x} \in C$ may then be interpreted as the entire family of trajectories based at C .

Eq. 3.1 represents an autonomous non-linear differential equation. This means that the vector field \mathbf{f} is not an explicit function of time but rather it is merely a function of the phase space points \mathbf{x} . Intuitively this means that 3.1 is a "stand-alone" system, independent (autonomous) of the evolution of some other system.

The question arises as to how to study the global dynamical behavior of 3.1. A logical starting point is to revert at first to linear autonomous systems for which the dynamics is analytically known. A system such as $\dot{\mathbf{x}} = A\mathbf{x}$ where A is a constant matrix, gives rise to a dynamical behavior which is completely described by the spectral properties of A , that is by its 'flat' eigenspaces and constant eigenvalues. In fact, as we will shortly see, the (generalized) eigenspaces of A provide the natural geometrical setting in which to view the dynamics. The vector field $A\mathbf{x}$ generates the analytical flow $\phi_t = e^{At} : R^n \rightarrow R^n$ whose eigenvectors (coinciding with the eigenvectors of A) are invariant under e^{At} , i.e. under the system's dynamics. Any vector belonging initially to one such eigenspace will stay in it as the dynamics evolves, contracting or elongating according to the real part of the corresponding eigenvalue. In particular, the phase space will be decomposed into the direct sum of a stable E^s , an unstable E^u and a central subspace E^c each spanned respectively by the eigenvectors of A whose eigenvalues have negative, positive or zero real part. Hence solutions lying in the stable subspace will exponentially decay, those lying in the unstable subspace will grow exponentially and those in the center subspace will neither grow nor decay.

The leap from autonomous linear dynamics to non linear dynamics is undoubtedly huge. A legitimate question however can arise as to whether and how it is possible to extend the notions of invariant geometrical properties to the realm of non linear dynamics. A classical starting point is to accept the fact that stability in the context of linear systems is a global property, carried along the phase space

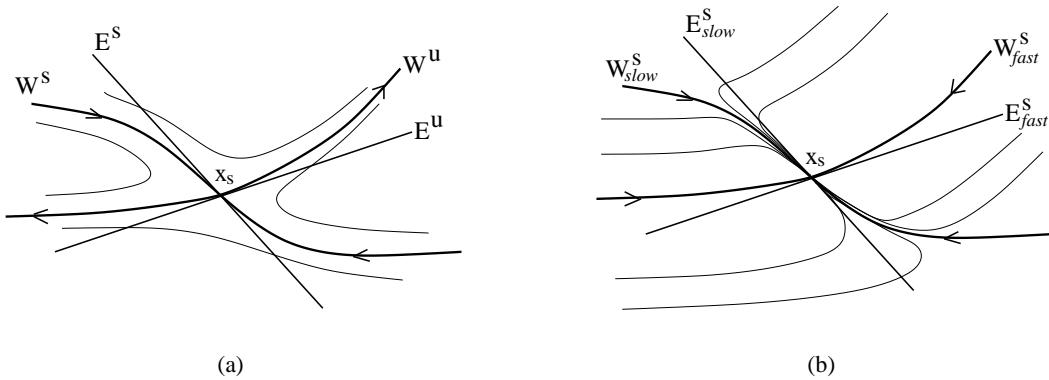


Fig. 3.1: The Hartman-Grobman theorem. (a) stable and unstable manifolds. (b) slow and fast stable manifolds by flat stable and unstable invariant subspaces, whereas it becomes a local property of equilibrium points (or invariant sets) with non linear systems. We may start to draw some parallels between the two cases exploiting the well known Hartman-Grobman theorem [67] according to which any non linear system is locally topologically equivalent near (in a neighborhood I_{x_s}) of its equilibrium point x_s (such that $f(x_s) = \mathbf{0}$) to the linearized system $\dot{\xi} = J\xi$ (where $J = Df(x_s)$ is the constant Jacobian matrix at the equilibrium point and $\xi = x - x_s, |\xi| \ll 1$) provided such linearized system is hyperbolic, i.e. J has no eigenvalues with zero real part (i.e. no center eigenspace). Non linear analogues to the flat stable and unstable eigenspaces of the linear problem can be defined naming them local invariant stable ($W_{loc}^s(x_s)$) and unstable ($W_{loc}^u(x_s)$) manifolds, where

$$W_{loc}^s(x_s) = \{x \in I_{x_s} | \phi_t(x) \rightarrow x_s \text{ as } t \rightarrow \infty, \text{ and } \phi_t(x) \in I_{x_s} \text{ for all } t \geq 0\}, \quad (3.2)$$

$$W_{loc}^u(x_s) = \{x \in I_{x_s} | \phi_t(x) \rightarrow x_s \text{ as } t \rightarrow -\infty, \text{ and } \phi_t(x) \in I_{x_s} \text{ for all } t \leq 0\} \quad (3.3)$$

Topological equivalence means that $W_{loc}^s(x_s)$ and $W_{loc}^u(x_s)$ have the same dimensions and are tangent at x_s to the stable and unstable eigenspaces of the linear system, this implying that the non linear dynamics is in every respect locally 'similar' to the linear dynamics as shown in Fig.(3.1). Moreover we can 'lift' or 'carry over' [67, 68] the local invariant manifolds and define through the images of $W_{loc}^s(x_s)$ and $W_{loc}^u(x_s)$ stable and unstable global invariant manifolds by letting points in W_{loc}^s and W_{loc}^u flow backward and forward in time respectively:

$$W^s(x_s) = \bigcup_{t \leq 0} \phi_t(W_{loc}^s(x_s)) \quad (3.4)$$

$$W^u(x_s) = \bigcup_{t \geq 0} \phi_t(W_{loc}^u(x_s)). \quad (3.5)$$

Indeed much of the complexity of non linear systems arises from the fact that stable and unstable

manifolds (of the same or of distinct equilibrium points) can and often do intersect this implying that the qualitative behavior around the equilibrium point holds only locally.

Note that singularly perturbed systems possessing an equilibrium point for which the CSP theory holds can be viewed as having a stable manifold and no unstable manifold. In turn, in presence of a sufficient temporal gap between fast and slow dynamics, the stable manifold will decompose into a slow and a fast stable manifold as shown in Fig.(3.1).

Non linear systems allow for limit sets that are more complex than fixed points. Limit cycles constitute an example and for them too, with the aid of Poincare maps and their linearization, it is possible to define local and global stable and unstable invariant manifolds in the sense discussed above. A far more complex behavior is observed for those (deterministic) systems exhibiting chaotic limit sets for which an infinite set of invariant manifolds exists (stemming from dense sets of unstable limit cycles) giving rise to a so called stable/unstable invariant foliation the definition and analysis of which would go far beyond the scope of the present dissertation. It is however important to underline the fact that in the presence of hyperbolicity the local stability analysis is a very powerful tool leading to such geometrical concepts as invariant manifolds and ultimately to bifurcation theory and normal form theory [69]. In the presence of very complex dynamical behavior (and high dimensionality), however, it can become impossible to completely characterize a system with the tools discussed so far. Indeed for the limited class of singularly perturbed systems possessing a fixed equilibrium point, CSP theory has shown that a slow invariant manifold can be defined, provided all trajectories are stable and a sufficient timescale separation exists. The theory however fails whenever one of its prerequisites is absent.

An alternative approach must therefore be sought. Our original question of defining, for non linear dynamical systems, some sort of substitute for deterministic linear algebra, which represents a 'closed' theory for linear systems, still remains open. In particular we would be interested in defining a class of spectral objects playing an equivalent role to eigenspaces and eigenvalues of A in linear systems. The statistical/geometrical approach summarized in the *multiplicative ergodic theorem* of Oseledec [68], [70], as we will see in greater depth, seems to constitute a valid tool capable of answering our question. Such theorem provides a closed spectral theory for linear non autonomous systems described by $\dot{\mathbf{v}} = A(\mathbf{x}(t))\mathbf{v}$ where $\mathbf{x}(t)$ is the solution of some non linear autonomous system Eq.(3.1). In particular if $A(\mathbf{x}(t)) = D(\mathbf{f}(\mathbf{x}(t)))$ then $\dot{\mathbf{v}} = D(\mathbf{f}(\mathbf{x}(t)))\mathbf{v}$ is the so called variational equation describing the evolution of normalized perturbations of trajectories of system Eq.(3.1), viewed as vector elements, within the tangent bundle TC covering the smooth manifold

C on which system Eq.(3.1) is defined. Our hope is therefore that of characterizing the pointwise invariant geometrical structure of Eq.(3.1) and its characteristic timescale spectrum in terms of subspaces defined within the tangent space that are invariant for $\dot{\mathbf{v}} = A(\mathbf{x}(t))\mathbf{v}$, and numbers (Lyapunov exponents) expressing the time-averaged growth of vectors within such subspaces.

We should emphasize at this point that studying the eigenvectors and eigenvalues of $A(\mathbf{x}(t)) = D(\mathbf{f}(\mathbf{x}(t)))$ is misleading and unappropriate as they can have little or nothing to do with both the invariant geometrical properties of the non linear system and with its characteristic timescales and can definitely fail, from what we have learned, in characterizing the asymptotic properties on limit cycles and chaotic attractors.

Before proceeding further let us define the variational equation describing, as mentioned, the tangent bundle evolution of normalized perturbations of trajectories. Given the system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t))$, let us consider two distinct nearby trajectories $\mathbf{x}_1(t)$ and $\mathbf{x}_2(t)$. The evolution of the displacement vector is given by

$$\begin{aligned}\frac{d(\mathbf{x}_2 - \mathbf{x}_1)}{dt} &= \mathbf{f}(\mathbf{x}_2) - \mathbf{f}(\mathbf{x}_1) \\ &= \mathbf{f}(\mathbf{x}_1) + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_2 - \mathbf{x}_1) + \mathcal{O}(\|\mathbf{x}_2 - \mathbf{x}_1\|^2) - \mathbf{f}(\mathbf{x}_1).\end{aligned}\quad (3.6)$$

Placing $\|\mathbf{x}_2(t = 0) - \mathbf{x}_1(t = 0)\| = \varepsilon$ and defining the normalized perturbation vector as $\mathbf{v}(t) = \lim_{\varepsilon \rightarrow 0} \frac{\mathbf{x}_2 - \mathbf{x}_1}{\varepsilon}$, thereby implying $\mathbf{v}(t = 0) = \lim_{\varepsilon \rightarrow 0} \frac{\mathbf{x}_2(0) - \mathbf{x}_1(0)}{\varepsilon} = \mathbf{1}$, then by dividing Eq.(3.6) by ε and letting $\varepsilon \rightarrow 0$ we have

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\mathbf{v} + \lim_{\varepsilon \rightarrow 0} \frac{\mathcal{O}(\|\mathbf{x}_2 - \mathbf{x}_1\|^2)}{\varepsilon}$$

Note that if $\|\mathbf{f}(\mathbf{x}(t))\| \leq L\|\mathbf{x}(t)\|$, $\forall t$, where L is a non-negative constant, then by Gronwall's lemma $\|\mathbf{x}(t)\| \leq \|\mathbf{x}(t = 0)\|e^{Lt}$. Then $\|\mathbf{x}_2(t) - \mathbf{x}_1(t)\| \leq \|\mathbf{x}_2(t)\| - \|\mathbf{x}_1(t)\| \leq (\|\mathbf{x}_2(t = 0)\| - \|\mathbf{x}_1(t = 0)\|)e^{Lt} \leq \|\mathbf{x}_2(t = 0) - \mathbf{x}_1(t = 0)\|e^{Lt} = \varepsilon e^{Lt}$. Therefore $\lim_{\varepsilon \rightarrow 0} (\mathcal{O}(\|\mathbf{x}_2 - \mathbf{x}_1\|^2)/\varepsilon) = \lim_{\varepsilon \rightarrow 0} (\mathcal{O}(\varepsilon^2 e^{Lt})/\varepsilon) = 0$ and the variational equation reduces to

$$\frac{d\mathbf{v}}{dt} = \frac{\partial \mathbf{f}(\mathbf{x}(t))}{\partial \mathbf{x}}\mathbf{v} \quad (3.7)$$

with the initial condition $\mathbf{v}(t = 0) = \mathbf{1}$. Note that the variational equation is coupled with Eq(3.1) and will form with it a skew-product system which will lie at the core of the techniques to be later discussed.

The following section will be entirely devoted to the spectral theory of linear autonomous systems

and the techniques of linear algebra needed to tackle them. This will serve as an introduction to non-linear systems and their related linear non autonomous variational equation for which Oseledec's theorem will formulate a spectral theory.

3.2 Linear dynamics

Consider a generic dynamical system

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x})$$

Where $\mathbf{x} \in C \subseteq R^n$ and $\mathbf{f}(\mathbf{x}) \in TC_{\mathbf{x}}$ is a vector field belonging to the tangent space¹ $TC_{\mathbf{x}}$ in \mathbf{x} and $\mathbf{x}(t = 0) = \mathbf{x}_0$. Such system will induce the evolution of vectors $\mathbf{a} \in TC_{\mathbf{x}}$ by means of the skew-product system:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \quad (3.8)$$

$$\dot{\mathbf{a}} = J(\mathbf{x})\mathbf{a} \quad (3.9)$$

where $\mathbf{a}(t = 0) = \mathbf{a}(0)$ and $J(t) = J(\mathbf{x}(t)) = J(\phi_t(\mathbf{x}_0)) = \partial f^i / \partial x_j$ is the jacobian matrix of the vector field $\mathbf{f}(\mathbf{x})$, and where $\phi_t : C \rightarrow C$ is the phase flow associated with Eq. (3.8), $\phi_t(\mathbf{x}_0) = \mathbf{x}(t)$ being the solution of the Cauchy problem of Eq. (3.8) itself.

If $\mathbf{f}(\mathbf{x})$ is a linear function (i.e. $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$) then $J = A$ where A is a constant matrix and the system is decoupled. Vector dynamics is autonomous and simply reduces to:

$$\dot{\mathbf{a}} = A\mathbf{a} \quad (3.10)$$

Let us suppose that matrix A has real and distinct eigenvalues each with algebraic multiplicity equal to 1 with corresponding eigenvectors with geometric multiplicity equal to 1 and let us define

$$L = [\mathbf{w}^1, \dots, \mathbf{w}^n]^T ; \quad R = [\mathbf{v}_1, \dots, \mathbf{v}_n] \quad (3.11)$$

as respectively the left and right constant eigenvector matrices of A with $L = R^{-1}$ and where $\mathbf{v}_i \in TC_{\mathbf{x}}$ are the left column eigenvectors of A and $\mathbf{w}^i \in TC_{\mathbf{x}}^*$ are the right row eigenvectors of A belonging to the dual tangent space $TC_{\mathbf{x}}^*$. Note that row vectors may be considered as linear

¹A tangent space $TC_{\mathbf{x}}$ is the set of all tangent vectors to the smooth manifold C at \mathbf{x} . A tangent bundle TC is the collection of all tangent vectors, along with the information of the point to which they are tangent, i.e. $TC = \{(\mathbf{x}, \mathbf{v}) : \mathbf{x} \in C, \mathbf{v} \in TC_{\mathbf{x}}\}$

functionals transforming vectors into scalars according to the usual row by column product.

Let us define a new variable $\boldsymbol{\eta}(t) = \mathbf{L} \mathbf{a}(t)$ so that $\mathbf{a}(t) = \mathbf{R} \boldsymbol{\eta}(t)$. The system can then be re-written as

$$\mathbf{R} \dot{\boldsymbol{\eta}} = \mathbf{A} \mathbf{R} \boldsymbol{\eta}$$

and premultiplying by \mathbf{L} we fully diagonalize matrix \mathbf{A} :

$$\dot{\boldsymbol{\eta}} = \mathbf{L} \mathbf{A} \mathbf{R} \boldsymbol{\eta} = \Lambda \boldsymbol{\eta} \quad (3.12)$$

where $\Lambda = \text{diag}(\lambda_i)$ is the diagonal eigenvalue matrix of \mathbf{A} , λ_i being such eigenvalues, and $\boldsymbol{\eta}(t=0) = \boldsymbol{\eta}(0)$. Eq. (3.12) has the following solution:

$$\boldsymbol{\eta}(t) = \begin{pmatrix} \eta^1(0)e^{\lambda_1 t} \\ \dots \\ \eta^n(0)e^{\lambda_n t} \end{pmatrix}$$

where

$$\boldsymbol{\eta}(0) = \mathbf{L} \mathbf{a}(0) = \begin{pmatrix} \mathbf{w}^1 \mathbf{a}(0) \\ \dots \\ \mathbf{w}^n \mathbf{a}(0) \end{pmatrix}$$

Reverting to the original vector variable $\mathbf{a}(t)$ we obtain

$$\begin{aligned} \mathbf{a}(t) = \mathbf{R} \boldsymbol{\eta}(t) &= [\mathbf{v}_1, \dots, \mathbf{v}_n] \boldsymbol{\eta}(t) = \sum_{i=1}^n \mathbf{v}_i e^{\lambda_i t} \eta^i(0) \\ &= \sum_{i=1}^n \mathbf{v}_i e^{\lambda_i t} (\mathbf{w}^i \mathbf{a}(0)). \end{aligned} \quad (3.13)$$

In a dual manner we may consider the dynamics of linear functionals $\mathbf{c} \in \text{TC}_x^*$ within the dual tangent space, as they evolve according to the equation $\dot{\mathbf{c}} = \mathbf{c} \mathbf{A}$. As before we may define a new variable $\boldsymbol{\gamma} = \mathbf{c} \mathbf{R}$, itself a linear functional, so that $\mathbf{c} = \boldsymbol{\gamma} \mathbf{L}$ and solve the equation $\dot{\boldsymbol{\gamma}} = \boldsymbol{\gamma} \mathbf{L} \mathbf{A} \mathbf{R} = \boldsymbol{\gamma} \Lambda$ whose solution is

$$\boldsymbol{\gamma}(t) = (\gamma_1(0)e^{\lambda_1 t}, \dots, \gamma_n(0)e^{\lambda_n t})$$

where

$$\boldsymbol{\gamma}(0) = \mathbf{c}(0) \mathbf{R} = (\mathbf{c}(0)\mathbf{v}_1, \dots, \mathbf{c}(0)\mathbf{v}_n)$$

and again reverting to the original variable $\mathbf{c}(t)$ we obtain

$$\begin{aligned}\mathbf{c}(t) = \gamma(t) \mathbf{L} = \gamma(t) [\mathbf{w}^1, \dots, \mathbf{w}^n]^T &= \sum_{i=1}^n \gamma_i(0) e^{\lambda_i t} \mathbf{w}^i \\ &= \sum_{i=1}^n (\mathbf{c}(0) \mathbf{v}_i) e^{\lambda_i t} \mathbf{w}^i.\end{aligned}\quad (3.14)$$

The same results can be obtained by making use of the Magnus series expansion [71, 72, 73] which gives the solution of the skew-product system of Eqs. (3.8), (3.9) in the following form:

$$\mathbf{a}(t) = e^{\Omega(\mathbf{x}(t))} \mathbf{a}(0) \quad (3.15)$$

where the matrix exponential² $e^{\Omega(\mathbf{x}(t))} = e^{\Omega(t, \phi_t(\mathbf{x}_0))} = e^{\Omega(t, \mathbf{x}_0)} : \text{TC}_{\mathbf{x}_0} \rightarrow \text{TC}_{\mathbf{x}(t)}$ is an operator directly mapping the initial vector $\mathbf{a}(0)$ into the vector $\mathbf{a}(t)$ at time t . The matrix function $\Omega(\mathbf{x}(t))$ can be expressed by the Magnus series expansion³

$$\begin{aligned}\Omega(\mathbf{x}(t)) &= \int_0^t \mathbf{J}(\mathbf{x}(\tau)) d\tau \\ &+ \frac{1}{2} \int_0^t \left[\mathbf{J}(\mathbf{x}(\tau)), \int_0^\tau \mathbf{J}(\mathbf{x}(\xi)) d\xi \right] d\tau \\ &+ \frac{1}{4} \int_0^t \left[\mathbf{J}(\mathbf{x}(\tau)), \int_0^\tau \left[\mathbf{J}(\mathbf{x}(\xi)), \int_0^\zeta \mathbf{J}(\mathbf{x}(\zeta)) d\zeta \right] d\xi \right] d\tau \\ &+ \frac{1}{12} \int_0^t \left[\left[\mathbf{J}(\mathbf{x}(\tau)), \int_0^\tau \mathbf{J}(\mathbf{x}(\xi)) d\xi \right], \int_0^\tau \mathbf{J}(\mathbf{x}(\zeta)) d\zeta \right] d\tau + \dots\end{aligned}\quad (3.16)$$

where $[\mathbf{P}, \mathbf{Q}] = \mathbf{P} \mathbf{Q} - \mathbf{Q} \mathbf{P}$ is the commutator of two generic operators.

The terms of the Magnus series following the first give a measure of the non-commutativity of the local jacobian $\mathbf{J}(\mathbf{x}(\tau))$ along the system's trajectory $\mathbf{x}(\tau)$ up to time t and in some sense express the extent to which the generic non linear system is history-dependent. These issues will be thoroughly examined in the next section. In the case of a linear system such terms simply vanish because $\mathbf{J}(\mathbf{x}(t)) = \mathbf{A}$ and therefore jacobians commute exactly along any trajectory. In such case Eq.(3.15) reduces to:

$$\mathbf{a}(t) = e^{\mathbf{A} t} \mathbf{a}(0) \quad (3.17)$$

²A matrix exponential is to be regarded as a function of an operator. In a finite-dimensional space where linear operators are matrices, for example, the matrix exponential is defined as $e^{\mathbf{A}} = f(\mathbf{A}) = \sum_{h=1}^{\infty} \frac{\mathbf{A}^h}{h!}$. With the concept of series expansion any other operator function may be defined.

³The Magnus expansion extends the solution of a scalar linear equation $\dot{a} = j(t)a$ to the case the system is defined in R^N . In the scalar case the solution is trivially $a(t) = a(t=0)e^{\int_0^t j(\tau)d\tau}$

That this is a solution of Eq.(3.10) clearly shows by substituting Eq.(3.17) thus obtaining $d\mathbf{a}/dt = \mathbf{A} e^{\mathbf{A}t} \mathbf{a}(0) = \mathbf{A} \mathbf{a}(t)$. Expanding the matrix exponential and substituting $\boldsymbol{\eta}(t) = \mathbf{L} \mathbf{a}(t)$ (so that $\mathbf{a}(t) = \mathbf{R} \boldsymbol{\eta}(t)$) we obtain the same expression of Eq.(3.13):

$$\begin{aligned}
 \mathbf{a}(t) &= e^{\mathbf{A}t} \mathbf{a}(0) \\
 &= \left(\sum_{h=0}^n \frac{\mathbf{A}^h t^h}{h!} \right) \mathbf{a}(0) \\
 &= \left(\sum_{h=0}^n \frac{\mathbf{A}^h \mathbf{R} t^h}{h!} \right) \boldsymbol{\eta}(0) \\
 &= \left(\sum_{h=0}^n \frac{\mathbf{R} \Lambda^h t^h}{h!} \right) \mathbf{L} \mathbf{a}(0) \\
 &= \mathbf{R} (e^{\Lambda t}) \mathbf{L} \mathbf{a}(0) \\
 &= \sum_{i=1}^n \mathbf{v}_i e^{\lambda_i t} (\mathbf{w}^i \mathbf{a}(0))
 \end{aligned} \tag{3.18}$$

where, being $\mathbf{A} \mathbf{R} = \mathbf{R} \Lambda$, we have substituted $\mathbf{A}^h \mathbf{R} = \mathbf{A}^{h-1} \mathbf{A} \mathbf{R} = \mathbf{A}^{h-1} \mathbf{R} \Lambda = \mathbf{A}^{h-2} \mathbf{R} \Lambda^2 = \dots = \mathbf{R} \Lambda^h$, and being Λ a diagonal matrix⁴ then $(e^{\Lambda t})_j^i = \delta_j^i e^{\lambda_i t}$ where δ_j^i is the Kronecker symbol.

Note that if we choose any right eigenvector of \mathbf{A} as the initial vector, i.e. $\mathbf{a}(0) = \mathbf{v}_k$, then by Eq.(3.18) we have:

$$\mathbf{a}(t) = e^{\mathbf{A}t} \mathbf{v}_k = \mathbf{R} (e^{\Lambda t}) \mathbf{L} \mathbf{v}_k = \sum_{i=1}^n \mathbf{v}_i e^{\lambda_i t} \delta_k^i = \mathbf{v}_k e^{\lambda_k t} \tag{3.19}$$

The latter expression⁵ clearly shows that the set $\{\mathbf{v}_i\}_{i=1}^n$ of right eigenvectors of \mathbf{A} represent an invariant set of directions, or subspaces of the tangent space, under the vector dynamics $\dot{\mathbf{a}} = \mathbf{A} \mathbf{a}$ whose solution is given by Eq.(3.18). Invariance arises because the solution operator $e^{\mathbf{A}t}$, representing the action of vector dynamics for a time t , maps such eigenvectors onto themselves for any t .

The preceding considerations can be generalized more rigorously. We say that vector space E is the direct sum of r subspaces E_1, \dots, E_r if every vector $\mathbf{x} \in E$ can be expressed uniquely as

⁴We have implicitly retained the assumption that \mathbf{A} has real and distinct eigenvalues so that $\Lambda = \text{diag}(\lambda_i)$. This is of course the simplest case for which the change of base $\boldsymbol{\eta}(t) = \mathbf{L} \mathbf{a}(t)$ yields the diagonalization of the operator \mathbf{A} so that $e^{\mathbf{A}t} = \mathbf{R} e^{\Lambda t} \mathbf{L} = \mathbf{R} \text{diag}(e^{\lambda_i t}) \mathbf{L}$. Incidentally for all other cases we will search for a proper change of base $\tilde{\mathbf{a}}(t) = \mathbf{T} \mathbf{a}(t)$ in which the operator attains the similar form $\tilde{\mathbf{A}} = \mathbf{T}^{-1} \mathbf{A} \mathbf{T}$ where $\tilde{\mathbf{A}} = \text{diag}(e^{\tilde{\lambda}_i t})$ is diagonal and $e^{\mathbf{A}t} = \mathbf{T} e^{\tilde{\mathbf{A}}t} \mathbf{T}^{-1} = \mathbf{T} \text{diag}(e^{\tilde{\lambda}_i t}) \mathbf{T}^{-1}$. The latter is also the simplest way of evaluating a matrix exponential.

⁵Note that Eq.(3.19) also shows that the set $\{\mathbf{v}_i\}_{i=1}^n$ of right eigenvectors of \mathbf{A} is also the set of right eigenvectors of the matrix exponential $e^{\mathbf{A}}$ and that its eigenvalues are e^{λ_i}

$\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_r$ where $\mathbf{x}_i \in E_i$. We denote this by

$$E = E_1 \oplus \cdots \oplus E_r \quad (3.20)$$

If every vector $\mathbf{x}_i \in E_i$ can be expressed by means of a basis $\{(\mathbf{e}_i)_j\}_{j=1}^{r_i}$ spanning E_i ($\dim E_i = r_i$), i.e. $\mathbf{x}_i = ((x_i)_1, \dots, (x_i)_{r_i})^T$, then the union of such r bases constitutes a basis for E ($\dim E = \sum_i r_i$) and $\mathbf{x} = ((x_1)_1, \dots, (x_1)_{r_1}, \dots, (x_r)_1, \dots, (x_r)_{r_r})^T$.

If $A : E \rightarrow E$ and $A_i : E_i \rightarrow E_i$, $i = 1, r$ are operators then we say that A is the direct sum of the A_i , i.e.

$$A = A_1 \oplus \cdots \oplus A_r \quad (3.21)$$

if $E = \bigoplus_{i=1}^r E_i$, if each E_i is invariant under A , i.e. $A(E_i) \subset E_i$, and if $A\mathbf{x} = A_i\mathbf{x}$ if $\mathbf{x} \in E_i$. Then, using the basis $\{(\mathbf{e}_i)_j\}_{j=1}^{r_i}$ to represent $\mathbf{x}_i = ((x_i)_1, \dots, (x_i)_{r_i})^T \in E_i$, the operator A_i is a $r_i \times r_i$ matrix. Because the union of each basis of E_i is a basis for E then the operator A may be represented in such union basis by the diagonal composition of matrices A_i all other entries being zero (because there is no mapping of one subspace to another) :

$$A = \begin{bmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{bmatrix}$$

This being said, then it follows directly that the eigenspaces of operator $A : TC \rightarrow TC$ in vector dynamics $\dot{\mathbf{a}} = A\mathbf{a}$ (or equivalently of the solution operator e^{At}), provide an invariant direct sum decomposition of the tangent space TC and that A is diagonal in the right eigenvector basis provided eigenvalues are real and distinct. If λ_i are such real and distinct eigenvalues and \mathbf{v}_i the corresponding eigenvectors , then TC has an invariant direct sum decomposition and correspondingly A has a direct sum decomposition:

$$TC = TC_1 \oplus \cdots \oplus TC_n \quad (3.22)$$

$$A = A_1 \oplus \cdots \oplus A_n \quad (3.23)$$

where each subspace TC_i is one dimensional and spanned by \mathbf{v}_i and is invariant under A , i.e. $A(TC_i) \subset TC_i$ and each operator $A_i : TC_i \rightarrow TC_i$ is such that $A\mathbf{v}_i = A_i\mathbf{v}_i$ if $\mathbf{v}_i \in TC_i$. Note that the existence of such decomposition for A implies that in the right eigenvector base it

can be written as $A = \text{diag}(A_i) = \text{diag}(\lambda_i)$. Equivalently, in such base the solution operator $e^{At} = e^{A_1 t} \oplus \dots \oplus e^{A_n t}$ will attain the diagonal form $e^{At} = \text{diag}(e^{\lambda_i t})$.

In fact this can be generalized to the case where A has distinct eigenvalues not necessarily all real, say $\lambda_1, \dots, \lambda_r$ real with corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_r$ and the remaining $\mu_1, \bar{\mu}_1, \dots, \mu_i, \bar{\mu}_i$ nonreal with corresponding eigenvectors $\mathbf{f}_1, \bar{\mathbf{f}}_1, \dots, \bar{\mathbf{f}}_i, \mathbf{f}_i$. In this case we have the following invariant direct sum decomposition of TC and the corresponding direct sum decomposition for A

$$\text{TC} = \text{TC}_a \oplus \text{TC}_b \quad A = A_a \oplus A_b \quad (3.24)$$

where TC_a is spanned by $\{\mathbf{v}_k\}$ and TC_b is spanned by $\{\mathbf{f}_k, \bar{\mathbf{f}}_k\}$ and where $A_a : \text{TC}_a \rightarrow \text{TC}_a$ has real distinct eigenvalues $\lambda_1, \dots, \lambda_r$ and $A_a(\text{TC}_a) \subset \text{TC}_a$ and $A_b : \text{TC}_b \rightarrow \text{TC}_b$ has nonreal distinct eigenvalues $\mu_1, \bar{\mu}_1, \dots, \mu_i, \bar{\mu}_i$ and $A_b(\text{TC}_b) \subset \text{TC}_b$.

In general, therefore, any dynamical system such as (3.10) may be rewritten as two systems $\dot{\mathbf{a}}_a = A_a \mathbf{a}_a$ and $\dot{\mathbf{a}}_b = A_b \mathbf{a}_b$ where $\mathbf{a}_a \in \text{TC}_a$ and $\mathbf{a}_b \in \text{TC}_b$. We have shown before that A_a is diagonalizable by expressing it in the right eigenvector base $\{\mathbf{v}_k\}$ so that $A_a = \text{diag}(\lambda_i)$. Similarly it can be shown that there exists an invariant direct sum decomposition of $\text{TC}_b = \oplus_i \text{TC}_{bi}$ and a corresponding direct sum decomposition of $A_b = \oplus_i A_{bi}$ such that each TC_{bi} is twodimensional and A_{bi} has eigenvalues $\mu_i, \bar{\mu}_i$. If $\mu_i = a + ib$ and the corresponding eigenvector is $\phi_i = \mathbf{u} + i\mathbf{v}$ with $\mathbf{u}, \mathbf{v} \in \mathbb{R}^2$ then operator A_{bi} may be expressed in the basis $\{\mathbf{u}, \mathbf{v}\}$, i.e. using the change of base $Q = [\mathbf{u}, \mathbf{v}]$, as

$$A = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

which has the following matrix exponential form

$$e^A = e^a \begin{bmatrix} \cos b & -\sin b \\ \sin b & \cos a \end{bmatrix}$$

which can be used in Eq.(3.17) to yield the solution of the dynamical system.

An even greater generalization can be achieved by recalling [74] that any operator may be decomposed uniquely as $A = S + N$ where S is always diagonalizable into a canonical form (whose exponential can be calculated as shown before), N is a nilpotent operator (i.e $N^r = 0$ indicates that N is nilpotent of order r) and where S and N commute, i.e. $[S, N] = SN - NS = 0$. This is a

consequence of the fact that the vector space E on which A operates is the direct sum of the *generalized eigenspaces*⁶ of A (invariant under A) each having a dimension equal to the multiplicity of the corresponding eigenvalue. So if A has only one eigenvalue λ of multiplicity $n = \dim E$, by placing $S = \lambda I$ and $N = A - \lambda I$ we find $A = S + N$ and $[S, N] = 0$. Moreover N is nilpotent of order n , i.e. $N^n = 0$, because $E = \text{Ker } N^n$ this implying that, being $N^n \mathbf{v}_1 = 0, \dots, N^n \mathbf{v}_n = 0$ i.e $N^n(\mathbf{v}_1 + \dots + \mathbf{v}_n) = 0$, the only way for $\{\mathbf{v}_i\}_{i=1}^n$ to be a set of n linearly independent vectors and hence span E is that $N^n = 0$. Now, because $e^{B+C} = e^B e^C$ is true only if B and C commute, then we can write $e^A = e^S e^N = e^{\lambda I} \sum_{k=1}^{n-1} \frac{N^k}{k!}$ which is easily determined. We should emphasize that for generic operators with distinct eigenvalues and given algebraic multiplicities the usual procedure is finding a basis that puts the operator in the so called Jordan form (or real canonical form) where S is diagonal (or block diagonal for complex eigenvalues) and N is in nilpotent canonical form (i.e. with 1's just below the diagonal and 0's elsewhere). The procedure is iterative and beyond the scope of the present chapter.

In conclusion, given the linear dynamical system $\dot{\mathbf{a}} = A \mathbf{a}$, the (generalized) eigenspaces of A (or equivalently of the solution e^{At}) give us a complete geometrical and timescale characterization of the dynamical system itself as they represent the invariant decomposition of the tangent space, meaning that any vector within a subspace concurring to such decomposition will remain in such subspace during its evolution. Moreover the corresponding eigenvalues will characterize the timescale spectrum of the system as each eigenvector will be mapped by the solution operator into $e^{At} \mathbf{v}_i = e^{\lambda t} \mathbf{v}_i$ (or at most into $e^{\lambda t} e^{Nt} \mathbf{v}_i$ with N nilpotent). Note that this decomposition is unique, providing the most efficient setting in which the linear dynamical system can be studied as it allows us to establish an unbiased temporal ordering of the slow/fast and stable/unstable components of the dynamics based on the intrinsic time scales without any 'mode' mixing between such components. Any other setting would induce mode mixing and would not allow the definition of a spectrum of intrinsic timescales. Finding the system-invariant geometrical setting and hence the spectrum of intrinsic timescales is exactly the prerequisite needed to achieve a simplification/reduction of the system as this would allow for the unambiguous neglection of the fastest stable modes leaving to the 'most relevant' modes the task of describing the asymptotic behaviour of the system.

How can these notions be applied to non linear systems? In particular, can we define a system-invariant decomposition for the non linear system? Can we define an unambiguous spectrum of intrinsic time scales? The next section will be devoted to finding answers to these questions.

⁶We define *generalized eigenspace* of A related to the eigenvalue λ_k the subspace $E_k = \text{Ker}(A - \lambda_k I)^{n_k} = \{\mathbf{v} \in E_k \mid (A - \lambda_k I)^{n_k} \mathbf{v} = 0\} \subset E$

3.3 Nonlinear dynamics

In the following sections, we will tackle the issues that were brought forward in the introduction. In particular we will show how the vector dynamics within the tangent bundle, described by the variational equation (3.7) coupled to the dynamical system (3.1), constitutes the basic tool to define invariance properties for the system itself. Some concepts that will be explored in this section can also be found in the Skodje and Davis work [75], in the context of chemical kinetic simplification, and in the Mease et. al. work [76], in the context of control theory.

3.4 Vector dynamics and invariance properties

We have seen how the evolution of normalized perturbations induced by a generic non linear system on a smooth manifold C is governed by the skew-product system of the original autonomous dynamical system and the linear non autonomous variational equation which we may re-write as:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) \quad \mathbf{x}(t=0) = \mathbf{x}_o \quad (3.25)$$

$$\dot{\mathbf{a}} = \mathbf{J}(\mathbf{x}(t)) \mathbf{a} \quad \mathbf{a}(t=0) = \mathbf{a}(0) \quad (3.26)$$

where $\mathbf{a}(0) \in TC_{\mathbf{x}_o}$ is a vector belonging to the tangent space $TC_{\mathbf{x}_o}$ at $\mathbf{x}_o \in C$ and $\mathbf{J}(t) = \mathbf{J}(\mathbf{x}(t)) = \mathbf{J}(\phi_t(\mathbf{x}_o))$ is the Jacobian matrix of $\mathbf{f}(\mathbf{x})$. We have seen in the previous section (Eq.(3.15)) that the formal solution of Eq.(3.26) may be written as:

$$\mathbf{a}(t) = e^{\Omega(t, \mathbf{x}_0)} \mathbf{a}(0) \quad \mathbf{a}(t) \in TC_{\phi_t(\mathbf{x}_o)} \quad (3.27)$$

where the matrix function $\Omega(t, \mathbf{x}_0)$ is given by the Magnus series expansion of Eq.(3.17). The Magnus expansion is interesting as it highlights the important role of the non-commutativity of the local Jacobian matrices along the trajectory. This feature is absent in linear systems (for which the matrix function $\Omega(t, \mathbf{x}_0) = A t$ and the solution operator reduces to e^{At}) and it is this very feature that draws a line between linear and non linear systems. In a non linear system all the commutation terms following the first in the Magnus expansion are non-zero because local Jacobians are a function of $\mathbf{x}(t)$ owing to the non-autonomous nature of the variational equation. To give us an idea of the influence of non-commutativity of local Jacobians on the solution operator $e^{\Omega(t, \mathbf{x}_0)}$ it might be instructive to examine the behavior of a characteristic commutator element appearing in the Magnus

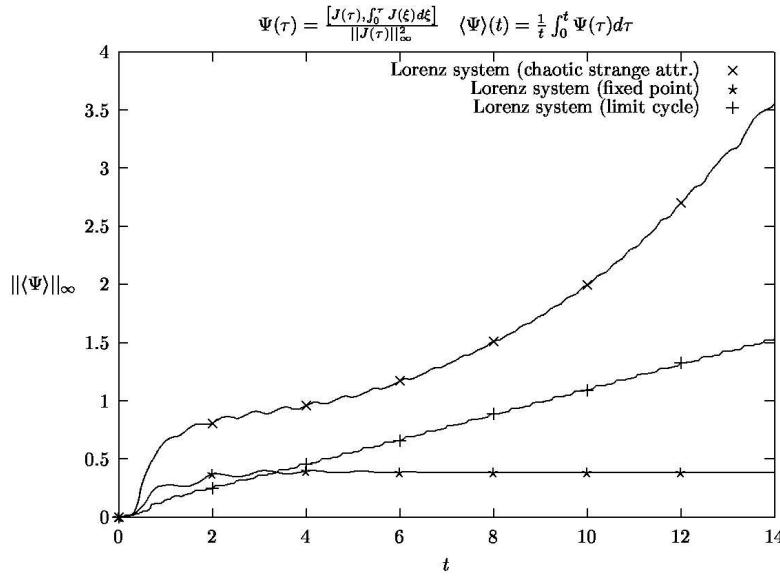


Fig. 3.2: The role of non-commutativity along a trajectory for the Lorenz system

series expansion. Let us consider the Lorenz system:

$$\begin{aligned}
 \dot{x} &= \sigma(y - x) \\
 \dot{y} &= rx - y - xz \\
 \dot{z} &= xy - cz
 \end{aligned} \tag{3.28}$$

where $\sigma = 10$, $c = 8/3$ and r was varied so as to embrace the three representative cases in which the system exhibits a stationary equilibrium point ($r = 10$), a limit cycle ($r = 260$) and a chaotic attractor ($r = 28$). For the three cases the normalized matrix function $\Psi(\tau) = [J(\tau), \int_0^\tau J(\xi)d\xi] / \|J(\tau)\|_\infty^2$ was evaluated and time-averaged along a trajectory to yield $\langle \Psi \rangle(t) = \frac{1}{t} \int_0^t \Psi(\tau)d\tau$. The norm $\|\langle \Psi \rangle\|_\infty$ is reported in Fig.(3.2) as a function of time along the trajectory. The extent of the non-commutativity in the three cases is different and could be used as a classification for the different behaviors exhibited by a dynamical system. In the equilibrium point case $\|\langle \Psi \rangle\|_\infty$ seems to saturate to a constant value this implying that commutators tend to zero in norm as the trajectory approaches the equilibrium point. In the limit cycle case the same quantity seems to be growing linearly with time, this implying that the role of non-commutativity is self-similar along each period T i.e. $\frac{1}{T} \int_0^T \Psi(\tau)d\tau = \Psi_T = \text{const.}$ and $\langle \Psi \rangle(nT) = n\Psi_T$. In the chaotic case, on the other hand, the growth is faster than in the limit cycle case.

Discrete dynamical systems provide an effective visualization of the role of non-commutativity on the choice of the pointwise invariant directions and the extent to which these can differ from local

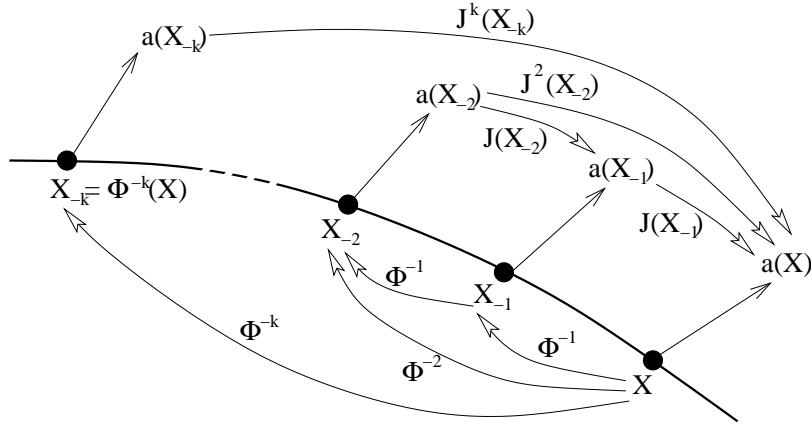


Fig. 3.3: Vector dynamics for a discrete system

Jacobian eigenspaces. Let us therefore consider a generic map $\Phi : C \rightarrow C \in R^N$ and the vector dynamics associated to it:

$$\mathbf{x}_{n+1} = \Phi(\mathbf{x}_n) \quad (3.29)$$

$$\mathbf{a}(\mathbf{x}_{n+1}) = J(\mathbf{x}_n)\mathbf{a}(\mathbf{x}_n) \quad (3.30)$$

where $J(\mathbf{x}) = \partial\Phi(\mathbf{x})/\partial\mathbf{x}$. Consider now the generic point $\mathbf{x} \in C$ and the vector $\mathbf{a}(\mathbf{x}) \in TC_{\mathbf{x}}$. Evidently, an arbitrarily long compositions of local Jacobians will have mapped $\mathbf{a}(\mathbf{x}_{-k}) \in TC_{\mathbf{x}_{-k}}$ into $\mathbf{a}(\mathbf{x})$ (see Fig.(3.3)):

$$\mathbf{a}(\mathbf{x}) = J(\mathbf{x}_{-1})\mathbf{a}(\mathbf{x}_{-1}) = \quad (3.31)$$

$$= J(\mathbf{x}_{-1})J(\mathbf{x}_{-2})\mathbf{a}(\mathbf{x}_{-2}) = \quad (3.32)$$

$$= J(\mathbf{x}_{-1}) \cdots J(\mathbf{x}_{-k})\mathbf{a}(\mathbf{x}_{-k}) = \quad (3.33)$$

$$= J(\Phi^{-1}(\mathbf{x})) \cdots J(\Phi^{-k}(\mathbf{x}))\mathbf{a}(\mathbf{x}_{-k}) = \quad (3.34)$$

$$= J^k(\Phi^{-k}(\mathbf{x}))\mathbf{a}(\mathbf{x}_{-k}) \quad (3.35)$$

where we have defined the k-fold composition of Φ as $\Phi^k = \Phi \circ \cdots \circ \Phi$ and $J^k = \partial\Phi^k/\partial\mathbf{x}$. Clearly the system-invariant directions at \mathbf{x} will have little or nothing to do with the eigenspaces of the local Jacobian $J(\mathbf{x})$ but rather will have to be sought as the eigenspaces associated with an arbitrarily long composition of local Jacobians $J(\mathbf{x}_{-1}) \cdots J(\mathbf{x}_{-k})$ embracing the history up to point \mathbf{x} . Of course if these Jacobians commuted they would share the same eigenvectors ⁷ and thus local invariance would also imply global invariance (which is what happens in linear systems).

⁷Given two commuting matrices $RM = MR$ and $M\mathbf{v} = \lambda\mathbf{v}$ then $M(R\mathbf{v}) = R(M\mathbf{v}) = R(\lambda\mathbf{v}) = \lambda(R\mathbf{v})$ and therefore $R\mathbf{v} = \theta\mathbf{v}$ where θ is some constant, which shows that M and R share the same eigenvectors.

The operator $V(\mathbf{x}, k) = J^k(\Phi^{-k}(\mathbf{x}))$ mapping vectors in $TC_{\Phi^{-k}(\mathbf{x})}$, i.e vectors in the tangent space at the k -th pre-image of \mathbf{x} , into vectors in $TC_{\mathbf{x}}$ is evidently not only a pointwise function of \mathbf{x} but also of k , that is a function of how far back in discrete time we consider the pre-image of \mathbf{x} . By defining the eigenvectors of $V(\mathbf{x}, k)$ as the set $\{\mathbf{e}_h(\mathbf{x}, k)\}$ with $h = 1, N$, then the pointwise subspaces of $TC_{\mathbf{x}}$ in a generic point \mathbf{x} of the phase space which are invariant under vector dynamics are obtained from such eigenvectors by taking the limit $k \rightarrow \infty$, i.e. by considering an infinitely long sequence of pre-images of \mathbf{x} . We can therefore define such invariant directions [77] as $\mathbf{e}_h(\mathbf{x}) = \lim_{k \rightarrow \infty} \mathbf{e}_h(\mathbf{x}, k)$, provided there exists a diverging sequence $k_1 < k_2 < \dots < k_n$ such that the corresponding sequence of eigenvectors $\{\mathbf{e}_h(\mathbf{x}, k_1)\}, \dots, \{\mathbf{e}_h(\mathbf{x}, k_n)\}$ converges towards a limit vector basis $\{\mathbf{e}_h(\mathbf{x})\}$.

For continuous time dynamical systems we can reach the same conclusion. With reference to Eq.(3.27) we must consider the action of the operator $e^{\Omega(t, \mathbf{x}_0)}$ starting from the counterimage point $\mathbf{x}_0 = \Phi_{-t}(\mathbf{x})$:

$$\mathbf{a}(t) = e^{\Omega(t, \Phi_{-t}(\mathbf{x}))} \mathbf{a}(0). \quad (3.36)$$

For the operator $V : TC_{\Phi_{-t}(\mathbf{x})} \rightarrow TC_{\mathbf{x}}$ defined as $V(\mathbf{x}, t) = e^{\Omega(t, \Phi_{-t}(\mathbf{x}))}$ we can define a set of pointwise eigenvectors $\{\mathbf{e}_h(\mathbf{x}, t)\}$ and equivalently as before, provided the eigenvectors converge, we can define the pointwise invariant vectors (subspaces) as $\mathbf{e}_h(\mathbf{x}) = \lim_{t \rightarrow \infty} \mathbf{e}_h(\mathbf{x}, t)$. We may therefore imagine for the tangent space at any point $\mathbf{x} \in C$ an invariant decomposition $TC_{\mathbf{x}} = \bigoplus_{h=1}^n E^h(\mathbf{x})$ where each subspace $E^h(\mathbf{x})$ is spanned by $\mathbf{e}_h(\mathbf{x})$.

It is now worth pointing out that for the stationary equilibrium case of the Lorenz system, the role of non-commutativity is, as we have seen, bounded (Fig.(3.2)) as the commutators tend to zero in norm as the trajectory approaches the equilibrium point. At this point we could add that the counterimage of the equilibrium point is the equilibrium point itself as it represents an invariant set: $\mathbf{x}_0 = \mathbf{x}_s$ and therefore $\mathbf{x}_0 = \phi_{-t}(\mathbf{x}_s) = \mathbf{x}_s = \phi_0(\mathbf{x}_s)$. The Magnus function would hence reduce to $\Omega(t, \phi_{-t}(\mathbf{x}_s)) = \Omega(t, \mathbf{x}_s) = \mathbf{J}(\mathbf{x}_s)t$ all the other terms in the expansion being zero as $\mathbf{J}(\mathbf{x}(t)) = \mathbf{J}(\mathbf{x}_s) = \text{const.}$. This shows clearly that $V(\mathbf{x}, t) = e^{\mathbf{J}(\mathbf{x}_s)t}$ and hence the invariant subspaces at the equilibrium point coincide with the local eigenspaces of the Jacobian just as the Hartman-Grobman theorem states.

In the case $\mathbf{x} = \mathbf{x}^* \in \gamma$ where γ is a limit cycle, we could imagine of replacing the continuous dynamical system with a discrete one $\mathbf{x}^* = \Phi_T(\mathbf{x}^*)$ where Φ_T is the return map and T the period. Vector dynamics would then be governed by $\mathbf{a}(\mathbf{x}^*(T)) = J_T(\mathbf{x}^*)\mathbf{a}(\mathbf{x}^*(0))$ where

$J_T = \partial\Phi_T(\mathbf{x})/\partial\mathbf{x}$ is the Jacobian of the return map at \mathbf{x} . Then the operator V would reduce to $V(\mathbf{x}^*, t) = V(\mathbf{x}^*, nT) = J_T^n(\Phi_T^{-n}(\mathbf{x}^*)) = J_T^n(\mathbf{x}^*) = \text{const.}$ Since the eigenvectors of $J_T^n(\mathbf{x}^*)$ coincide with the eigenvectors of $J_T(\mathbf{x}^*)$ then the pointwise invariant directions on the limit cycle are simply given by the local eigenvectors of the Jacobian $J_T(\mathbf{x})$ of the return (Poincaré') map, where $\mathbf{x} \in \gamma$, and where the corresponding eigenvalues are known as Floquet exponents (see [67]). The foregoing issues can be illustrated by means of an analytical example referring to a prototypical system possessing a slow manifold and an equilibrium point: the two-dimensional Davis-Skodje model [36, 35]. Note that Skodje and Davis in [75] use their model to illustrate their 'Global Eigenvalue Method' on which part of the present issues are grounded, obtaining similar results. The Davis-Skodje model reads:

$$\frac{dy}{dt} = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right) - \frac{z}{(1+z)^2}, \quad \frac{dz}{dt} = -z \quad (3.37)$$

for which an analytical expression for the flow exists:

$$\mathbf{x}(t) = \phi_t(\mathbf{x}_o) = \left(\frac{z_o}{e^t + z_o} + e^{-t/\varepsilon} \left(y_o - \frac{z_o}{1+z_o} \right), e^{-t} z_o \right)^T, \quad (3.38)$$

where $\mathbf{x}_o = (y_o, z_o)^T$. This allows us to define \mathbf{x}_o by means of the pre-image of \mathbf{x} :

$$\mathbf{x}_o = \phi_{-t}(\mathbf{x}(t)) = \left(\frac{z}{e^{-t} + z} + e^{t/\varepsilon} \left(y - \frac{z}{1+z} \right), e^t z \right)^T. \quad (3.39)$$

Vector dynamics is described by the linear non-autonomous system:

$$\dot{a}_y = -\frac{1}{\varepsilon} a_y + \frac{1+z-\varepsilon+z\varepsilon}{(1+z)^3 \varepsilon} a_z, \quad \dot{a}_z = -a_z \quad (3.40)$$

which admits a closed-form solution given by

$$\begin{aligned} a_y(t) &= \frac{e^{-t/\varepsilon} \{(1+z_o)^2 (e^t + z_o)^2 a_y(0)\}}{(1+z_o)^2 (e^t + z_o)^2} + \\ &\quad + \frac{e^{-t/\varepsilon} \{[-(e^t + z_o)^2 + e^{t(1+1/\varepsilon)} (1+z_o)^2] a_z(0)\}}{(1+z_o)^2 (e^t + z_o)^2} \\ a_z(t) &= e^{-t} a_z(0). \end{aligned} \quad (3.41)$$

The latter is by definition in the form $\mathbf{a}(t) = e^{\Omega(t, \mathbf{x}_0)} \mathbf{a}(0)$ where $\mathbf{a}(t) = (a_y(t), a_z(t))^T$. In order to obtain a closed form expression of the operator $V(\mathbf{x}, t) = e^{\Omega(t, \Phi_{-t}(\mathbf{x}))}$ we will now substitute $\mathbf{x}_o = (y_o, z_o)^T$, i.e. Eq.(3.39), into Eq.(3.41) thus obtaining

$$V(\mathbf{x}, t) = e^{\Omega(t, \Phi_{-t}(\mathbf{x}))} = \begin{pmatrix} e^{-t/\varepsilon} & q(t, z) \\ 0 & e^{-t} \end{pmatrix} \quad (3.42)$$

$$q(t, z) = \frac{e^{-t} + 2z + e^t z^2 - e^{-t/\varepsilon}(1+z)^2}{(1+z)^2 (1+e^t z)^2}. \quad (3.43)$$

The eigenvalues of $V(\mathbf{x}, t)$ are $\lambda_1(t) = e^{-t/\varepsilon}$ and $\lambda_2(t) = e^{-t}$ whereas the associated normalized eigenvectors are respectively

$$\mathbf{e}_1(\mathbf{x}, t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2(\mathbf{x}, t) = \begin{pmatrix} \frac{Q(t, z)}{\sqrt{1+Q(t, z)^2}} \\ \frac{1}{\sqrt{1+Q(t, z)^2}} \end{pmatrix} \quad (3.44)$$

where $Q(t, z) = q(t, z)/(e^{-t} - e^{-t/\varepsilon})$. It can be easily shown that $\lim_{t \rightarrow \infty} Q(t, z) = 1/(1+z)^2$ and therefore for $t \rightarrow \infty$, $\mathbf{e}_1(\mathbf{x}, t)$ and $\mathbf{e}_2(\mathbf{x}, t)$ converge to

$$\mathbf{e}_1(\mathbf{x}) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2(\mathbf{x}) = \begin{pmatrix} \frac{1}{(1+z)^2} \frac{1}{\sqrt{1+(1/(1+z)^4)}} \\ \frac{1}{\sqrt{1+(1/(1+z)^4)}} \end{pmatrix} \quad (3.45)$$

which represent the system's pointwise invariant subspaces. Invariance can be verified by substituting in turn $\mathbf{a}(0) = \mathbf{e}_1(\mathbf{x}_o) = (1, 0)^T$ and $\mathbf{a}(0) = \mathbf{e}_2(\mathbf{x}_o) = (1, 1/(1+z_o)^2)^T$ as initial conditions into the vector dynamics Eq(3.41) thus obtaining in $TC_{\mathbf{x}}$ respectively

$$\mathbf{a}(t) = e^{-t/\varepsilon} \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{a}(t) = e^{-t} \begin{pmatrix} \frac{1}{(1+z)^2} \\ 1 \end{pmatrix} \quad (3.46)$$

Note that we have substituted $z_o = ze^t$. Note also⁸ that each vector is multiplied respectively by the eigenvalues $\lambda_1(t) = e^{-t/\varepsilon}$ and $\lambda_2(t) = e^{-t}$ corresponding to $\mathbf{e}_1(\mathbf{x})$ and $\mathbf{e}_2(\mathbf{x})$. This means that $\tau_1 = \varepsilon$ and $\tau_2 = 1$ may be considered as the intrinsic timescales of the system: any unit norm vector $\mathbf{v}(\mathbf{x}_o) \in \mathbf{e}_h(\mathbf{x}_o)$ will be mapped into a vector belonging to $\mathbf{e}_h(\mathbf{x})$ its norm being multiplied $\lambda_h(\mathbf{x}(t))$. It is intuitive that by following the natural evolution of vectors within invariant subspaces, if the time-averaged value of $\lambda_h(\mathbf{x}(t))$ is negative it will be safe to define the corresponding invariant subspace as stable (and vice versa), because vectors within such subspace will have contracted.

The invariant subspaces also provide the most effective basis onto which the system's vector field

⁸Before we give a definition of Lyapunov exponents, it is instructive to note that if $\mathbf{a}(0) = \mathbf{e}_1(\mathbf{x}_o)$ and $|\mathbf{e}_1(\mathbf{x}_o)| = 1$ then $\frac{1}{t} \lim_{t \rightarrow \infty} |\mathbf{a}(t)| = -1/\varepsilon$ and if $\mathbf{a}(0) = \mathbf{e}_2(\mathbf{x}_o)$ and $|\mathbf{e}_2(\mathbf{x}_o)| = 1$ then $\frac{1}{t} \lim_{t \rightarrow \infty} |\mathbf{a}(t)| = -1$. $-1/\varepsilon$ and -1 are in fact the Lyapunov exponents of the system. Note that they are independent of \mathbf{x}_o .

$\mathbf{g}(\mathbf{x})$ (source term) can be projected. Utilizing the classical CSP approach, let us rewrite Eq(3.37) in the form $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) = f^1(\mathbf{x})\mathbf{e}_1(\mathbf{x}) + f^2(\mathbf{x})\mathbf{e}_2(\mathbf{x})$ where the projections $f^1(\mathbf{x})$ and $f^2(\mathbf{x})$ are found through the use of the dual vectors (linear functionals) $\mathbf{e}^1(\mathbf{x})$ and $\mathbf{e}^2(\mathbf{x})$ (dual to \mathbf{e}_1 and \mathbf{e}_2), i.e. $f^1(\mathbf{x}) = \mathbf{e}^1 \mathbf{g}$ and $f^2(\mathbf{x}) = \mathbf{e}^2 \mathbf{g}$. It is straight forward to find

$$f^1(\mathbf{x}) = \frac{1}{\varepsilon} \left(-y + \frac{z}{1+z} \right), \quad f^2(\mathbf{x}) = -z \quad (3.47)$$

If we wish to follow the evolution of the two projections along a trajectory we should substitute $\mathbf{x} = \Phi_t(\mathbf{x}_o)$ to obtain:

$$f^1(\mathbf{x}) = \frac{1}{\varepsilon} \left(-y_o + \frac{z_o}{1+z_o} \right) e^{-t/\varepsilon} = f^1(\mathbf{x}_o) e^{-t/\varepsilon}, \quad f^2(\mathbf{x}) = -z_o e^{-t} = f^2(\mathbf{x}_o) e^{-t}. \quad (3.48)$$

which can be rewritten as

$$\begin{pmatrix} f^1(\mathbf{x}) \\ f^2(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} e^{-t/\varepsilon} & 0 \\ 0 & e^{-t} \end{pmatrix} \begin{pmatrix} f^1(\mathbf{x}_o) \\ f^2(\mathbf{x}_o) \end{pmatrix} \quad (3.49)$$

which shows that the evolution of each projection is independent and fully uncoupled from the evolution of the other. This is a direct consequence of the invariance properties of the subspaces $\mathbf{e}_1(\mathbf{x})$ and $\mathbf{e}_2(\mathbf{x})$ onto which $\mathbf{g}(\mathbf{x})$ was projected, each projection being mapped into the same invariant subspace, its norm being multiplied by the corresponding eigenvalue λ_h .

The knowledge of an unambiguous time scale spectrum is essential for the definition of an invariant slow manifold, let us see why. In the present example for $\varepsilon \ll 1$ we can affirm without doubt that the projection $f^1(\Phi_t(\mathbf{x}_o))$ of $\mathbf{g}(\mathbf{x})$ onto the subspace $\mathbf{e}_1(\mathbf{x})$ decreases faster than the other projection $f^2(\Phi_t(\mathbf{x}_o))$, this being true along any given trajectory. We should keep in mind that the latter statement is possible only because Eq.(3.49) is uncoupled: any projection basis other than the invariant basis $\mathbf{e}_h(\mathbf{x})$ would lead to a coupled system and hence to the impossibility of defining an intrinsic time scale spectrum. This being said we could term $\mathbf{e}_1(\mathbf{x})$ the fast invariant direction and $\mathbf{e}_2(\mathbf{x})$ the slow invariant direction. Supposing the analysis of a system's intrinsic time scale spectrum establishes M fast invariant direction and $N - M$ slow invariant direction, then we can imagine the following invariant decomposition of the tangent space $TC_{\mathbf{x}}$ in every point $\mathbf{x} \in C$:

$$TC_{\mathbf{x}} = E^{fast}(\mathbf{x}) \oplus E^{slow}(\mathbf{x}), \quad E^{fast}(\mathbf{x}) = \bigoplus_{h=1}^M E^h(\mathbf{x}), \quad E^{slow}(\mathbf{x}) = \bigoplus_{h=M+1}^N E^h(\mathbf{x}) \quad (3.50)$$

where each subspace $E^h(\mathbf{x})$ is spanned by the invariant direction $\mathbf{e}_h(\mathbf{x})$.

The slow invariant manifold $\mathcal{M} \subset C$ is in general a lower-dimensional subspace of C , or a graph in C , which can be defined as the locus of phase space points $\mathbf{x} \in C$ where the fastest intrinsic timescales are fully exhausted. Having defined such intrinsic timescales through the evolution of vectors within the invariant subspaces, then \mathcal{M} can be defined as the locus of points $\mathbf{x} \in C$ such that $\mathbf{g}(\mathbf{x})$ bears no component on the fast subspace $E^{fast}(\mathbf{x})$

$$\mathcal{M} = \{\mathbf{x} \in C | \mathbf{g}(\mathbf{x}) \in E^{fast}(\mathbf{x})\} = \{\mathbf{x} \in C | f^h(\mathbf{x}) = 0, h = 1, M\}. \quad (3.51)$$

In the case of the Davis-Skodje example, the locus $f^1(\mathbf{x}) = 0$ is, from Eq.(3.47), evidently

$$y = \frac{z}{1+z} \quad (3.52)$$

Note that the invariance of \mathcal{M} under the action of the dynamical system $\mathbf{x} = \Phi_t(\mathbf{x}_o)$ follows directly from Eq.(3.49) and its being uncolupled. So if $\mathbf{x}_o \in \mathcal{M}$ then $f^1(\mathbf{x}_o) = 0$ and from Eq.(3.49) it follows that $f^1(\Phi_t(\mathbf{x}_o)) = f^1(\mathbf{x}) = 0$ from which we conclude that $\mathbf{x} \in \mathcal{M}$.

The preceding issues and in particular the determination of a pointwise invariant decomposition for the system through the eigenstructure of $V(\mathbf{x}, t)$, has been achieved independently of the singularly perturbed nature of the Davis-Skodje system, that is of the existing fast/slow timescale separation (in this case, independently of parameter ε). This is quite remarkable since perturbative methods like CSP or the Fraser-Rousel method base their efficiency (and their success) solely on the singularly perturbed nature of the system and hence on the extent of the fast/slow timescale separation. This also suggests that for systems that do not have a singularly perturbed nature, such methods should fail.

Fig.(3.4)(a) displays the norm of the fast projection f^1 along a trajectory for the Davis-Skodje system. When using the invariant subspaces as the projection basis, $f^1(\mathbf{x}(t))$ goes to zero (see Eq.(3.48)) at the same time as the trajectory reaches the slow manifold ($|y(t) - \tilde{y}(t)|_{\mathcal{M}} = 10^{-13}$), that is within an intrinsic fast timescale. On the contrary, when using local eigenvectors of the Jacobian matrix or a refined CSP basis as the projection basis, residual coupling with f^2 makes it an ambiguous task that of defining an intrinsic fast timescale. Moreover the resulting slow manifold, defined as $\tilde{\mathcal{M}} = \{\mathbf{x} \in C | \tilde{f}^h(\mathbf{x}) = 0, h = 1, M\}$ where the \tilde{f}^h 's are defined with a projection basis other than the invariant one, would only approximate \mathcal{M} of Eq.(3.52) and would not be invariant for the system's dynamics. Fig.(3.4)(b) shows the disalignement of one of the two⁹ normalized

⁹The other is trivially $\mathbf{v}_2 = (1, 0)^T$

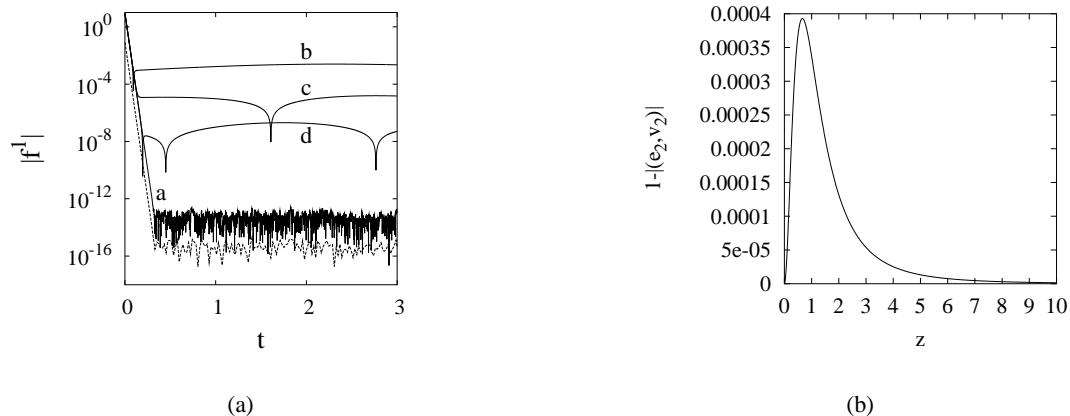


Fig. 3.4: Time behavior of the fast component $|f^1|$ along a system trajectory (starting point $(y_o, z_o)^T = (1, 10)^T$) for the Davis-Skodje model Eq. (3.37) (with $\varepsilon = 0.01$). (a) using the invariant set of vectors/covectors, ;(b) using eigenvectors of the local Jacobian matrix; (c) using CSP vectors/covectors with 1 refinement (starting from eigenvectors of the local Jacobian matrix); (d) using CSP vectors/covectors with 2 refinements (starting from eigenvectors of the local Jacobian matrix). The broken line represents the behavior of $|y(t) - \hat{y}(t)|_{\mathcal{M}}$.

eigenvectors of the local Jacobian from the normalized locally invariant subspace $\mathbf{e}_2(\mathbf{x})$, where $(\mathbf{e}_2, \mathbf{v}_2)$ indicates a scalar product.

3.5 A Lagrangian approach

In the previous section we have shown, with the help of the Davis-Skodje system, that the pointwise invariant decomposition of the tangent space TC_x at a point x can be achieved from the eigensystem of the operator $V(x, t) = e^{\Omega(t, \Phi_{-t}(x))}$ mapping vectors belonging to $TC_{\Phi_{-t}(x)}$ to vectors belonging to TC_x in the limit $t \rightarrow \infty$. This can be regarded as an "Eulerian" approach in that the knowledge of operator $V(x, t)$ allows us to determine the invariant subspaces pointwisely. The system's vector field $\mathbf{g}(x)$ at TC_x may be projected onto such invariant subspaces $\{\mathbf{e}_h(x)\}$ and an evolution equation for the projections $f^h(x)$ starting from $f^h(x_o)$ may be obtained such as that of Eq.(3.49). This equation turns out to be fully uncoupled in the form

$$f^h(\mathbf{x}) = \lambda_h(\mathbf{x}) f^h(\mathbf{x}_o) \quad (3.53)$$

where $\lambda_h(\mathbf{x})$ is the associated eigenvalue to $\{\mathbf{e}_h(\mathbf{x})\}$. The uncoupling is a direct consequence of the invariance of $\{\mathbf{e}_h(\mathbf{x})\}$ under the action of vector dynamics¹⁰, i.e. invariance under the action of

¹⁰It should be pointed out that this latter statement, in spite of it being proven for the Davis-Skodje system, was not proven in general. In spite of its seemingly general validity, a rigorous proof of the invariance of subspaces $\{\mathbf{e}_h(\mathbf{x})\}$ under the action of the skew-product vector dynamics is an open issue that needs further work.

$$e^{\Omega(t, \Phi_{-t}(\mathbf{x}))}.$$

The foregoing issues can also be treated by following a different approach. The source term projections onto the invariant base are in the form $f^h(\mathbf{x}) = \mathbf{e}^h(\mathbf{x})\mathbf{g}(\mathbf{x})$ where $\{\{\mathbf{e}_h\}, \{\mathbf{e}^h\}\}$ $h = 1 \cdots n$ is a system of biorthonormal vectors/covectors. Supposing we are interested in the evolution of such projections along a trajectory $\mathbf{x}(t) = \Phi_t(\mathbf{x}_o)$, such evolution will be governed by the following equation:

$$\begin{aligned} \frac{df^h(\mathbf{x})}{dt} &= \frac{d\mathbf{e}^h(\mathbf{x})}{dt} \frac{d\mathbf{x}}{dt} + \mathbf{e}^h(\mathbf{x}) \frac{d\mathbf{g}(\mathbf{x})}{dt} = \\ &= \frac{d\mathbf{e}^h(\mathbf{x})}{dt} \mathbf{g}(\mathbf{x}) + \mathbf{e}^h(\mathbf{x}) J(\mathbf{x}) \mathbf{g}(\mathbf{x}) = \\ &= \sum_{j=1}^n \left(\frac{d\mathbf{e}^h(\mathbf{x})}{dt} + \mathbf{e}^h(\mathbf{x}) J(\mathbf{x}) \right) \mathbf{e}_j(\mathbf{x}) f^h(\mathbf{x}) \end{aligned} \quad (3.54)$$

where $J(\mathbf{x}) = d\mathbf{g}(\mathbf{x})/d\mathbf{x}$ and $\mathbf{g}(\mathbf{x}) = d\mathbf{x}/dt = \sum_{j=1}^n f^h(\mathbf{x}) \mathbf{e}_h(\mathbf{x})$. By looking at Eq.(3.54) we observe that the evolution of the projections $f^h(\mathbf{x})$ clearly depend upon the evolution of the system $\{\{\mathbf{e}_h\}, \{\mathbf{e}^h\}\}$. This is no surprise since $\mathbf{g}(\mathbf{x})$ is being projected onto the vector basis $\mathbf{e}_h(\mathbf{x})$, biorthonormal to $\mathbf{e}^h(\mathbf{x})$. If we wish to make explicit such dependance, then we should consider the evolution of normalized (unit norm) vectors and their corresponding dual covectors (whose norm is generally different from 1). Such evolution is obtained from the usual skew product variational equation by adding a renormalization coefficient, so that the evolution of vectors and covectors read respectively

$$\frac{d\mathbf{e}_h(\mathbf{x})}{dt} = J(\mathbf{x})\mathbf{e}_h(\mathbf{x}) - (J(\mathbf{x})\mathbf{e}_h(\mathbf{x}), \mathbf{e}_h(\mathbf{x})) \mathbf{e}_h(\mathbf{x}), \quad \mathbf{e}_h(\mathbf{x}) \in TC_{\mathbf{x}} \quad (3.55)$$

$$\frac{d\mathbf{e}^h(\mathbf{x})}{dt} = -\mathbf{e}^h(\mathbf{x})J(\mathbf{x}) + (J(\mathbf{x})\mathbf{e}_h(\mathbf{x}), \mathbf{e}_h(\mathbf{x})) \mathbf{e}^h(\mathbf{x}), \quad \mathbf{e}^h(\mathbf{x}) \in TC_{\mathbf{x}}^* \quad (3.56)$$

Note that by taking the scalar product of the first of the above equations with $\mathbf{e}_h(\mathbf{x})$ we obtain $d||\mathbf{e}_h||^2/dt = 2(J\mathbf{e}_h, \mathbf{e}_h)(1 - ||\mathbf{e}_h||^2)$ which is satisfied $\forall t$ if $||\mathbf{e}_h|| = 1, \forall t$. Note also that by considering the un-renormalized vector dynamics $d\tilde{\mathbf{e}}_h/dt = J\tilde{\mathbf{e}}_h$ and taking the scalar product with $\tilde{\mathbf{e}}_h$ we obtain $d||\tilde{\mathbf{e}}_h||^2/dt = 2(J\tilde{\mathbf{e}}_h, \tilde{\mathbf{e}}_h) = 2(J\mathbf{e}_h, \mathbf{e}_h)||\tilde{\mathbf{e}}_h||^2$ where $\mathbf{e}_h = \tilde{\mathbf{e}}_h/||\tilde{\mathbf{e}}_h||$, whose solution is

$$||\tilde{\mathbf{e}}_h(\mathbf{x}(t))|| = ||\tilde{\mathbf{e}}_h(\mathbf{x}_o)|| e^{\int_0^t (J(\mathbf{x}(\tau))\mathbf{e}_h(\mathbf{x}(\tau)), \mathbf{e}_h(\mathbf{x}(\tau))) d\tau}. \quad (3.57)$$

So by following the dynamics of invariant unit vectors $\mathbf{e}_h(\mathbf{x}(t))$, any initial vector $\tilde{\mathbf{e}}_h(\mathbf{x}_o)$ belonging to the invariant subspace $E_h(\mathbf{x}_o)$ (i.e. parallel to $\mathbf{e}_h(\mathbf{x}_o)$) will be mapped into a vector $\mathbf{e}_h(\mathbf{x})||\tilde{\mathbf{e}}_h(\mathbf{x}_o)|| e^{\int_0^t (J(\mathbf{x}(\tau))\mathbf{e}_h(\mathbf{x}(\tau)), \mathbf{e}_h(\mathbf{x}(\tau))) d\tau}$ belonging to $E_h(\mathbf{x}(t))$ (i.e. parallel to $\mathbf{e}_h(\mathbf{x}(t))$), this

highlighting the role of the renormalizing factor ($J\mathbf{e}_h, \mathbf{e}_h$) in the dynamics of invariant vectors.

Eq.(3.56) for covectors stems directly from Eq.(3.55) and from the differentiation of the biorthonormality condition $\mathbf{e}^i \mathbf{e}_j = \delta_j^i$. Note that if the set $\{\mathbf{e}_h(\mathbf{x}_o)\}$ is a basis of $TC_{\mathbf{x}_o}$, if $\|\mathbf{e}_h(\mathbf{x}_o)\| = 1 \forall h$ and if $\mathbf{e}^i(\mathbf{x}_o) \mathbf{e}_j(\mathbf{x}_o) = \delta_j^i$ then the system $(\{\mathbf{e}_h(\mathbf{x}(t))\}, \{\mathbf{e}^h(\mathbf{x}(t))\})$ is a system of biorthonormal vectors/covectors $\forall t$, i.e. (3.55)-(3.55) preserve biorthonormality. Note however that Eq.(3.56) does not preserve unit norm. We also note that by considering covector dynamics $d\tilde{\mathbf{e}}^h/dt = -\tilde{\mathbf{e}}^h J$, which preserves biorthogonality with the un-renormalized system of naturally evolving vectors $\{\tilde{\mathbf{e}}_h\}$, and again taking the scalar product with $\tilde{\mathbf{e}}^h$, we obtain $d\|\tilde{\mathbf{e}}^h\|^2/dt = -2(\tilde{\mathbf{e}}^h J, \tilde{\mathbf{e}}^h) = -2(\mathbf{e}^h J, \mathbf{e}^h)\|\tilde{\mathbf{e}}^h\|^2$ where $\mathbf{e}^h = \tilde{\mathbf{e}}^h/\|\tilde{\mathbf{e}}^h\|$, and whose solution is

$$\|\tilde{\mathbf{e}}^h(\mathbf{x}(t))\| = \|\tilde{\mathbf{e}}^h(\mathbf{x}_o)\| e^{\int_0^t -(\mathbf{e}^h(\mathbf{x}(\tau)) J(\mathbf{x}(\tau)), \mathbf{e}^h(\mathbf{x}(\tau))) d\tau}. \quad (3.58)$$

This would suggest that an initial covector $\tilde{\mathbf{e}}^h(\mathbf{x}_o)$ dual to $\tilde{\mathbf{e}}_h(\mathbf{x}_o) \in E_h(\mathbf{x}_o)$ (where $\tilde{\mathbf{e}}^i(\mathbf{x}_o) \tilde{\mathbf{e}}_j(\mathbf{x}_o) = \delta_j^i$), is mapped by $d\tilde{\mathbf{e}}^h/dt = -\tilde{\mathbf{e}}^h J$ into a covector $\tilde{\mathbf{e}}^h(\mathbf{x}) = \mathbf{e}^h(\mathbf{x}) \|\tilde{\mathbf{e}}^h(\mathbf{x}_o)\| e^{\int_0^t -(\mathbf{e}^h(\mathbf{x}(\tau)) J(\mathbf{x}(\tau)), \mathbf{e}^h(\mathbf{x}(\tau))) d\tau}$ dual to $\tilde{\mathbf{e}}_h(\mathbf{x}) \in E_h(\mathbf{x})$.

By substituting Eq.(3.56) into Eq.(3.54) we obtain

$$\frac{df^h(\mathbf{x})}{dt} = (J(\mathbf{x}) \mathbf{e}_h(\mathbf{x}), \mathbf{e}_h(\mathbf{x})) f^h(\mathbf{x}) \quad (3.59)$$

which, again, is fully decoupled and whose solution is

$$f^h(\mathbf{x}) = e^{\int_0^t (J(\mathbf{x}(\tau)) \mathbf{e}_h(\mathbf{x}(\tau)), \mathbf{e}_h(\mathbf{x}(\tau))) d\tau} f^h(\mathbf{x}_o) \quad (3.60)$$

The latter approach, because it is based on following unit invariant vectors in their natural evolution dictated by Eq.(3.55), may be referred to as a "Lagrangian" approach in contrast to the preceding pointwise Eulerian approach. It goes without saying that in the Lagrangian approach we have implied a prior, pointwise, knowledge of the invariant subspaces and have not specified an alternative method, other than the Eulerian, for determining them. We will shortly see that the Lagrangian approach, provided a suitable vector-unfolding procedure is enforced, is indeed a way of establishing invariant subspaces this being true due to a "dominance principle" of the most unstable subspaces. This concepts lie at the core of the Natural Tangent Dynamics with Recurrent Biorthonormalizations method to be discussed later.

The two different approaches yielded the two expressions Eq.(3.53) and Eq.(3.60) respectively. We should ask ourselves how these two seemingly equivalent equations are related. To this goal

let us refer to the Davis-Skodje system for which we can evaluate analytical expressions for the renormalizing factor $(J\mathbf{e}_h, \mathbf{e}_h)$, being the normalized pointwise invariant subspaces known from Eq.(3.45) and being

$$J(\mathbf{x}) = \begin{pmatrix} -\frac{1}{\varepsilon} & \frac{1+z-\varepsilon+z\varepsilon}{(1+z)^3\varepsilon} \\ 0 & -1 \end{pmatrix}. \quad (3.61)$$

After substituting $\mathbf{x} = \Phi_t(\mathbf{x}_o)$ we have the following expressions

$$(J(\mathbf{x})\mathbf{e}_1(\mathbf{x}), \mathbf{e}_1(\mathbf{x})) = -\frac{1}{\varepsilon} \quad (3.62)$$

$$(J(\mathbf{x})\mathbf{e}_2(\mathbf{x}), \mathbf{e}_2(\mathbf{x})) = \frac{1}{1 + \frac{1}{(1+e^{-t}z_0)^4}} \left(-1 + \frac{e^{-t}z_0 - 1}{(e^{-t}z_0 + 1)^5} \right) \quad (3.63)$$

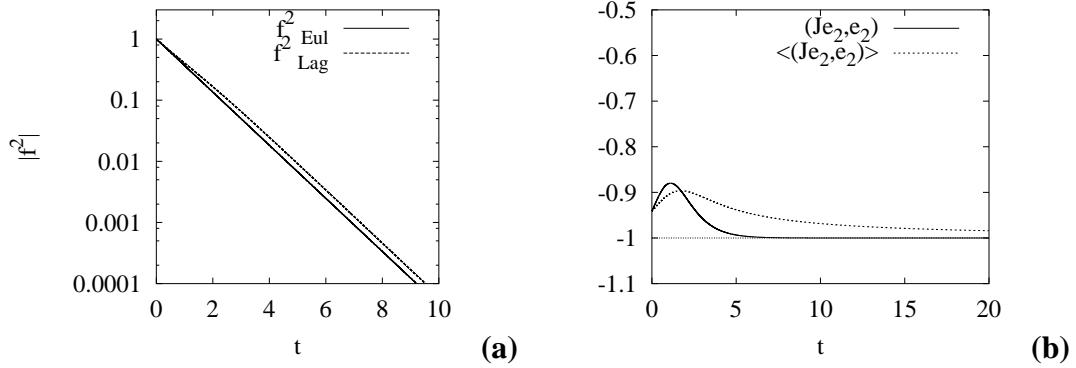


Fig. 3.5: (a) Solutions to Eq(3.53) and (3.60) with $h = 2$ and $f^2(\mathbf{x}_o) = -z_0 = -1$. (b) Renormalization coefficient $(J\mathbf{e}_2, \mathbf{e}_2)$ and its time average $\langle(J\mathbf{e}_2, \mathbf{e}_2)\rangle$ along a trajectory with $(y_0, z_0) = (1, 1)$

We notice that $e^{\int_0^t (J\mathbf{e}_1, \mathbf{e}_1)d\tau} = e^{-t/\varepsilon} = \lambda_1$. This should not mislead us into thinking that $e^{\int_0^t (J\mathbf{e}_h, \mathbf{e}_h)d\tau} = \lambda_h$ is a general result, in fact it isn't. In this particular case however the system can be regarded as semilinear because $\mathbf{e}_1 = (1, 0)^T$ is a flat invariant subspace, just as in linear systems. Indeed in the other case $e^{\int_0^t (J\mathbf{e}_2, \mathbf{e}_2)d\tau} \neq \lambda_2 = e^{-t}$. By comparing Eq.(3.53), i.e $f_{Eul}^2(\mathbf{x}) = \lambda_2 f^2(\mathbf{x}_o) = e^{-t} f^2(\mathbf{x}_o)$, with Eq.(3.60), i.e $f_{Lag}^2(\mathbf{x}) = e^{\int_0^t (J\mathbf{e}_2, \mathbf{e}_2)d\tau} f^2(\mathbf{x}_o)$, however, we notice that after an initial transient, $f_{Lag}^2 \sim e^{-t}$ as shown in Fig.(3.5)(a), i.e f_{Lag}^2 behaves like f_{Eul}^2 asymptotically.

Fig.(3.5)(b) reveals that $\lim_{t \rightarrow \infty} (J\mathbf{e}_2, \mathbf{e}_2) = -1$ and in fact the time average of the renormalizing factor tends to -1 , i.e. $\lim_{t \rightarrow \infty} \langle(J\mathbf{e}_2, \mathbf{e}_2)\rangle = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t (J\mathbf{e}_2, \mathbf{e}_2)d\tau = \lim_{t \rightarrow \infty} \frac{1}{t} \log e^{\int_0^t (J\mathbf{e}_2, \mathbf{e}_2)d\tau} = -1$. On the other hand applying the same reasoning to λ_2 we have $\lim_{t \rightarrow \infty} \frac{1}{t} \log \lambda_2 = \lim_{t \rightarrow \infty} \frac{1}{t} \log e^{-t} = -1$. These rather empirical results would tend to lead us to conclude that

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \lambda_h(\mathbf{x}(t)) = \lim_{t \rightarrow \infty} \frac{1}{t} \log e^{\int_0^t (J(\mathbf{x})\mathbf{e}_h(\mathbf{x}), \mathbf{e}_h(\mathbf{x}))d\tau} = \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\tilde{\mathbf{e}}_h(\mathbf{x})\| \quad (3.64)$$

where $\tilde{\mathbf{e}}_h(\mathbf{x})$ are the unnormalized invariant subspaces, eigenvectors of $V(\mathbf{x}, t)$ for $t \rightarrow \infty$, as they naturally evolve through the system's vector dynamics, $\lambda_h(\mathbf{x}(t))$ are the corresponding eigenvalues and $\mathbf{e}_h(\mathbf{x})$ are the normalized invariant subspaces. It would seem, therefore, that each invariant subspace is characterized not merely by its corresponding pointwise eigenvalue $\lambda_h(\mathbf{x}(t))$ but by an even more general quantity, of statistical nature, represented by Eq.(3.64), embedding the asymptotic average stretching/contracting features of invariant vectors as they evolve naturally within the corresponding invariant subspace through the vector dynamics dictated by the original autonomous system. Luckily these concepts were generalized and proven in the celebrated Oseledec's theorem (1968) which we will see in the next section.

3.6 Oseledec's Multiplicative Ergodic Theorem

The invariant geometric structure of dynamical systems possessing nontrivial invariant sets and its relation with the statistical long-term dynamical characterization is described by the Oseledec's theorem [78, 66, 70]. A formulation of the Oseledec's theorem suited for the aim of this dissertation is the following.

Theorem 3.6.1 *Let $\phi_t(\mathbf{x})$ be the phase flow associated with the dynamical system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$, which is ergodic in C and let $\mu^*(\mathbf{x})$ be its ergodic measure. If $\int_C \log^+ \|\mathbf{J}(\mathbf{x})\| d\mu^*(\mathbf{x}) < \infty$, (where $\log^+ y = \max(0, \log y)$, and $\|\mathbf{J}\|$ is the norm of the operator \mathbf{J}), then the tangent and cotangent spaces $TC_{\mathbf{x}}$ and $TC_{\mathbf{x}}^*$ at any point $\mathbf{x} \in C$ can be decomposed into the direct sum of subspaces*

$$TC_{\mathbf{x}} = \bigoplus_{h=1}^n E^h(\mathbf{x}) \quad TC_{\mathbf{x}}^* = \bigoplus_{h=1}^n F^h(\mathbf{x}), \quad (3.65)$$

such that:

1. For each vector $\mathbf{v}(0) \in E^h(\mathbf{x}_o)$ and covector $\mathbf{w}(0) \in F^h(\mathbf{x}_o)$ at the point $\mathbf{x}(0) = \mathbf{x}_o$, the evolution of the skew-product system

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) \quad \dot{\mathbf{v}}(t) = J(\mathbf{x}(t)) \mathbf{v}(t) \quad \dot{\mathbf{w}}(t) = \mathbf{w}(t) J(\mathbf{x}(t)) \quad (3.66)$$

is such that:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mathbf{v}(t)\| = \Lambda_h \quad \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mathbf{w}(t)\| = \Lambda_h^* \quad (3.67)$$

where the real numbers Λ_h , Λ_h^* do not depend on \mathbf{x}_o , $\mathbf{v}(0)$ and $\mathbf{w}(0)$, but exclusively on $h = 1, \dots, N$.

2. Each subbundle $\{E^h(\mathbf{x})\}_{\mathbf{x} \in C}$, $\{F^h(\mathbf{x})\}_{\mathbf{x} \in C}$, $h = 1, \dots, N$ is invariant under the action of the dynamical system. This means that for any $\mathbf{x}_o \in C$, $\mathbf{v}(0) \in E^h(\mathbf{x}_o)$, $\mathbf{w}(0) \in F^h(\mathbf{x}_o)$, the vector $\mathbf{v}(t)$ and the covector $\mathbf{w}(t)$, solution at time t of the equations Eq. (3.66), lie respectively in the subspaces $E^h(\phi_t(\mathbf{x}_o))$, $F^h(\phi_t(\mathbf{x}_o))$ belonging to the tangent and cotangent spaces $TC_{\phi_t(\mathbf{x}_o)}$, $TC_{\phi_t(\mathbf{x}_o)}^*$ of the image point $\phi_t(\mathbf{x}_o)$. ■

The remarkable aspects of Oseledec's theorem lie in the fact that the system's tangent space invariant decomposition is defined with great generality, independently of any particular feature the system may possess except it being ergodic in C . This means that an invariant decomposition may be sought even when the system should not exhibit the features needed by CSP theory, namely that of being singularly perturbed and having a trivial invariant set (i.e. an equilibrium point).

We have learned from Eq.(3.57) that given any initial vector $\mathbf{v}(\mathbf{x}_o) \in E^h(\mathbf{x}_o)$, its image through Eq.(3.66) $\mathbf{v}(t) \in E^h(\phi_t(\mathbf{x}_o)) = E^h(\mathbf{x}(t))$ will have a norm given by

$$\|\mathbf{v}(\mathbf{x}(t))\| = \|\mathbf{v}(\mathbf{x}_o)\| e^{\int_0^t (J(\mathbf{x}(\tau))\mathbf{e}_h(\mathbf{x}(\tau)), \mathbf{e}_h(\mathbf{x}(\tau))) d\tau}, \quad (3.68)$$

where $\mathbf{e}_h(\mathbf{x}(t))$ is the unit vector spanning $E^h(\mathbf{x}(t))$. We therefore conclude that Liapunov exponents may be determined by

$$\Lambda_h = \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mathbf{v}(\mathbf{x}(t))\| = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t (J(\mathbf{x}(t))\mathbf{e}_h(\mathbf{x}(t)), \mathbf{e}_h(\mathbf{x}(t))). \quad (3.69)$$

If we had a prior knowledge of the Liapunov-Oseledec invariant subspaces $E^h(\mathbf{x})$ we could follow the dynamics of unit vectors $\mathbf{e}_h(\mathbf{x}) \in E^h(\mathbf{x})$ within them and use such vectors to project the system's source term $\mathbf{g}(\mathbf{x})$. As a result, and as Eq.(3.60) shows, i.e.

$$f^h(\mathbf{x}) = e^{\int_0^t (J\mathbf{e}_h, \mathbf{e}_h) d\tau} f^h(\mathbf{x}_o) = e^{t \langle (J\mathbf{e}_h, \mathbf{e}_h) \rangle} f^h(\mathbf{x}_o), \quad (3.70)$$

the projections of $\mathbf{g}(\mathbf{x})$ would evolve in a decoupled way, this allowing us to define a spectrum of unambiguous intrinsic timescales, which for sufficiently large t would correspond to the Liapunov spectrum because $f^h \sim e^{\Lambda_h t}$.

The structure of the invariant subspaces can be conveniently reformulated by introducing the concept of invariant filtrations. Given a finite-dimensional vector space W , a collection $\mathcal{V} = \{V_k, k = 0, \dots, m \leq N\}$ of linear subspaces of W , satisfying the condition $\{\mathbf{0}\} = V_0 \subset V_1 \subset \dots \subset V_m = W$ is called a *filtration* of W .

For the invariant subspaces $\{E^h(\mathbf{x})\}_{h=1}^N$, $\{F^h(\mathbf{x})\}_{h=1}^N$ at a point $\mathbf{x} \in C$, the invariant filtrations $\mathcal{V}_x = \{\overline{E}_x^h\}_{h=0}^N$, $\mathcal{V}_x^* = \{\overline{F}_x^h\}_{h=0}^N$ in the tangent and cotangent spaces can be defined as:

$$\begin{aligned}\overline{E}_x^1 &= E_x^1 & \overline{F}_x^1 &= F_x^1 \\ \overline{E}_x^2 &= E_x^1 \oplus E_x^2 & \overline{F}_x^2 &= F_x^1 \oplus F_x^2 \\ &\vdots & &\vdots \\ \overline{E}_x^h &= \bigoplus_{k=1}^h E_x^k & \overline{F}_x^h &= \bigoplus_{k=1}^h F_x^k & h = 1, \dots, N.\end{aligned}\tag{3.71}$$

The filtrations \mathcal{V}_x and \mathcal{V}_x^* can be referred to as *the Lyapunov-Oseledec filtrations* at a point $\mathbf{x} \in C$. The purpose of the geometric characterization of a dynamical system is to obtain the Lyapunov-Oseledec filtrations at each point of the invariant set and the spectrum of the Lyapunov exponents. This will allow us to unambiguously identify the most stable components of the dynamics thereby making it possible to carry out a model reduction by neglecting such components. In the next sections we will introduce the NTDRB technique which accomplishes the task of determining the Lyapunov-Oseledec filtrations by following the natural vector/covector dynamics provided an "unfolding" biorthonormalization procedure is enforced to contrast the dominance principle which will be the topic of the following section.

3.7 A dominance principle

In the section where we introduced the Lagrangian approach we followed the evolution of unit invariant vectors and their dual covectors by means of Eqs.(3.55),(3.56). We have seen that any initial vector $\tilde{\mathbf{e}}_h(\mathbf{x}_o)$ belonging to the invariant subspace $E_h(\mathbf{x}_o)$ will be mapped into a vector $\tilde{\mathbf{e}}_h(\mathbf{x}) = \mathbf{e}_h(\mathbf{x})||\tilde{\mathbf{e}}_h(\mathbf{x}_o)||e^{\int_0^t (J(\mathbf{x}(\tau))\mathbf{e}_h(\mathbf{x}(\tau)), \mathbf{e}_h(\mathbf{x}(\tau)))d\tau}$ belonging to $E_h(\mathbf{x}(t))$. Suppose we consider a generic vector $\mathbf{a}(\mathbf{x}_o) \in TC_{\mathbf{x}_o}$ that has non-zero projections on all invariant subspaces $E_h(\mathbf{x}_o)$, and let us consider its evolution through the skew product system

$$\frac{d\mathbf{x}}{dt} = \mathbf{g}(\mathbf{x}) \quad ; \quad \frac{d\mathbf{a}(\mathbf{x})}{dt} = J(\mathbf{x})\mathbf{a}(\mathbf{x}).\tag{3.72}$$

Vector $\mathbf{a}(\mathbf{x}(t)) \in TC_{\mathbf{x}(t)}$ may be at all times projected onto the invariant unit basis at $TC_{\mathbf{x}(t)}$, i.e. $\mathbf{a}(\mathbf{x}(t)) = \gamma^1(\mathbf{x}(t))\mathbf{e}_1(\mathbf{x}(t)) + \dots + \gamma^n(\mathbf{x}(t))\mathbf{e}_n(\mathbf{x}(t))$ where each component is mapped invariantly as $\gamma^h(\mathbf{x}(t))\mathbf{e}_h(\mathbf{x}(t)) = \gamma^h(\mathbf{x}_o)\mathbf{e}_h(\mathbf{x})e^{\int_0^t (J(\mathbf{x}(\tau))\mathbf{e}_h(\mathbf{x}(\tau)), \mathbf{e}_h(\mathbf{x}(\tau)))d\tau}$, (note that $||\mathbf{e}_h(\mathbf{x}_o)|| = 1$),

where $\gamma^h(\mathbf{x}_o) = \mathbf{e}^h(\mathbf{x}_o)\mathbf{a}(\mathbf{x}_0)$. Therefore we may write for vector $\mathbf{a}(\mathbf{x}(t))$:

$$\mathbf{a}(\mathbf{x}(t)) = \sum_{h=1}^n e^{\int_0^t (J(\mathbf{x}(\tau))\mathbf{e}_h(\mathbf{x}(\tau)), \mathbf{e}_h(\mathbf{x}(\tau))) d\tau} \left(\mathbf{e}^h(\mathbf{x}_o)\mathbf{a}(\mathbf{x}_0) \right) \mathbf{e}_h(\mathbf{x}(t)) \quad (3.73)$$

Suppose that the invariant subspaces $E_h(\mathbf{x})$ are ordered so that for the relative Liapunov exponents we can write $\Lambda_1 > \Lambda_2 > \dots > \Lambda_n$. Eq.(3.73) may then be rewritten as

$$\begin{aligned} \mathbf{a}(\mathbf{x}) &= e^{\int_0^t (J\mathbf{e}_1, \mathbf{e}_1) d\tau} \left[(\mathbf{e}^1(\mathbf{x}_o)\mathbf{a}(\mathbf{x}_0)) \mathbf{e}_1(\mathbf{x}) + \right. \\ &\quad \left. + \sum_{h=2}^n e^{\int_0^t (J\mathbf{e}_h, \mathbf{e}_h) d\tau - \int_0^t (J\mathbf{e}_1, \mathbf{e}_1) d\tau} \left(\mathbf{e}^h(\mathbf{x}_o)\mathbf{a}(\mathbf{x}_0) \right) \mathbf{e}_h(\mathbf{x}) \right] \\ &= e^{t\langle (J\mathbf{e}_1, \mathbf{e}_1) \rangle(t)} \left[(\mathbf{e}^1(\mathbf{x}_o)\mathbf{a}(\mathbf{x}_0)) \mathbf{e}_1(\mathbf{x}) + \right. \\ &\quad \left. + \sum_{h=2}^n e^{t\langle (J\mathbf{e}_h, \mathbf{e}_h) \rangle(t) - t\langle (J\mathbf{e}_1, \mathbf{e}_1) \rangle(t)} \left(\mathbf{e}^h(\mathbf{x}_o)\mathbf{a}(\mathbf{x}_0) \right) \mathbf{e}_h(\mathbf{x}) \right] \end{aligned} \quad (3.74)$$

For $t \rightarrow \infty$, $\langle (J\mathbf{e}_h, \mathbf{e}_h) \rangle(t) \rightarrow \Lambda_h$, therefore for sufficiently large t , because $\Lambda_h - \Lambda_1 < 0 \forall h > 2$ we conclude that $\mathbf{a}(\mathbf{x}(t)) \sim e^{t\Lambda_1} \mathbf{e}_1(\mathbf{x})$. This is what we can define as a dominance principle: a generic initial vector $\mathbf{a}(\mathbf{x}_o)$ evolving according to Eq.(3.73) will progressively align itself towards the most unstable invariant subspace, i.e. the subspace exhibiting the largest Liapunov exponent (moreover we notice that $\lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mathbf{a}(\mathbf{x}(t))\| = \Lambda_1$). This circumstance is intuitive if we think that a generic vector will initially almost everywhere¹¹ possess a non-zero component in the most unstable invariant subspace, a component that either stretches more or contracts less than all the others and will therefore eventually dominate. On the other hand, note that if $\mathbf{a}(\mathbf{x}_o)$ had no component on $E_1(\mathbf{x}_o)$, i.e. if $\mathbf{e}^1(\mathbf{x}_o)\mathbf{a}(\mathbf{x}_0) = 0$, then the dominance principle would act on the second most unstable subspace and so on. The latter concept will prove crucial in the application of the NTDRB technique which is to be illustrated further.

Note that the dominance principle is straight forward for linear autonomous systems. In the case of real eigenvalues Eq.(3.18) may be rewritten as

$$\mathbf{a}(t) = e^{\lambda_1 t} \left(\mathbf{v}_1(\mathbf{w}^1 \mathbf{a}(0)) + \sum_{i=2}^n \mathbf{v}_i e^{(\lambda_i - \lambda_1)t} (\mathbf{w}^i \mathbf{a}(0)) \right) \quad (3.75)$$

and if $\lambda_1 > \dots > \lambda_n$ and if $\mathbf{w}^i \mathbf{a}(0) \neq 0$ then for sufficiently large t , $\mathbf{a}(t) \sim e^{\lambda_1 t} \mathbf{v}_1(\mathbf{w}^1 \mathbf{a}(0))$.

¹¹This means that $\mathbf{e}^1(\mathbf{x}_o)\mathbf{a}(\mathbf{x}_0) = 0$ for $\mathbf{x}_o \in X \subset C$ and $X = \emptyset$ i.e. X is a null set (of measure zero).

The dominance principle, as we will see, has profound consequences on the development of the present theory. Suppose we wish to apply the classical CSP approach to our dynamical system projecting the vector field $\mathbf{g}(\mathbf{x})$ onto a set of independent unit basis vectors $\{\mathbf{a}_i(\mathbf{x}(t))\}$, $i = 1, \dots, n$, so that each projection is $f^i(\mathbf{x}) = \mathbf{b}^i(\mathbf{x})\mathbf{g}(\mathbf{x})$ (note that the dependance from t is sometimes neglected to lighten the notation). Suppose further that the unit basis vectors $\{\mathbf{a}_i(\mathbf{x}(t))\}$ and its corresponding dual covector basis $\{\mathbf{b}^i(\mathbf{x})\}$ are chosen by means of a Lagrangian approach, i.e. as the natural unit vector and covector evolution, where $\mathbf{x}(t) = \Phi_t(\mathbf{x}_o)$, according to

$$\frac{d\mathbf{a}_i(\mathbf{x})}{dt} = J(\mathbf{x})\mathbf{a}_i(\mathbf{x}) - (J(\mathbf{x})\mathbf{a}_i(\mathbf{x}), \mathbf{a}_i(\mathbf{x}))\mathbf{a}_i(\mathbf{x}), \quad \mathbf{a}_i(\mathbf{x}) \in TC_{\mathbf{x}} \quad (3.76)$$

$$\frac{d\mathbf{b}^i(\mathbf{x})}{dt} = -\mathbf{b}^i(\mathbf{x})J(\mathbf{x}) + (J(\mathbf{x})\mathbf{a}_i(\mathbf{x}), \mathbf{a}_i(\mathbf{x}))\mathbf{b}^i(\mathbf{x}), \quad \mathbf{b}^i(\mathbf{x}) \in TC_{\mathbf{x}}^* \quad (3.77)$$

starting from a generic basis $\{\mathbf{a}_i(\mathbf{x}_o)\} \in TC_{\mathbf{x}_o}$ whose only prerequisite is $\|\mathbf{a}_i(\mathbf{x}_o)\| = 1$ for $i = 1, \dots, n$ and $\mathbf{b}^i(\mathbf{x}_o)\mathbf{a}_j(\mathbf{x}_o) = \delta_j^i$ for $i = 1, \dots, n$. We know that this can be done because Eqs.(3.76),(3.77) preserve biorthonormality and therefore $\{\mathbf{b}^i(\mathbf{x})\}$ always retain their role of projection operators (note that $\{\mathbf{b}^i(\mathbf{x})\}$ do not preserve unit norm). For this particular choice of basis vectors, which differ from the system's invariant vectors, we know that the amplitudes $f^i(\mathbf{x})$ evolve according to (see Eq.(3.59))

$$\frac{df^i(\mathbf{x})}{dt} = (J(\mathbf{x})\mathbf{a}_i(\mathbf{x}), \mathbf{a}_i(\mathbf{x}))f^i(\mathbf{x}) \quad (3.78)$$

whose solution is

$$f^i(\mathbf{x}) = e^{\int_0^t (J\mathbf{a}_i, \mathbf{a}_i) d\tau} f^i(\mathbf{x}_o) = e^{t\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(t)} f^i(\mathbf{x}_o). \quad (3.79)$$

What we have learned from the dominance principle is that any initial vector will almost everywhere progressively converge to the most unstable invariant direction $\mathbf{e}_1(\mathbf{x})$ corresponding to the greatest Liapunov exponent Λ_1 . This means that the whole basis of unit vectors $\{\mathbf{a}_i(\mathbf{x}(t))\}$ will collapse onto $\mathbf{e}_1(\mathbf{x}(t)) \forall i$ and for sufficiently large t . Moreover, as $t \rightarrow \infty$, $\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(t) \rightarrow \Lambda_1$, and therefore $f^i(\mathbf{x}(t)) \sim e^{t\Lambda_1} f^i(\mathbf{x}_o)$. This shows that the choice of $\{\mathbf{a}_i(\mathbf{x}(t))\}$ as a projection basis leads to the impossibility of determining the intrinsic timescale spectrum for the system (represented by the spectrum of Liapunov exponents), but merely results into the determination of the most unstable timescale Λ_1 and the corresponding invariant subspace.

To conclude we may ask what effect the dominance principle has on the basis of covectors $\mathbf{b}^i(\mathbf{x}(t))$, evolving according to $d\mathbf{b}^i/dt = -\mathbf{b}^i J$, biorthogonal to the vector basis $\mathbf{a}_i(\mathbf{x}(t))$, evolving according to $d\mathbf{a}_i/dt = J\mathbf{a}_i$. As we projected the set $\mathbf{a}_i(\mathbf{x}(t)) = \gamma_i^1(\mathbf{x}(t))\mathbf{e}_1(\mathbf{x}(t)) + \dots + \gamma_i^n(\mathbf{x}(t))\mathbf{e}_n(\mathbf{x}(t))$

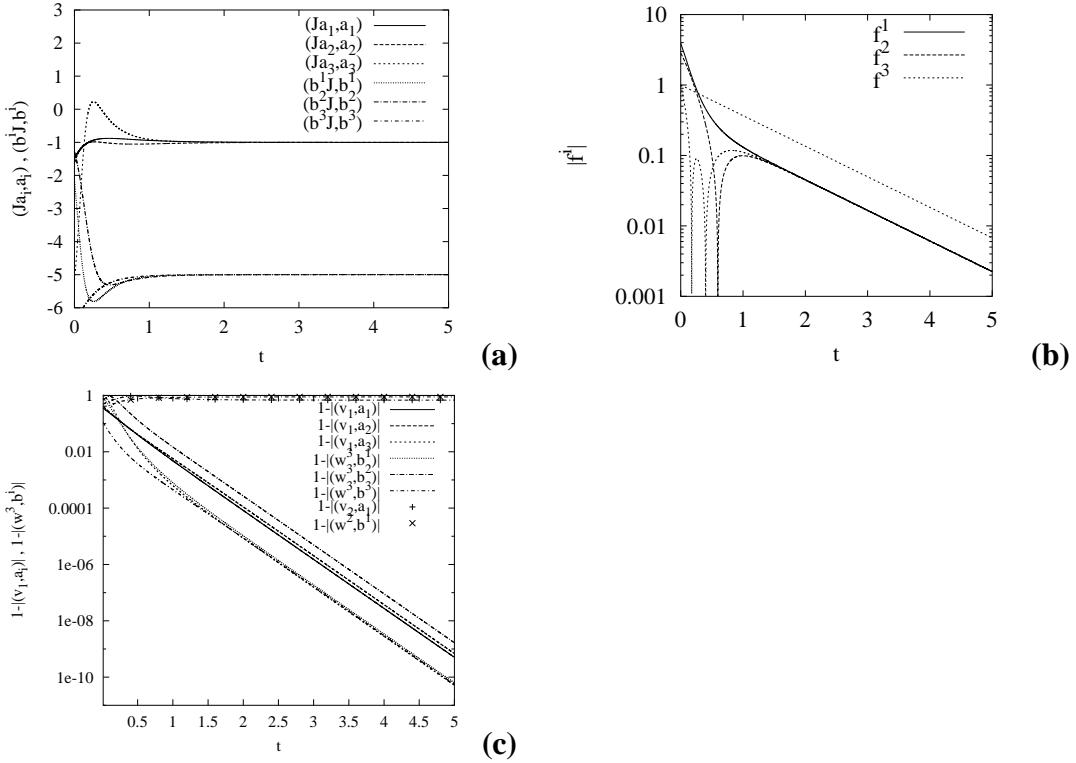


Fig. 3.6: Linear system Eq.(3.81). (a)Behavior of (Ja_i, a_i) and $(b^i J, b^i)$ along a trajectory stemming from $\mathbf{x}_o = (1, 1, 1)^T$ (b)Corresponding behavior of $f^i(\mathbf{x}(t))$, the dotted straight line represents e^{-t} . (c)Folding of vectors $\{\mathbf{a}_i\}$ onto \mathbf{v}_1 and covectors $\{\mathbf{b}^i\}$ onto \mathbf{w}^3 .

with $\gamma_i^h = \mathbf{e}_h(\mathbf{x}(t))\mathbf{a}_i(\mathbf{x}(t))$, we can project $\mathbf{b}^i(\mathbf{x}(t)) = \alpha_1^i(\mathbf{x}(t))\mathbf{e}^1(\mathbf{x}(t)) + \dots + \alpha_n^i(\mathbf{x}(t))\mathbf{e}^n(\mathbf{x}(t))$ where $\alpha_h^i(\mathbf{x}(t)) = \mathbf{b}^i(\mathbf{x}(t))\mathbf{e}_h(\mathbf{x}(t))$. The latter statement follows from $\mathbf{b}^i\mathbf{a}_j = \sum_h \alpha_h^i \gamma_j^h = \sum_h \mathbf{b}^i \mathbf{e}_h \mathbf{e}^h \mathbf{a}_j = \mathbf{b}^i I \mathbf{a}_j = \delta_j^i$. If we recall Eq.(3.58), whereby we hypothesized that each component is mapped as $\alpha_h^i(\mathbf{x})\mathbf{e}^h(\mathbf{x}) = \alpha_h^i(\mathbf{x}_o)\mathbf{e}^h(\mathbf{x})e^{\int_0^t -\langle (\mathbf{e}^h(\mathbf{x})J(\mathbf{x}), \mathbf{e}^h(\mathbf{x})) \rangle d\tau}$, we could write

$$\mathbf{b}^i(\mathbf{x}) = \sum_{h=1}^n (\mathbf{b}^i(\mathbf{x}_o)\mathbf{e}_h(\mathbf{x}_o)) \mathbf{e}^h(\mathbf{x}) e^{-\langle (\mathbf{e}^h(\mathbf{x})J(\mathbf{x}), \mathbf{e}^h(\mathbf{x})) \rangle}. \quad (3.80)$$

if we define the numbers $\lim_{t \rightarrow \infty} \langle (\mathbf{e}^h J, \mathbf{e}^h) \rangle = \Lambda^h$ (which we may refer to as dual Liapunov exponents), and supposing $\Lambda^1 > \dots > \Lambda^n$, we conclude, applying the same reasoning as before, that for a sufficiently large t , because of the minus sign in the above equation, $\mathbf{b}^i(\mathbf{x}(t)) \sim e^{-t\Lambda^n} \mathbf{e}^n(\mathbf{x}(t))$. It would appear therefore that the set of covectors $\{\mathbf{b}^i(\mathbf{x}(t))\}$ naturally collapse onto the subspace spanned by $\mathbf{e}^n(\mathbf{x})$ dual to $\mathbf{e}_n(\mathbf{x}) \in E_n(\mathbf{x})$ where $E_n(\mathbf{x})$ is the invariant subspace characterized by the smallest Liapunov exponent.

To illustrate the dominance principle and its implications let us consider two examples. First, let us

consider a linear example in R^3

$$\dot{\mathbf{x}} = A\mathbf{x}, \quad A = \begin{pmatrix} -2 & -0.5 & -1 \\ -4 & -3 & 4 \\ 1 & 0.5 & -4 \end{pmatrix} \quad (3.81)$$

and the associated normalized vector dynamics Eq.(3.76) and relative biorthonormal covector dynamics Eq.(3.77) (where clearly, $J = A$). The system has three invariant subspaces and corresponding intrinsic timescales which are identified with the eigensystem of matrix A . The three eigenvalues are $\lambda_1 = -1$ (corresponding to the most unstable invariant subspace), $\lambda_2 = -3$ and $\lambda_3 = -5$ which control the system's exponential relaxation towards the stable equilibrium $\mathbf{x}_S = \mathbf{0}$. Fig.(3.6)(a) shows how as $t \rightarrow \infty$ the quantity $(J\mathbf{a}_i, \mathbf{a}_i) \rightarrow \lambda_1 = -1$, when the initial unit basis $\{\mathbf{a}_i(\mathbf{x}_o)\}$ is chosen as a random basis. Fig.(3.6)(b) shows correspondingly that for sufficiently large t each projection along a trajectory $f^i(\mathbf{x}(t)) \sim e^{\lambda_1 t} = e^{-t}$.

Interestingly, Fig.(3.6)(a) also shows that the quantity $(\mathbf{b}^i J, \mathbf{b}^i) \rightarrow \lambda_3 = -5$, when $\{\mathbf{b}^i(\mathbf{x}_o)\}$ is biorthonormal to $\{\mathbf{a}_i(\mathbf{x}_o)\}$ and $\{\mathbf{b}^i(\mathbf{x})\}$ is renormalized at each step. This confirms that the dominance principle acts on covectors \mathbf{b}^i so as to force them to fold onto \mathbf{e}^n dual to $\mathbf{e}_n \in E_n$ where E_n is the most stable eigenpace (with the smallest eigenvalue). This can also be seen in Fig.(3.6)(c) where it is shown that $1 - |(\mathbf{v}_1, \mathbf{a}_i)| \rightarrow 0$ and $1 - |(\frac{\mathbf{w}^3}{\|\mathbf{w}^3\|}, \mathbf{b}^i)| \rightarrow 0 \forall i$ and where \mathbf{v}_i and \mathbf{w}^i are the set of right and left eigenvectors of A , namely for $\lambda_1 = -1$, $\mathbf{v}_1 = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0\right)^T$ and $\mathbf{w}^1 = \left(-\frac{\sqrt{5}}{2}, \frac{\sqrt{5}}{4}, \frac{\sqrt{5}}{2}\right)$, for $\lambda_2 = -3$, $\mathbf{v}_2 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)^T$ and $\mathbf{w}^2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{2\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, and for $\lambda_3 = -5$, $\mathbf{v}_3 = \left(0, -\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right)^T$ and $\mathbf{w}^3 = \left(-\frac{\sqrt{5}}{2}, -\frac{\sqrt{5}}{4}, \frac{\sqrt{5}}{2}\right)$, (where \mathbf{w}^i just shown do not possess unit norm).

The second example refers to the Lorenz system Eq.(3.28) in its chaotic attractor. Again, the evolution of an initial random vector basis was followed and the quantity $\langle(J\mathbf{a}_i, \mathbf{a}_i)\rangle$ was monitored. Fig.(3.7)(a) shows that as $t \rightarrow \infty$, $\langle(J\mathbf{a}_i, \mathbf{a}_i)\rangle \rightarrow \Lambda_1 = 0.91 \forall i$, which corresponds to the greatest Liapunov exponent for the Lorenz system, as reported in the literature, e.g.[79], where it is accurately calculated using exterior algebra methods. Fig.(3.7)(b) shows on the other hand that $\langle(\mathbf{b}^i J, \mathbf{b}^i)\rangle \rightarrow \Lambda^3 = -14.45 \forall i$, which corresponds to the smallest dual Liapunov exponent for the Lorenz system.

An interesting issue arises when we consider vector dynamics applied not to a generic vector but to the system's autonomous vector field $\mathbf{g}(\mathbf{x})$. This is trivially governed by the following evolution equation $\dot{\mathbf{g}}(\mathbf{x}(t)) = (\partial \mathbf{g}(\mathbf{x}(t))/\partial \mathbf{x})d\mathbf{x}/dt = J(\mathbf{x}(t))\mathbf{g}(\mathbf{x}(t))$ which, as we know has a solution in

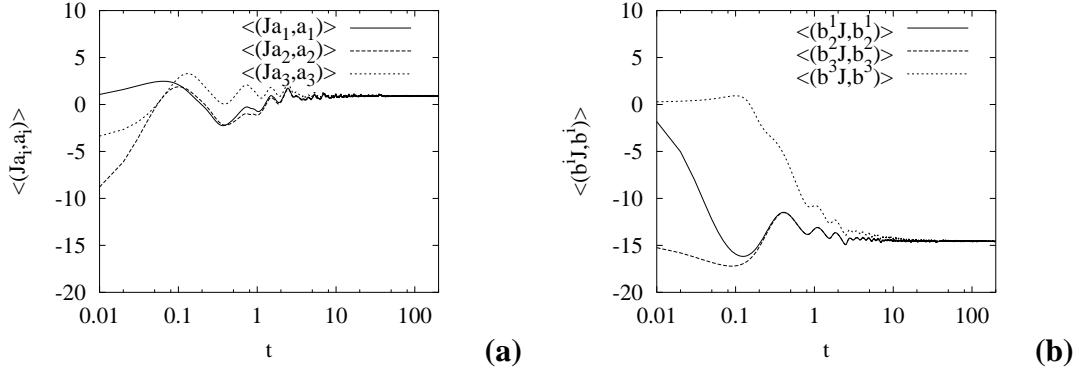


Fig. 3.7: Lorenz system. $\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(t)$ and $\langle (\mathbf{b}^i J, \mathbf{b}^i) \rangle(t)$ along the system's trajectory starting from $\mathbf{x}_o = (10, 10, 10)^T$.

the form $\mathbf{g}(\mathbf{x}(t)) = e^{\Omega(t, \mathbf{x}_o)} \mathbf{g}(\mathbf{x}_o)$. For any point \mathbf{x}_o , the vector $\mathbf{g}(\mathbf{x}_o) \in TC_{\mathbf{x}_o}$ which is tangent to the trajectory at \mathbf{x}_o , will be mapped into the vector $\mathbf{g}(\mathbf{x}(t)) \in TC_{\mathbf{x}(t)}$ tangent to the trajectory at $\mathbf{x}(t) = \Phi_t(\mathbf{x}_o)$. Let us now define the notion of invariant set [67] \mathcal{C} for our dynamical system, i.e. for the flow Φ_t in C , as a subset $\mathcal{C} \subseteq C$ such that for $\mathbf{x} \in \mathcal{C}$, $\Phi_t(\mathbf{x}) \in \mathcal{C} \forall t$ (or simply $\Phi_t(\mathcal{C}) = \mathcal{C}$). Examples of invariant sets are the stable or unstable manifolds of a fixed point or of a limit cycle. The asymptotic behavior of a dynamical system, however, takes place in the so called attracting set¹² $\mathcal{A} \subseteq C$ which is (this is among the possible definitions) an invariant set such that there exists a neighborhood U of \mathcal{A} for which $\Phi_{t \geq 0}(\mathbf{x}) \in U$ and $\Phi_t(\mathbf{x}) \rightarrow \mathcal{A}$ as $t \rightarrow \infty \forall \mathbf{x} \in U$. Examples of such attracting sets are limit cycles or chaotic attractors and within them $\|\mathbf{g}\|_{\mathcal{A}}$ is bounded¹³. By defining $G_M = \max_{\mathcal{A}} \|\mathbf{g}\| > 0$ then $0 < \|\mathbf{g}(\mathbf{x}(t))\| \leq G_M, \forall t$. With this in mind we may ask what invariant subspace $\mathbf{g}(\mathbf{x}(t))$ will span as $t \rightarrow \infty$. Clearly the Liapunov exponent associated to such invariant subspace will be $0 \leq \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mathbf{g}(\mathbf{x}(t))\| \leq \lim_{t \rightarrow \infty} \frac{1}{t} \log G_M = 0$ and therefore $\Lambda_C = \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\mathbf{g}(\mathbf{x}(t))\| = 0$, as shown in Fig.(3.7). This means that $\mathbf{g}(\mathbf{x}(t))$ will span the invariant subspace $E_C(\mathbf{x})$ corresponding to the Liapunov exponent $\Lambda_C = 0$, i.e. the central invariant subspace.

By projecting $\mathbf{g}(\mathbf{x})$ onto the invariant basis the Lagrangian evolution equation Eq.(3.73) for $\mathbf{g}(\mathbf{x})$ becomes

$$\mathbf{g}(\mathbf{x}(t)) = e^{\int_0^t (J(\mathbf{x}(\tau)) \mathbf{e}_C(\mathbf{x}(\tau)), \mathbf{e}_C(\mathbf{x}(\tau))) d\tau} (\mathbf{e}^C(\mathbf{x}_o) \mathbf{g}(\mathbf{x}_o)) \mathbf{e}_C(\mathbf{x}(t)) \quad (3.82)$$

¹²An attractor is defined as an attracting set containing a dense orbit (see [67] §1.5).

¹³For trivial limit sets such as an equilibrium point there still exists some debate over the definition of a central subspace.

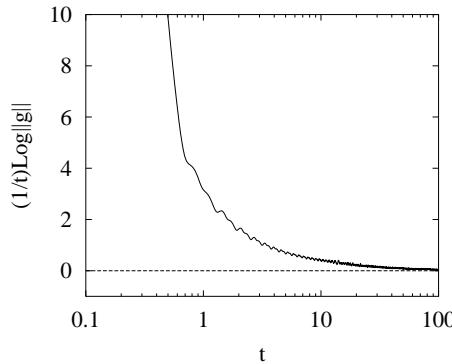


Fig. 3.8: Lorenz system. The quantity $\frac{1}{t} \log \|g(\mathbf{x}(t))\| \rightarrow 0$ as $t \rightarrow \infty$ ($\mathbf{x}_o = (1, 1, 1)^T$). The system's vector field $g(\mathbf{x})$ spans the central invariant subspace $E_C(\mathbf{x})$ for which $\Lambda_C = 0$.

(where $e_C(\mathbf{x})$ spans $E_C(\mathbf{x})$) thus indicating that as $t \rightarrow \infty$, $g(\mathbf{x}(t)) \sim e^{0t} \|g(\mathbf{x}_o)\| \frac{g(\mathbf{x}(t))}{\|g(\mathbf{x})\|}$ and holding this for any \mathbf{x}_o or any t , the relation simply reduces to a trivial identity which merely reminds us of the autonomous nature of the system.

3.8 The NTDRB technique to estimate invariant filtrations and Lyapunov exponents

The foregoing arguments have laid a basis for the definition of the "Natural Tangent Dynamics with Recurrent Biorthonormalizations" technique (NTDRB for short) which was already announced, and whose objective is that of determining the Liapunov-Oseledec invariant subspaces (we will see shortly that the technique determines the invariant filtrations rather than the invariant subspaces) and the relative Liapunov exponents, thereby performing a geometrical and timescale characterization of a dynamical system which is a crucial step towards the system's reduction. Let us see what we have learned so far:

- (i) the dynamical system in $\mathbf{x} \in C$ admits a unique invariant decomposition of $TC_{\mathbf{x}}$ which constitutes the optimal basis onto which we can project the system's vector field $g(\mathbf{x})$. The reason for this is that by following the evolution of unit vectors $e_h(\mathbf{x}) \in E^h(\mathbf{x})$ into each invariant subbundle by means of Eq.(3.55), not only the amplitudes of the vector field are decoupled, as shown by Eq.(3.70), but for large t we can also define a spectrum of intrinsic timescales given by the Liapunov exponents, because $f^h \sim e^{\Lambda_h t}$, and where the exponents may be estimated by Eq.(3.69);

- (ii) the natural evolution of a generic system of vectors $\{\mathbf{a}_i(\mathbf{x})\}$ is subject to a dominance principle whereby such system of vectors would asymptotically collapse onto the most unstable invariant subspace $E^1(\mathbf{x})$ corresponding to the greatest Liapunov exponent Λ_1 . The corresponding natural evolution of covectors $\{\mathbf{b}^i(\mathbf{x})\}$, biorthogonal to $\{\mathbf{a}_i(\mathbf{x})\}$, is also subject to the dominance principle so that the system of dual covectors would collapse onto the subspace spanned by $\mathbf{e}^n(\mathbf{x})$ dual to $\mathbf{e}_n(\mathbf{x}) \in E^n(\mathbf{x})$ corresponding to the smallest Liapunov exponent. Similarly by considering the natural evolution of generic unnormalized vectors and covectors as established in Oseledec's theorem Eq.(3.66), i.e.

$$\dot{\mathbf{x}}(t) = \mathbf{g}(\mathbf{x}(t)) \quad \dot{\mathbf{a}}_i(t) = J(\mathbf{x}_i(t)) \mathbf{a}_i(t) \quad \dot{\mathbf{c}}^i(t) = \mathbf{c}^i(t) J(\mathbf{x}^i(t)) \quad (3.83)$$

which do not preserve biorthogonality, the dominance principle will still be effective, forcing the system of vectors $\{\mathbf{a}_i(\mathbf{x})\}$ to collapse onto the invariant subspace $E^1(\mathbf{x})$. For the system of covectors $\{\mathbf{c}^i(\mathbf{x})\}$ we can write a similar equation to Eq.(3.73), where each invariant subbundle $F^h(\mathbf{x})$ is spanned by $\mathbf{f}^h(\mathbf{x})$:

$$\mathbf{c}^i(\mathbf{x}) = \sum_{h=1}^n e^{t\langle \mu_h^* \rangle} (\mathbf{c}^i(\mathbf{x}_o) \mathbf{f}_h(\mathbf{x}_o)) \mathbf{f}^h(\mathbf{x}) \quad i = 1, \dots, n \quad (3.84)$$

where $\mu_h^* = (\mathbf{c}^i J, \mathbf{c}^i)$. From the latter equation the dominance principle will act in such a way as to force $\{\mathbf{c}^i(\mathbf{x})\}$ to collapse onto the invariant subspace $F^1(\mathbf{x})$ spanned by $\mathbf{f}^1(\mathbf{x})$ corresponding to the greatest Liapunov exponent Λ_1^* . From the latter statement we also conclude that the system of invariant subspaces $\{E^i(\mathbf{x}), F^i(\mathbf{x})\}$, as defined in Oseledec's theorem, is in general not biorthonormal since Eqs.(3.83) do not preserve biorthogonality.

The basic idea at the core of the NTDRB technique is to follow the natural evolution of a generic unit basis $\{\mathbf{a}_i\}$ and recurrently contrast the dominance principle by "unfolding" such basis into a new basis $\{\tilde{\mathbf{a}}_i\}$ such that $\tilde{\mathbf{a}}_1 = \mathbf{a}_1$, i.e. it is left unchanged, whereas $\tilde{\mathbf{a}}_2$ is defined in such a way as to bear no projection on $\tilde{\mathbf{a}}_1$, $\tilde{\mathbf{a}}_3$ defined so as to have no projection on $\tilde{\mathbf{a}}_1$ and $\tilde{\mathbf{a}}_2$ and so on. If the new basis $\{\tilde{\mathbf{a}}_i\}$ is considered as a new initial condition for the natural vector evolution Eq.(3.76) then in spite of the tendency of $\tilde{\mathbf{a}}_1$ to still collapse towards $E^1(\mathbf{x})$, the second vector $\tilde{\mathbf{a}}_2$ will now tend to collapse onto invariant subspace $E^2(\mathbf{x})$ corresponding to the second largest Liapunov exponent for the reason that $\tilde{\mathbf{a}}_2$ was forced to have no projection onto $\tilde{\mathbf{a}}_1$. The same of course goes for the other vectors when they are forced to have no component along the subspace spanned by the preceding vectors and are let to evolve naturally according to vector dynamics. The end effect is that vectors

$\tilde{\mathbf{a}}_i$ will each be funnelled into the corresponding invariant subspace E^i with a natural ordering of decreasing Liapunov exponent.

Because the basis $\{\tilde{\mathbf{a}}_i\}$ is in general not orthogonal, we need a set of projection operators capable of yielding the projections of any vector onto the basis $\{\tilde{\mathbf{a}}_i\}$. More specifically, given two bases $\{\tilde{\mathbf{a}}_i\}$ and $\{\tilde{\mathbf{c}}^i\}$ of $TC_{\mathbf{x}}$ and $TC_{\mathbf{x}}^*$ respectively, the couple $\{\{\tilde{\mathbf{a}}_i\}, \{\tilde{\mathbf{c}}^i\}\}$ is said to be a biorthonormal system if $\tilde{\mathbf{c}}^i \tilde{\mathbf{a}}_j = \delta_j^i$. The set of linear functionals $\tilde{\mathbf{c}}^i : TC_{\mathbf{x}} \rightarrow R$ provide such a set of projection operators, so that for any $\mathbf{v} \in TC_{\mathbf{x}}$, the real number $\tilde{\mathbf{c}}^i \mathbf{v}$ is the projection of \mathbf{v} onto $\tilde{\mathbf{a}}_i$.

It appears clear, at this point, that the unfolding procedure of naturally evolving unit vectors $\{\mathbf{a}_i\}$ should in fact be a *biorthonormalizations procedure* of the couple $\{\{\mathbf{a}_i\}, \{\mathbf{c}^i\}\}$ (with $i = 1, \dots, n$) where $\{\mathbf{c}^i\}$ is a naturally evolving basis of unit covectors. Much like the skew product system of Eq.(3.66) in Oseledec's theorem, we will follow the natural evolution of unit vectors and covectors by means of the following system of equations:

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{g}(\mathbf{x}(t)) \quad (3.85)$$

$$\frac{d\mathbf{a}_i(\mathbf{x})}{dt} = J(\mathbf{x})\mathbf{a}_i(\mathbf{x}) - (J(\mathbf{x})\mathbf{a}_i(\mathbf{x}), \mathbf{a}_i(\mathbf{x})) \mathbf{a}_i(\mathbf{x}), \quad \mathbf{a}_i(\mathbf{x}) \in TC_{\mathbf{x}} \quad (3.86)$$

$$\frac{d\mathbf{c}^i(\mathbf{x})}{dt} = \mathbf{c}^i(\mathbf{x})J(\mathbf{x}) - (\mathbf{c}^i(\mathbf{x})J(\mathbf{x}), \mathbf{c}^i(\mathbf{x})) \mathbf{c}^i(\mathbf{x}), \quad \mathbf{c}^i(\mathbf{x}) \in TC_{\mathbf{x}}^* \quad (3.87)$$

and recurrently apply the biothogonalization procedure $\{\{\mathbf{a}_i\}, \{\mathbf{c}^i\}\} \xrightarrow{\text{Biorth}} \{\{\tilde{\mathbf{a}}_i\}, \{\tilde{\mathbf{c}}^i\}\}$.

In order to avoid any misunderstanding let us dissipate some doubts that may arise at this point, namely:

1. what would happen without the biorthonormalizations procedure? As we have seen in point (ii) above, system $\{\mathbf{a}_i\}$ would collapse onto $E^1(\mathbf{x})$ and $\{\mathbf{c}^i\}$ onto $F^1(\mathbf{x})$;
2. why didn't we choose Eq.(3.77) for the unit covector dynamics instead of Eq.(3.87)? This is because Eq.(3.77) preserves biorthogonality with system $\{\mathbf{a}_i\}$. This means that if the initial vector and covector bases are biorthogonal, the biorthonormalization procedure would simply coincide with the identity transformation leaving the two bases unchanged. This is to no avail to us because, as seen in point (ii) above, the two systems of vectors and covectors tend to collapse onto $E^1(\mathbf{x})$ and onto the dual of $E^n(\mathbf{x})$ respectively.

A rather cumbersome but effective way to visualize the action of the NTDRB technique is the following: we know from the dominance principle that if the unnormalized vectors and covectors $\mathbf{a}_1(t)$ and $\mathbf{c}^1(t)$ are left undisturbed in their natural evolution, then for sufficiently large t we have

that $\mathbf{a}_1(t) \sim e^{t\Lambda_1} \mathbf{e}_1(t)$ and $\mathbf{c}^1(t) \sim e^{t\Lambda_1^*} \mathbf{f}^1(t)$ where $\mathbf{e}_1 \in E^1$ and $\mathbf{f}^1 \in F^1$ where each covector \mathbf{f}^h lies in the invariant subspace F^h . Suppose further, that the system of invariant subspaces $\{E^h(\mathbf{x}), F^h(\mathbf{x})\}$ is a biorthogonal system. We will see later that this latter statement is not true, however for the sake of the present argument this is not strictly relevant. If this is so then each invariant subspace F^h is spanned by $\mathbf{e}^h \equiv \mathbf{f}^h$ (biorthonormal to $\mathbf{e}_h \in E^h$). In general for the set of vectors $\{\mathbf{a}_i\}$ we can use Eq.(3.73) (and Eq.(3.84) for the set of covectors $\{\mathbf{c}^i\}$) and placing $\mu_h = (J\mathbf{e}_h, \mathbf{e}_h)$ and $\mu^h = (\mathbf{e}^h J, \mathbf{e}^h)$ we can write for $k\tau < t < (k+1)\tau$

$$\mathbf{a}_i(t) = \sum_{h=1}^n e^{t\langle \mu_h \rangle} (\mathbf{e}^h(k\tau) \tilde{\mathbf{a}}_i(k\tau)) \mathbf{e}_h(t) \quad i = 1, \dots, n \quad (3.88)$$

$$\mathbf{c}^i(t) = \sum_{h=1}^n e^{t\langle \mu^h \rangle} (\tilde{\mathbf{c}}^i(k\tau) \mathbf{e}_h(k\tau)) \mathbf{e}^h(t) \quad i = 1, \dots, n \quad (3.89)$$

where we have taken the new biorthogonalized vector $\tilde{\mathbf{a}}_i$ at time $k\tau$ (where k is the number of biorthonormalizations performed up to time t) as the initial condition for the natural vector evolution. Because $\mathbf{c}^1 \equiv \tilde{\mathbf{c}}^1 \rightarrow \mathbf{e}^1$ and $\mathbf{a}_1 = \tilde{\mathbf{a}}_1 \rightarrow \mathbf{e}_1$ then for $i = 2$, since by definition $\tilde{\mathbf{c}}^1 \tilde{\mathbf{a}}_2 = 0$ and $\tilde{\mathbf{c}}^2 \tilde{\mathbf{a}}_1 = 0$, then $\mathbf{e}^1 \tilde{\mathbf{a}}_2 \rightarrow 0$ and $\tilde{\mathbf{c}}^2 \mathbf{e}^1 \rightarrow 0$. Therefore by Eq.(3.88) and the dominance principle $\mathbf{a}_2 \rightarrow \mathbf{e}_2$ and $\mathbf{c}^2 \rightarrow \mathbf{e}^2$. The same reasoning can then be applied to the remaining vectors and covectors. Another important consequence is that each un-normalized vector and covector $\bar{\mathbf{a}}_i, \bar{\mathbf{c}}^i$ (i.e. resulting from the integration of Eqs.(3.83)), parallel to \mathbf{a}_i and \mathbf{c}^i , for sufficiently large t , and provided the biorthonormalization procedure is enforced, will scale according to its Liapunov exponent, i.e. $\bar{\mathbf{a}}_i(\mathbf{x}(t)) \sim e^{t\Lambda_i} \mathbf{a}_i(\mathbf{x}(t))$ and $\bar{\mathbf{c}}^i(\mathbf{x}(t)) \sim e^{t\Lambda^i} \mathbf{c}^i(\mathbf{x}(t))$ where $\Lambda_i = \lim_{t \rightarrow \infty} \langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle$ and $\Lambda^i = \lim_{t \rightarrow \infty} \langle (\mathbf{c}^i J, \mathbf{c}^i) \rangle$. It is important to note that for simplicity we have assumed the system $\{E^h(\mathbf{x}), F^h(\mathbf{x})\}$ biorthogonal. The fact that it really is not will cause the vector \mathbf{a}_i to converge towards the invariant filtration \mathcal{V}_x rather than towards the invariant subspace E^i . Similarly \mathbf{c}^i will converge towards \mathcal{V}_x^* rather than F^i . As a result, the scaling of vectors will also be different from the one just described. We will deal with this issues later.

The NTDRB technique, therefore, consists ultimately in the iterative application of two consecutive steps: (a) the integration of the natural unit vector/covector evolution Eqs.(3.86,3.87), coupled with Eq.(3.85), for a system of n linearly independent vectors and covectors $\{\mathbf{a}_i(t)\}$ and $\{\mathbf{c}^i(t)\}$ along a trajectory $\Phi_t(\mathbf{x}_o) = \mathbf{x}(t)$; (b) the recurrent application of a biorthonormalization procedure, every time interval τ , that starting from the vectors and covectors $\{\mathbf{a}_h(\tau)\}$ and $\{\mathbf{c}^h(\tau)\}$, yields a new system of biorthonormal vectors/covectors $(\{\tilde{\mathbf{a}}_n(\tau)\}, \{\tilde{\mathbf{c}}^n(\tau)\})$ (with the further condition, as we will see, that the basis vectors $\{\tilde{\mathbf{a}}_n\}$ possess unit norm). The biorthonormal system obtained at

the end of the biorthonormalization step is set as the new initial condition for the vector and covector dynamical evolution along the system's trajectory for the next time interval τ .

Let us now illustrate the biorthonormalization procedure. First we start by setting:

$$\tilde{\mathbf{a}}_1 = \mathbf{a}_1, \quad \tilde{\mathbf{c}}^1 = \frac{\mathbf{c}^1}{\mathbf{c}^1 \mathbf{a}_1} \quad (3.90)$$

so that $\tilde{\mathbf{c}}^1 \tilde{\mathbf{a}}_1 = 1$. Note therefore that the first vector is left unchanged, letting it collapse freely onto subspace $E^1(\mathbf{x})$.

For the second vector $\tilde{\mathbf{a}}_2$, we set

$$\tilde{\mathbf{a}}_2 = \frac{\hat{\mathbf{a}}_2}{\|\hat{\mathbf{a}}_2\|}, \quad \hat{\mathbf{a}}_2 = \mathbf{a}_2 + \alpha_2^1 \tilde{\mathbf{a}}_1 \quad (3.91)$$

where unit norm is enforced and where the coefficient α_2^1 is found by enforcing the condition $\tilde{\mathbf{c}}^1 \tilde{\mathbf{a}}_2 = 0$

$$\tilde{\mathbf{c}}^1 \tilde{\mathbf{a}}_2 = 0 \implies \tilde{\mathbf{c}}^1 \mathbf{a}_2 + \alpha_2^1 \tilde{\mathbf{c}}^1 \tilde{\mathbf{a}}_1 = \tilde{\mathbf{c}}^1 \mathbf{a}_2 + \alpha_2^1 = 0 \implies \alpha_2^1 = -\tilde{\mathbf{c}}^1 \mathbf{a}_2. \quad (3.92)$$

The second covector $\tilde{\mathbf{c}}^2$ is found in the same way as

$$\tilde{\mathbf{c}}^2 = \frac{\hat{\mathbf{c}}^2}{\hat{\mathbf{c}}^2 \tilde{\mathbf{a}}_2}, \quad \hat{\mathbf{c}}^2 = \mathbf{c}^2 + \beta_1^2 \tilde{\mathbf{c}}^1 \quad (3.93)$$

where the coefficient β_1^2 is found by enforcing the condition $\tilde{\mathbf{c}}^2 \tilde{\mathbf{a}}_1 = 0$

$$\tilde{\mathbf{c}}^2 \tilde{\mathbf{a}}_1 = 0 \implies \mathbf{c}^2 \tilde{\mathbf{a}}_1 + \beta_1^2 \tilde{\mathbf{c}}^1 \tilde{\mathbf{a}}_1 = \mathbf{c}^2 \tilde{\mathbf{a}}_1 + \beta_1^2 = 0 \implies \beta_1^2 = -\mathbf{c}^2 \tilde{\mathbf{a}}_1. \quad (3.94)$$

The biorthonormalization procedure proceeds by finding $\tilde{\mathbf{a}}_n$ and $\tilde{\mathbf{c}}^n$ sequentially, starting from \mathbf{a}_n , \mathbf{c}^n and from the two sets of $n - 1$ orthonormal basis vectors $\{\tilde{\mathbf{a}}_i\}_{i=1}^{n-1}$ and $\{\tilde{\mathbf{c}}^i\}_{i=1}^{n-1}$ which have already been found¹⁴.

For the generic n -th vector $\tilde{\mathbf{a}}_n$ we set

$$\tilde{\mathbf{a}}_n = \frac{\hat{\mathbf{a}}_n}{\|\hat{\mathbf{a}}_n\|}, \quad \hat{\mathbf{a}}_n = \mathbf{a}_n + \sum_{i=1}^{n-1} \alpha_i^n \tilde{\mathbf{a}}_i \quad (3.95)$$

¹⁴This procedure yields a biorthonormal systems in the tangent/cotangent spaces under the condition that the quantities $\tilde{\mathbf{c}}^n \tilde{\mathbf{a}}_n$, $n = 1, \dots, N$, ($\hat{\mathbf{c}}^n = \mathbf{c}^1$) are all different from zero. This means that the biorthonormalization procedure may not be applied starting from two generic vector and covector bases in the tangent and cotangent spaces. However, if the initial vector/covector system ($\{\mathbf{a}_n(0)\}_{n=1}^N, \{\mathbf{c}^n(0)\}_{n=1}^N$) is biorthogonal, it can be shown that there exists a positive finite value τ_M , such that if $\tau < \tau_M$, the vector system ($\{\mathbf{a}_n(\tau)\}_{n=1}^N, \{\mathbf{c}^n(\tau)\}_{n=1}^N$) is such that all the denominators $\hat{\mathbf{c}}^n \tilde{\mathbf{a}}_n$, $n = 1, \dots, N$ entering the biorthonormalization procedure are different from zero and consequently the procedure can be applied.

where the coefficients $\{\alpha_n^i\}_{i=1}^{n-1}$ are found by enforcing the $n - 1$ conditions $\tilde{\mathbf{c}}^i \tilde{\mathbf{a}}_n = 0$, $i = 1, \dots, n - 1$, thus obtaining

$$\alpha_n^i = -\tilde{\mathbf{c}}^i \mathbf{a}_n, \quad i = 1, \dots, n - 1 \quad (3.96)$$

Subsequently, we find $\tilde{\mathbf{c}}^n$ by setting

$$\tilde{\mathbf{c}}^n = \frac{\hat{\mathbf{c}}^n}{\hat{\mathbf{c}}^n \tilde{\mathbf{a}}_n}, \quad \hat{\mathbf{c}}^n = \mathbf{c}^n + \sum_{i=1}^{n-1} \beta_i^n \tilde{\mathbf{c}}^i \quad (3.97)$$

and the coefficients $\{\beta_i^n\}_{i=1}^{n-1}$ are obtained by enforcing the conditions $\tilde{\mathbf{c}}^n \tilde{\mathbf{a}}_i = 0$, $i = 1, \dots, n - 1$ which yields

$$\beta_i^n = -\mathbf{c}^n \tilde{\mathbf{a}}_i, \quad i = 1, \dots, n - 1 \quad (3.98)$$

Operatively, NTDRB can be summarized into the following steps:

1. Start with two generic initial sets of linearly independent unit vectors $\{\mathbf{a}_i(0)\}$ and covectors $\{\mathbf{c}^i(0)\}$ spanning the tangent $TC_{\mathbf{x}_o}$ and the cotangent space $TC_{\mathbf{x}_o}^*$ at the starting point $\mathbf{x}(0) = \mathbf{x}_o$, corresponding to time $t = 0$, with the unique constraint that they should form a biorthogonal system, i.e. $\mathbf{c}^i(0) \cdot \mathbf{a}_j(0) = \theta_i \delta_j^i$ for $i, j = 1, \dots, N$ where $0 < |\theta_i| < 1$.
2. Integrate the skew-product system

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) \quad (3.99)$$

$$\dot{\mathbf{a}}_i = \mathbf{J}(\mathbf{x}) \mathbf{a}_i - (\mathbf{J}(\mathbf{x}) \mathbf{a}_i, \mathbf{a}_i) \mathbf{a}_i, \quad \mathbf{a}_i \in TC_{\mathbf{x}(t)}, i = 1, \dots, N \quad (3.100)$$

$$\dot{\mathbf{c}}^i = \mathbf{c}^i \mathbf{J}(\mathbf{x}) - (\mathbf{c}^i \mathbf{J}(\mathbf{x}), \mathbf{c}^i) \mathbf{c}^i, \quad \mathbf{c}^i \in TC_{\mathbf{x}(t)}^*, i = 1, \dots, N \quad (3.101)$$

in the time interval $(k - 1)\tau < t \leq k\tau$, where k is a biorthonormalization counter, to obtain $\mathbf{x}(k\tau)$ and the two sets of basis vectors $\{\mathbf{a}_i(k\tau)\}$ and covectors $\{\mathbf{c}^i(k\tau)\}$. Note that it is just as well acceptable to integrate the system:

$$\dot{\mathbf{x}}(t) = \mathbf{g}(\mathbf{x}(t)) \quad \dot{\overline{\mathbf{a}}}_i = J(\mathbf{x}) \overline{\mathbf{a}}_i(t) \quad \dot{\overline{\mathbf{c}}}^i = \overline{\mathbf{c}}^i(t) J(\mathbf{x}) \quad (3.102)$$

and subsequently normalize, i.e. place $\mathbf{a}_i(k\tau) = \overline{\mathbf{a}}_i(k\tau) / \|\overline{\mathbf{a}}_i(k\tau)\|$ and $\mathbf{c}^i(k\tau) = \overline{\mathbf{c}}^i(k\tau) / \|\overline{\mathbf{c}}^i(k\tau)\|$.

3. Apply the k -th biorthonormalization procedure to find $\{\tilde{\mathbf{a}}_i(k\tau)\}$ and $\{\tilde{\mathbf{c}}^i(k\tau)\}$
4. Because $\{\{\tilde{\mathbf{a}}_i(k\tau)\}, \{\tilde{\mathbf{c}}^i(k\tau)\}\}$ is a biorthonormal system it can be used to project the system's vector field (CSP approach) in the following way $\mathbf{g}(\mathbf{x}(k\tau)) = \sum_i \tilde{\mathbf{a}}_i(k\tau) f^i(k\tau)$ where

the amplitudes are $f^i(k\tau) = \tilde{\mathbf{c}}^i(k\tau)\mathbf{g}(\mathbf{x}(k\tau))$

5. Set the new initial conditions for the next integration step

$$\mathbf{a}_i(k\tau^+) = \tilde{\mathbf{a}}_i(k\tau) \quad i = 1, \dots, N \quad (3.103)$$

$$\mathbf{c}^i(k\tau^+) = \tilde{\mathbf{c}}^i(k\tau)/\|\tilde{\mathbf{c}}^i(k\tau)\| \quad i = 1, \dots, N \quad (3.104)$$

6. Because $\{\mathbf{a}_i(k\tau^+)\}$ and $\{\mathbf{c}^i(k\tau^+)\}$ have unit norm, estimate current Liapunov spectrum through the quantities $\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(k\tau)$ and $\langle (\mathbf{c}^i J, \mathbf{c}^i) \rangle(k\tau)$.
7. Repeat the procedure from step 2 for the next time interval setting $k = k + 1$.

3.9 A CSP-based approach using NTDRB

Given the dynamical system

$$\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x}) \quad (3.105)$$

where $\mathbf{x} \in C \subseteq R^n$, the classical CSP approach projects the dynamical system's vector field $\mathbf{g}(\mathbf{x}) \in TC_{\mathbf{x}}$ onto a set of n linearly independent basis vectors $\{\mathbf{a}_i(\mathbf{x})\}$ spanning $TC_{\mathbf{x}}$, where the projections are established through the use of covectors $\{\mathbf{c}^i(\mathbf{x})\}$ spanning $TC_{\mathbf{x}}^*$ where the system $\{\{\mathbf{a}_i(\mathbf{x})\}, \{\mathbf{c}^i(\mathbf{x})\}\}$ represents a biorthonormal system. Immediately after the NTDRB biorthonormalization step (3) above, the system $\{\{\tilde{\mathbf{a}}_i(k\tau)\}, \{\tilde{\mathbf{c}}^i(k\tau)\}\}$ is indeed a biorthonormal system and can therefore be used as a projection system for the vector field

$$\mathbf{g}(\mathbf{x}(k\tau)) = \sum_i \tilde{\mathbf{a}}_i(k\tau) f^i(k\tau), \quad f^i(k\tau) = \tilde{\mathbf{c}}^i(k\tau) \mathbf{g}(\mathbf{x}(k\tau)) \quad (3.106)$$

We should observe that the biorthonormalization step represents a discontinuity in the natural evolution of the system of unit vectors/covectors of Eqs.(3.85-3.87), which does not preserve the biorthonormality of such system. During the natural evolution of unit vectors $\{\mathbf{a}_i\}$ we could theoretically still evaluate a set of covectors $\{\mathbf{b}^i\}$ biorthonormal to $\{\mathbf{a}_i\}$ by simply placing $\mathbf{b}^i(k\tau^+) = \tilde{\mathbf{c}}^i(k\tau)$ as an initial condition and integrate, for an interval τ the equations

$$\frac{d\mathbf{b}^i(\mathbf{x})}{dt} = -\mathbf{b}^i(\mathbf{x}) J(\mathbf{x}) + (J(\mathbf{x}) \mathbf{a}_i(\mathbf{x}), \mathbf{a}_i(\mathbf{x})) \mathbf{b}^i(\mathbf{x}), \quad (3.107)$$

which indeed preserve biorthonormality. We have seen that in this case the projections may be

calculated, through Eq.(3.78) (which was derived from Eq.(3.59)) i.e.

$$\frac{df^i(\mathbf{x})}{dt} = (J(\mathbf{x})\mathbf{a}_i(\mathbf{x}), \mathbf{a}_i(\mathbf{x})) f^i(\mathbf{x}) \quad (3.108)$$

and therefore

$$f^i(\mathbf{x}) = e^{\int_0^t (J\mathbf{a}_i, \mathbf{a}_i) d\tau} f^i(\mathbf{x}_o) = e^{t \langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(t)} f^i(\mathbf{x}_o). \quad (3.109)$$

In fact, as the above equation shows, there is no need to evaluate the covectors $\{\mathbf{b}^i\}$ at all in order to estimate the projections. Moreover, because we postulated that each unit vector \mathbf{a}_i will gradually span the invariant subspaces E^i , provided a frequent biorthonormalization procedure is applied, we conclude that for sufficiently large t the amplitudes will each decay according to the corresponding Liapunov exponent:

$$f^i(\mathbf{x}) \sim e^{\Lambda_i} f^i(\mathbf{x}_o). \quad (3.110)$$

Such projections, or mode amplitudes, will furthermore be naturally ordered from the most unstable to the most stable, this meaning that $f^1(\mathbf{x})$ will evolve according to the largest Liapunov exponent Λ_1 and $f^n(\mathbf{x})$ according to smallest Liapunov exponent Λ_n . Of course if all exponents are negative (and the system collapses onto an equilibrium point) then the ordering of modes will be from the slowest to the fastest. Such an ordering is a direct consequence of the natural ordering of vectors \mathbf{a}_i which is itself a consequence of the dominance principle: \mathbf{a}_1 being left uneffected by the biorthogonalization procedure will collapse on E^1 , \mathbf{a}_2 on E^2 and so on.

As a first example of the application of the NTDRB technique let us consider the linear example first seen in Eq.(3.81) when illustrating the dominance principle. Fig.3.9(a) shows the time behavior (along a solution trajectory) of the quantities $(J\mathbf{a}_i, \mathbf{a}_i)$ and $(\mathbf{c}^i J, \mathbf{c}^i)$ $i = 1, 2, 3$, and as expected they individually converge towards the eigenvalues of coefficient matrix A , namely $\lambda_1 = -1$, $\lambda_2 = -3$ and $\lambda_3 = -5$ respectively. This is to be compared with Fig.3.6 showing that the biorthonormalization procedure does indeed contrast the dominance principle. By monitoring the quantities $1 - |(\mathbf{v}_i, \mathbf{a}_i)|$ and $1 - |(\mathbf{w}^i, \mathbf{c}^i)|$, Fig.3.9(c) shows that vectors \mathbf{a}_1 , \mathbf{a}_2 and \mathbf{a}_3 progressively align with the right eigenvectors \mathbf{v}_1 , \mathbf{v}_2 and \mathbf{v}_3 respectively (representing invariant subspaces E^1 , E^2 and E^3). At the same time covectors \mathbf{c}^1 , \mathbf{c}^2 and \mathbf{c}^3 respectively align with the left eigenvectors \mathbf{w}_1 , \mathbf{w}_2 and \mathbf{w}_3 (representing invariant subspaces F^1 , F^2 and F^3). Fig. 3.10(a) shows the projections $f^i = \mathbf{c}^i g$ (calculated at each time step irrespective of how often the biorthonormalization is performed, in this case every $\tau = 10^{-4}$ corresponding to one integration step). Clearly the three projections evolve respectively as $f^1 \sim e^{t\lambda_1} = e^{-t}$, $f^2 \sim e^{t\lambda_2} = e^{-3t}$ and $f^3 \sim e^{t\lambda_3} = e^{-5t}$ i.e.

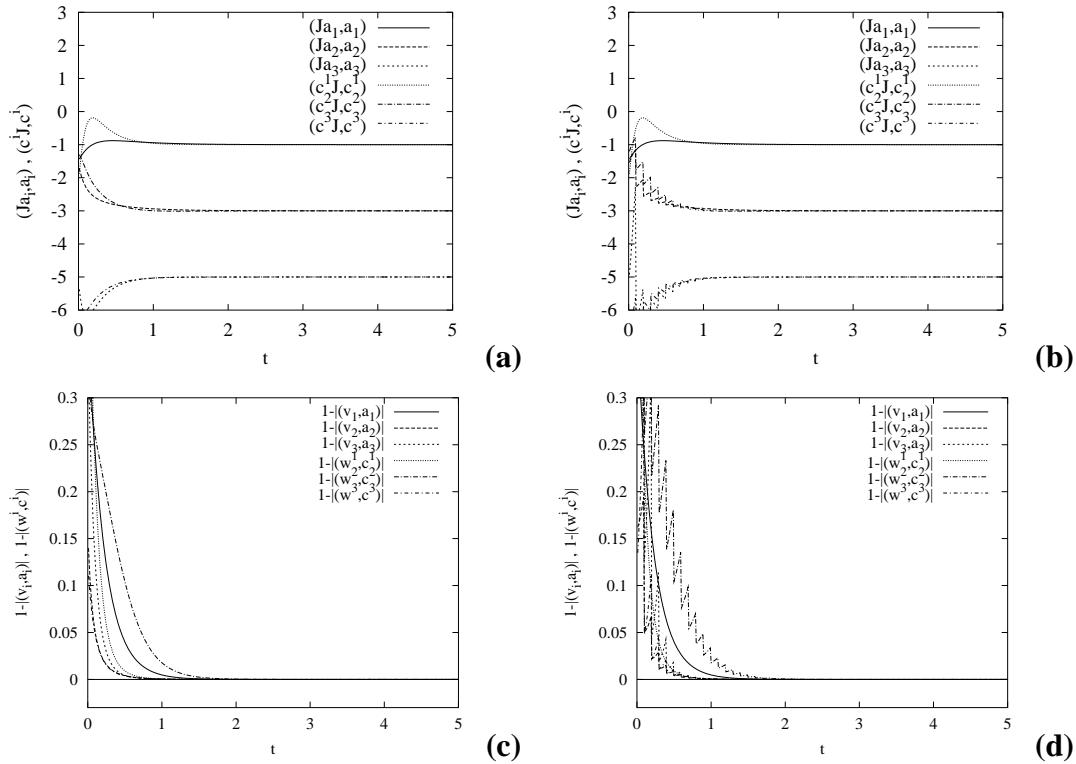


Fig. 3.9: Application of NTDRB: linear system Eq.(3.81). Figures on the left refer to $\tau = 10^{-4}$, figures on the right to $\tau = 0.1$. (a),(b) Behavior of $(J\mathbf{a}_i, \mathbf{a}_i)$ and $(\mathbf{c}^i J, \mathbf{c}^i)$ along a trajectory stemming from $\mathbf{x}_o = (1, 1, 1)^T$. (c),(d) Folding of vectors \mathbf{a}_i onto \mathbf{v}_i and covectors \mathbf{c}^i onto \mathbf{w}^i .

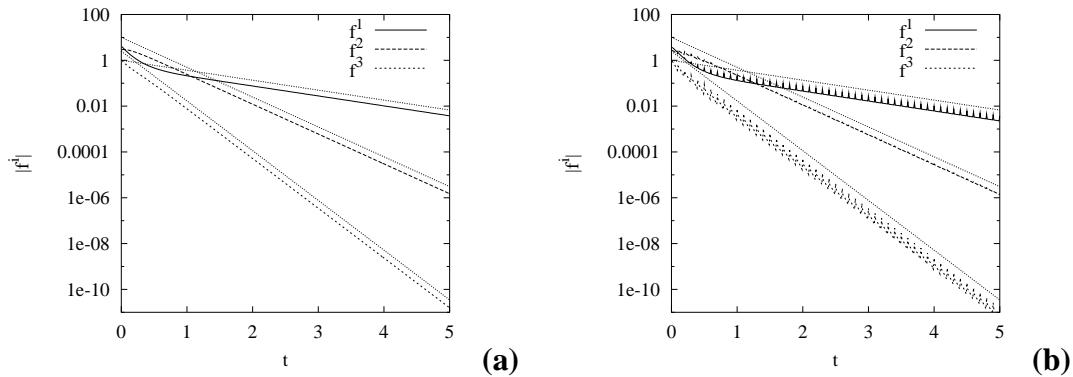


Fig. 3.10: Application of NTDRB: linear system Eq.(3.81). Figure on the left refers to $\tau = 10^{-4}$, figure on the right to $\tau = 0.1$. (a),(b) behavior of $f^i(\mathbf{x}(t))$, the dotted straight lines represent e^{-t} , e^{-3t} and e^{-5t} .

each with its own intrinsic timescale and free of mode coupling. Figs.3.9(b),(d) and Fig. 3.10(b) show the effects of adopting a far larger time interval ($\tau = 0.1$ corresponding to 10^3 integration steps) between two consecutive biorthonormalization steps. Initially the quantities $(J\mathbf{a}_2, \mathbf{a}_2)$ and $(J\mathbf{a}_3, \mathbf{a}_3)$, because the biorthogonalization is postponed for a long period of time, have a tendency to behave like $(J\mathbf{a}_1, \mathbf{a}_1)$ because they are left at the mercy of the dominance principle. Once the biorthonormalization procedure intervenes such quantities are eventually funnelled to their asymptotic values of $\lambda_2 = -3$ and $\lambda_3 = -5$. A very similar effect is also visible in Fig.3.9(d) where the

tendency of vectors \mathbf{a}_2 and \mathbf{a}_3 is that of aligning towards \mathbf{a}_1 if the biorthonormalization does not intervene. The discontinuous (saw-tooth) behavior characterizing the relaxation of $(J\mathbf{a}_i, \mathbf{a}_i)$ and $(\mathbf{c}^i J, \mathbf{c}^i)$ towards the eigenvalues fades away when the biorthonormalization step is more frequently applied. Indeed, the shorter τ is, the smoother is the convergence of the vectors \mathbf{a}_i towards the right eigenvectors. The time evolutions of $(J\mathbf{a}_i, \mathbf{a}_i)$ (and indeed of $(\mathbf{c}^i J, \mathbf{c}^i)$) obtained for different τ 's admit a single enveloping curve, which is the evolution obtained in the limit of $\tau \rightarrow 0$. The upper bound τ_c for τ ensuring that NTDRB will still yield the eigendirections of A is essentially set by the round-off errors associated with machine precision u_{mach} . Indeed, if λ_M and λ_m are the maximum and the minimum eigenvalue of the matrix A possessing real distinct eigenvalues, the upper bound τ_c for τ is given by: $\tau_c = -\log(u_{\text{mach}})/(\lambda_M - \lambda_m)$.

3.10 Non biorthogonality of the invariant system

In this section we will deal with an issue that was already briefly introduced, that of the lack of biorthogonality of the system of invariant subspaces $\{\{E^i(\mathbf{x})\}, \{F^i(\mathbf{x})\}\}$. In doing so we will also illustrate the application of the NTDRB technique to non linear problems.

It is intuitive, because NTDRB enforces a procedure whereby the system $\{\{\mathbf{a}_i(\mathbf{x})\}, \{\mathbf{c}^i(\mathbf{x})\}\}$ is biorthonormal after each biorthonormalization, that such system cannot converge to $\{\{E^i(\mathbf{x})\}, \{F^i(\mathbf{x})\}\}$ if the latter is not biorthogonal. The question then arises as to which geometric structure the NTDRB system $\{\{\mathbf{a}_i(\mathbf{x})\}, \{\mathbf{c}^i(\mathbf{x})\}\}$ is actually converging. Linear systems are not affected by this problem, as we have just seen in the previous section where the left and right eigenvectors do form a biorthogonal set of vectors and covectors constituting the invariant geometrical structure. We can however find evidence of the non biorthogonality of the invariant system in nonlinear problems (which is a consequence of the noncommutativity of the Jacobian $J(\mathbf{x})$ along trajectories). For this purpose let us concentrate on the Lorenz system in its chaotic attractor which we have used as a prototype of a chaotic system.

A straight forward technique to test the non biorthogonality of $\{\{E^i(\mathbf{x})\}, \{F^i(\mathbf{x})\}\}$ relies on the fact that the dominance principle assures us that the first vector \mathbf{a}_1 will unambiguously collapse onto the first invariant subspace $E^1(\mathbf{x})$ characterized by the greatest Liapunov exponent and at the same time the first covector \mathbf{c}^1 will collapse onto $F^1(\mathbf{x})$. So at least the first invariant subspace both in the tangent and cotangent spaces can be recovered by NTDRB (where we always start from a random biorthonormal vector/covector system). Moreover, as we have seen in a previous section, we know that the system's vector field $\mathbf{g}(\mathbf{x})$ spans at all times the central invariant subspace $E^C(\mathbf{x})$

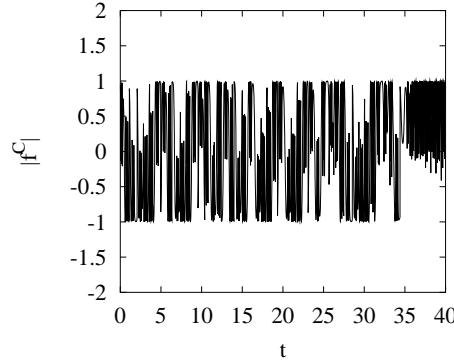


Fig. 3.11: Projection $f^1(t) = \mathbf{c}^1(\mathbf{x})\mathbf{e}_C(\mathbf{x})$ for the Lorenz system along a trajectory. The fact that $\mathbf{c}^1\mathbf{e}_C \neq 0$ is evidence of the non biorthogonality of the system $\{\{E^i(\mathbf{x})\}, \{F^i(\mathbf{x})\}\}$.

characterized by a Liapunov exponent equal to zero. A simple check on biorthogonality is therefore to consider the projection:

$$f^1(t) = \mathbf{c}^1(\mathbf{x})\mathbf{e}_C(\mathbf{x}). \quad (3.111)$$

where $\mathbf{e}_C(\mathbf{x}) = \frac{\mathbf{g}(\mathbf{x})}{\|\mathbf{g}(\mathbf{x})\|}$. If the system $\{\{E^i(\mathbf{x})\}, \{F^i(\mathbf{x})\}\}$ is biorthogonal one would expect $\mathbf{c}^1\mathbf{e}_C = 0$ at all times. In fact, for the Lorenz system we have $\mathbf{c}^1\mathbf{e}_C \neq 0 \forall t$ as Fig.(3.11) shows. Let

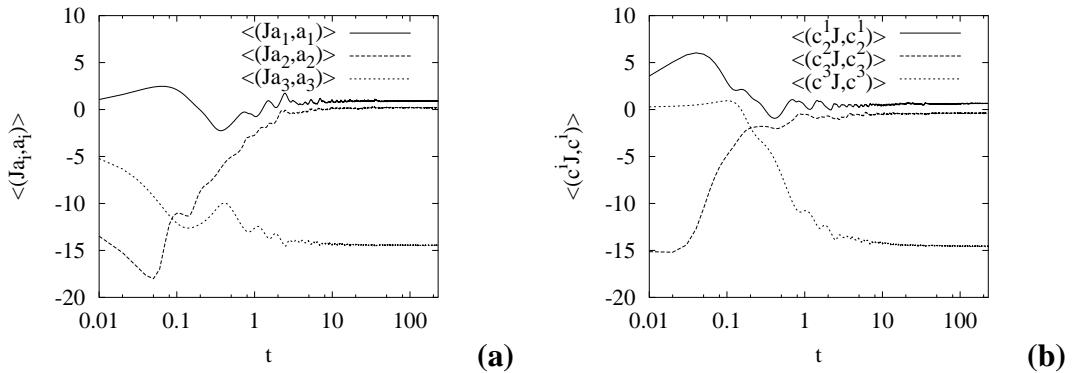


Fig. 3.12: Lorenz system. (a) $\langle(J\mathbf{a}_i, \mathbf{a}_i)\rangle(t)$ (b) $\langle(\mathbf{c}^i J, \mathbf{c}^i)\rangle(t)$, $i = 1, 2, 3$ as estimated via NTDRB.

us now monitor the time averaged quantities $\langle(J\mathbf{a}_i, \mathbf{a}_i)\rangle$ and $\langle(\mathbf{c}^i J, \mathbf{c}^i)\rangle$, $i = 1, 2, 3$ along a Lorenz system trajectory, as they are provided by the NTDRB technique. Figs.(3.12)(a) and (b) show such quantities as a function of time and from them it is evident that NTDRB contrasts the dominance principle funnelling each vector into the corresponding invariant subspace so that as $t \rightarrow \infty$ we have $\langle(J\mathbf{a}_i, \mathbf{a}_i)\rangle \rightarrow \Lambda_i$ and $\langle(\mathbf{c}^i J, \mathbf{c}^i)\rangle \rightarrow \Lambda_i^*$. This should definitely be compared with Fig.(3.7) where the dominance principle was not contrasted resulting into $\langle(J\mathbf{a}_i, \mathbf{a}_i)\rangle \rightarrow \Lambda_1$ i.e. vectors $\{\mathbf{a}_i\}$ all collapsing onto the first invariant subspace. We should observe that we have defined Liapunov exponents as the time-averaged logarithm of the norm of a vector element evolving within each

invariant subspace as $t \rightarrow \infty$. We have however implied, with NTDRB, the biorthogonality of the invariant subspaces, the lack of which may result in an imprecise estimation of the Liapunov exponents. A very accurate method for estimating Liapunov exponents is the exterior algebra method, which was already briefly discussed and which uses the time-averaged evolution of measure elements within each i -th dimensional subspace \overline{E}^i of the invariant filtration (rather than vectors within each one dimensional invariant subspace E^i). We will illustrate such a technique shortly (and explain why when applied to the NTDRB vectors it yields the correct Liapunov exponents), but for the time being let us summarize the results in the following table:

	Λ_1	Λ_2	Λ_3
$\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle$	0.91	0.19	-14.45
Exterior Algebra	0.91	0	-14.58

Although the discrepancies are qualitatively significant, they are not particularly relevant. Such discrepancies, however do exist, suggesting, once again the non alignement of the NTDRB vector/covector bases with respect to the invariant subspaces E^i and F^i . Note the zero Liapunov exponent detected by the exterior algebra method corresponding to the central invariant subspace $E^2 = E^C$.

Let us then answer the question as to which geometric structure the NTDRB system represented by $\{\{\mathbf{a}_i(\mathbf{x})\}, \{\mathbf{c}^i(\mathbf{x})\}\}$ is actually converging. Let us suppose that subspaces $E^h(\mathbf{x})$ and $F^h(\mathbf{x})$ are spanned by a biorthogonal system of unit vectors and of covectors $\mathbf{e}_h(\mathbf{x})$ and $\mathbf{f}^h(\mathbf{x})$ such that $\|\mathbf{e}_h\| = 1$, and $\mathbf{f}^h \mathbf{e}_h = 1$, $h = 1, \dots, N$. We know that after a transient, the first vector \mathbf{a}_1 and covector \mathbf{c}^1 obtained from NTDRB align with \mathbf{e}_1 and \mathbf{f}^1 respectively, i.e. $\mathbf{a}_1 = \mathbf{e}_1$, $\mathbf{c}^1 = \mathbf{f}^1$. It is possible to show analytically that, once \mathbf{a}_1 and \mathbf{c}^1 have converged towards the first invariant subspace, the second vector/covector \mathbf{a}_2 and \mathbf{c}^2 at the end of the biorthonormalization procedure attain the following expressions:

$$\mathbf{a}_2 = A^2(\mathbf{e}_2 - F_2^1 \mathbf{e}_1) \quad \mathbf{c}^2 = B_2(\mathbf{f}^2 - F_1^2 \mathbf{f}^1), \quad (3.112)$$

where A^2 , B_2 are normalization constants and $F_2^1 = \mathbf{f}^1 \mathbf{e}_2$, $F_1^2 = \mathbf{f}^2 \mathbf{e}_1$. Eq. (3.112) is obtained through elementary algebraic manipulations by enforcing the biorthonormalization procedure described in the previous section. This result can be generalized as follows: for a generic h , \mathbf{a}_h and \mathbf{c}^h converge, after the transient, to a linear combination of the first h vector/covectors of the invariant bases, i.e.

$$\mathbf{a}_h = A^h \left(\mathbf{e}_h + \sum_{k=1}^{h-1} M_h^k \mathbf{e}_k \right) \quad \mathbf{c}^h = B_h \left(\mathbf{f}^h + \sum_{k=1}^{h-1} N_h^k \mathbf{f}^k \right), \quad (3.113)$$

where A^h, B_h are normalization constants, and M_h^k, N_k^h are real coefficients related to the values of the functionals $\mathbf{f}^h \mathbf{e}_k$ characterizing the invariant bases.

Eqs. (3.112),(3.113) highlight that the discrepancies between the quantities $\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle$ and the Lyapunov spectrum (obtained by the exterior algebra method) are a consequence of the nonvanishing values of F_2^1, F_1^2 (and of M_h^k and N_k^h in general). The structure of Eqs. (3.112),(3.113) further indicates that albeit each of the one-dimensional invariant subspace $E^h(\mathbf{x}), F^h(\mathbf{x})$ cannot be recovered from the vectors $\mathbf{a}_h, \mathbf{c}^h$ obtained via NTDRB, the structure of the invariant filtrations $\mathcal{V}_{\mathbf{x}}$ and $\mathcal{V}_{\mathbf{x}}^*$, as defined by Liapunov's theorem, can be directly obtained from them, since the subspaces forming these two filtrations are given by:

$$\begin{aligned} \overline{E}_{\mathbf{x}}^1 &= \text{span}\{\mathbf{a}_1\} & \overline{F}_{\mathbf{x}}^1 &= \text{span}\{\mathbf{c}^1\} \\ \overline{E}_{\mathbf{x}}^2 &= \text{span}\{\mathbf{a}_1, \mathbf{a}_2\} & \overline{F}_{\mathbf{x}}^2 &= \text{span}\{\mathbf{c}^1, \mathbf{c}^2\} \\ \vdots & & \vdots & \\ \overline{E}_{\mathbf{x}}^h &= \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_h\} & \overline{F}_{\mathbf{x}}^h &= \text{span}\{\mathbf{c}^1, \dots, \mathbf{c}^h\} \\ \vdots & & \vdots & \\ \overline{E}_{\mathbf{x}}^n &= \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_n\} = TC_{\mathbf{x}} & \overline{F}_{\mathbf{x}}^n &= \text{span}\{\mathbf{c}^1, \dots, \mathbf{c}^n\} = TC_{\mathbf{x}}^*. \end{aligned} \tag{3.114}$$

This is the result that we have long anticipated: the NTDRB technique yields a pointwise estimate of the invariant geometrical structure of a dynamical system in terms of its h -dimensional Liapunov Oseledec invariant filtrations, $\overline{E}^h(\mathbf{x}), \overline{F}^h(\mathbf{x})$ as opposed to the one-dimensional invariant subspaces $E^h(\mathbf{x}), F^h(\mathbf{x}), h = 1, \dots, n$. This should not concern us because the concept of invariance applies to measure elements in filtrations just as it does for vectors in the one-dimensional invariant subspaces. This means that a vector in $E^1(\mathbf{x}_o) = \overline{E}^1(\mathbf{x}_o)$ is mapped invariantly (by the natural vector dynamics induced by the dynamical system) into a vector of $E^1(\mathbf{x}) = \overline{E}^1(\mathbf{x})$. At the same time, while a vector in $E^2(\mathbf{x}_o)$ is mapped invariantly into a vector of $E^2(\mathbf{x})$, an area element (which can be viewed as a vector product $\mathbf{a}_1 \times \mathbf{a}_2$) of the second filtration $\overline{E}^2(\mathbf{x}_o) = E^1(\mathbf{x}_o) \oplus E^2(\mathbf{x}_o)$ will be mapped invariantly into an area element of $\overline{E}^2(\mathbf{x}) = E^1(\mathbf{x}) \oplus E^2(\mathbf{x})$. Clearly a volume element of $\overline{E}^3(\mathbf{x}_o)$ (viewable as a triple product $\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3$) will be mapped into a volume element of $\overline{E}^3(\mathbf{x})$. For higher dimensions it is necessary to resort to exterior algebra in order to define measure elements. Fig.(3.13) is a visualization of the first two invariant filtrations for the Lorenz system. Fig.(3.13)(a) is a representation of the three components of the first NTDRB vector \mathbf{a}_1 i.e. (a_1^1, a_1^2, a_1^3) , whereas Fig.(3.13)(b) represents the three components of vector $\mathbf{n}_1 = \mathbf{a}_1 \times \mathbf{a}_2 / \|\mathbf{a}_1 \times \mathbf{a}_2\|$, i.e. (n_1^1, n_1^2, n_1^3) .

Liapunov exponents (defined as time averaged evolution of vector norm rate coefficients $\mu_i =$

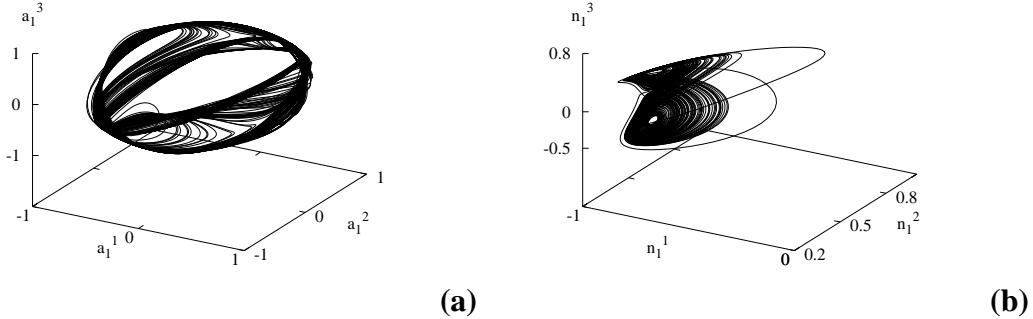


Fig. 3.13: Lorenz system. (a) components (a_1^1, a_1^2, a_1^3) . (b) components (n_1^1, n_1^2, n_1^3) of $\mathbf{n}_1 = \mathbf{a}_1 \times \mathbf{a}_2 / \|\mathbf{a}_1 \times \mathbf{a}_2\|$.

$(J\mathbf{e}_i, \mathbf{e}_i)$ within each invariant one-dimensional subspace) can be readily recovered by considering measure elements in their natural evolution within invariant filtrations. Defining $m_h(\mathbf{x}_o) = m_h(\tilde{\mathbf{e}}_1(\mathbf{x}_o), \dots, \tilde{\mathbf{e}}_h(\mathbf{x}_o))$ as the initial measure of an h -dimensional measure element (an hyper-rectangle) spanned by the invariant directions $\tilde{\mathbf{e}}_1(\mathbf{x}_o) \in E^1(\mathbf{x}_o), \dots, \tilde{\mathbf{e}}_h(\mathbf{x}_o) \in E^1(\mathbf{x}_o)$ it can be shown that such measure element is naturally mapped by the dynamical system¹⁵ as $m_h(\mathbf{x}) = m_h(\mathbf{x}_o)e^{t\langle \mu_h^G \rangle}$ where $\mu_h^G(\mathbf{x}(t))$ is a geometric measure element growth rate for which we can show¹⁶ that:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log m_h = \lim_{t \rightarrow \infty} \langle \mu_h^G \rangle = \Lambda_1 + \dots + \Lambda_h \quad (3.115)$$

and therefore for sufficiently large t , the measure of h -dimensional elements belonging to \overline{E}^h scale like $m_h(t) \sim e^{t(\Lambda_1 + \dots + \Lambda_h)}$.

We know that the evolution of the norm of a vector $\tilde{\mathbf{e}}_1 \in E^1 = \overline{E}^1$ is given by

$$\frac{d\|\tilde{\mathbf{e}}_1\|^2}{dt} = 2(J\mathbf{e}_1, \mathbf{e}_1)\|\tilde{\mathbf{e}}_1\|^2 \quad (3.116)$$

where \mathbf{e}_1 is a unit vector spanning E^1 . Therefore, as we know, the rate of growth of vector norm in $E^1 = \overline{E}^1$ is given by

$$\mu_1 = \mu_1^G = (J\mathbf{a}_1, \mathbf{a}_1) \quad (3.117)$$

where \mathbf{a}_1 is the first unit NTDRB vector which we know to eventually span $E^1 = \overline{E}^1$. Similarly,

¹⁵We could add at this point that a dynamical system may be considered as an "advection machine" of all possible geometrical entities definable within the phase space [79]. So the advection of points is described by $\dot{\mathbf{x}} = \mathbf{g}(\mathbf{x})$, the advection of vectors by $\dot{\mathbf{a}} = J(\mathbf{x})\mathbf{a}$. Similarly we can derive equations describing the advection of an h -dimensional measure element.

¹⁶A way to see this is to enforce a Gram-Schmidt procedure to yield a system of n vectors $\{\hat{\mathbf{e}}_h\}$ orthogonal to the invariant vectors $\{\tilde{\mathbf{e}}_h\}$, so that $\tilde{\mathbf{e}}_i \cdot \hat{\mathbf{e}}_j = 0$ for $i \neq j$. It can be easily shown that $m_h(\mathbf{x}) = \prod_{k=1}^h \|\hat{\mathbf{e}}_k(\mathbf{x})\|$ and since $\hat{\mathbf{e}}_h \in \text{span}\{\tilde{\mathbf{e}}_1, \dots, \tilde{\mathbf{e}}_h\}$ then $\hat{\mathbf{e}}_h \rightarrow \text{span}\{\tilde{\mathbf{e}}_h\}$. Therefore $\lim_{t \rightarrow \infty} \frac{1}{t} \log m_h = \sum_{k=1}^h \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\hat{\mathbf{e}}_k\| = \sum_{k=1}^h \Lambda_k$.

we can define μ_2^G as a geometric growth rate coefficient for the measure, $m_2 = \|\tilde{\mathbf{e}}_1 \times \tilde{\mathbf{e}}_2\|$, of an area element of \overline{E}^2 . Because $d\tilde{\mathbf{e}}_1/dt = J\tilde{\mathbf{e}}_1$ and $d\tilde{\mathbf{e}}_2/dt = J\tilde{\mathbf{e}}_2$ then

$$\frac{d(\tilde{\mathbf{e}}_1 \times \tilde{\mathbf{e}}_2)}{dt} = J\tilde{\mathbf{e}}_1 \times \tilde{\mathbf{e}}_2 + \tilde{\mathbf{e}}_1 \times J\tilde{\mathbf{e}}_2 \quad (3.118)$$

and by taking the scalar product with $\tilde{\mathbf{e}}_1 \times \tilde{\mathbf{e}}_2$ on both sides we obtain

$$\frac{d\|\tilde{\mathbf{e}}_1 \times \tilde{\mathbf{e}}_2\|^2}{dt} = \frac{2(\mathbf{e}_1 \times \mathbf{e}_2 + \mathbf{e}_1 \times J\mathbf{e}_2, \mathbf{e}_1 \times \mathbf{e}_2)}{\|\mathbf{e}_1 \times \mathbf{e}_2\|^2} \|\tilde{\mathbf{e}}_1 \times \tilde{\mathbf{e}}_2\|^2 \quad (3.119)$$

where \mathbf{e}_1 and \mathbf{e}_2 are two unit vectors such that $\overline{E}^2 = \text{span}\{\mathbf{e}_1, \mathbf{e}_2\}$. Therefore the growth rate coefficient for an area element of \overline{E}^2 may be calculated by

$$\mu_2^G = \frac{(\mathbf{a}_1 \times \mathbf{a}_2 + \mathbf{a}_1 \times J\mathbf{a}_2, \mathbf{a}_1 \times \mathbf{a}_2)}{\|\mathbf{a}_1 \times \mathbf{a}_2\|^2} \quad (3.120)$$

where \mathbf{a}_1 and \mathbf{a}_2 are the first two NTDRB unit vectors such that, eventually, $\overline{E}^2 = \text{span}\{\mathbf{a}_1, \mathbf{a}_2\}$. The same reasoning can be applied to a volume element of \overline{E}^3 whose growth rate coefficient is:

$$\mu_3^G = \frac{\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3}{|\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3|^2} [\mathbf{J}\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3 + \mathbf{a}_1 \cdot \mathbf{J}\mathbf{a}_2 \times \mathbf{a}_3 + \mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{J}\mathbf{a}_3]. \quad (3.121)$$

The above expressions can therefore be used in conjunction with NTDRB to yield the exact Liapunov exponents through Eq.(3.115). When generalized to higher dimensions this technique can be referred to as the exterior algebra technique for the determination of Liapunov exponents. Fig. 3.10 shows how the exact Liapunov spectrum may be estimated with this technique for the Lorenz system. Using Eq.(3.115), Λ_2 for instance may be recovered from the quantity $\langle \mu_2^G \rangle(t) - \langle \mu_1^G \rangle(t)$ and Λ_3 from $\langle \mu_3^G \rangle(t) - \langle \mu_2^G \rangle(t)$. On the same figure the evolution of vector norm growth rates $\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(t)$ (seen in Fig. 3.12) are also shown for comparison. Note that with the measre element evolution it is possible to recover the zero Liapunov exponent $\Lambda_2 = \Lambda_C = 0$ corresponding to the central subspace. As we had seen previously this was not possible using the vector norm growth rates.

We will now address yet a further aspect concerning NTDRB and its ability to detect the Liapunov-Oseledec invariant filtrations $\{\{\overline{E}^h\}, \{\overline{F}^h\}\}$ rather than the non biorthonormal system of invariant subspaces $\{\{E^h\}, \{F^h\}\}$. Namely we will examine the behavior of the projections $f^h(\mathbf{x}(t))$ of the system's vector field $\mathbf{g}(\mathbf{x})$ onto the NTDRB unit basis $\{\mathbf{a}_h(\mathbf{x})\}$ using the dual biorthonormal basis $\{\mathbf{c}^h(\mathbf{x})\}$. We will assume that the biorthonormalization procedure is carried out at each integration step ($\tau = \Delta t$) so that the effects of large τ are negeible. As we have seen, projections may be

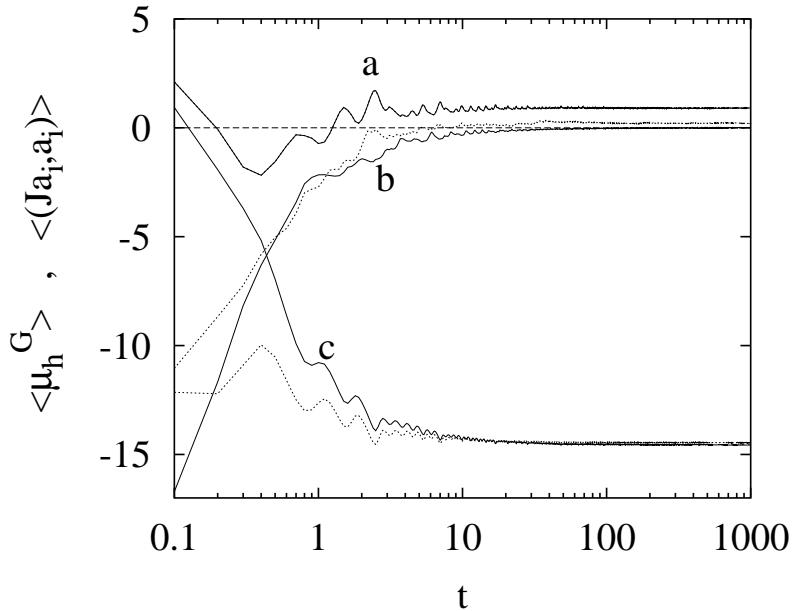


Fig. 3.14: Lorenz system: exterior algebra method for estimating Liapunov exponents (continuous lines). (a) $\langle \mu_1^G \rangle(t) = \langle (J\mathbf{a}_1, \mathbf{a}_1) \rangle(t)$. (b) $\langle \mu_2^G \rangle(t) - \langle \mu_1^G \rangle(t)$. (c) $\langle \mu_3^G \rangle(t) - \langle \mu_2^G \rangle(t)$. Dotted lines refer to vector norm growth rates $\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(t)$ as already shown in Fig. 3.12 (a).

calculated through Eq.(3.106) or alternatively by solving Eq.(3.108). In summary we have that $f^h(\mathbf{x}) = \mathbf{c}^h(\mathbf{x})\mathbf{g}(\mathbf{x})$ and their behavior is governed by $f^h(\mathbf{x}) = e^{t\langle (J\mathbf{a}_h, \mathbf{a}_h) \rangle(t)} f^h(\mathbf{x}_o)$ being \mathbf{a}_h the NTDRB unit vectors and \mathbf{c}^h their biorthonormal dual covectors. In the linear example we have seen that, because $\{\{E^h\}, \{F^h\}\}$ is a biorthonormal system and hence $\mathbf{a}_h \rightarrow E^h$, for sufficiently large t , $f^h \sim e^{t\Lambda_h}$. What happens more generally with non linear systems within their limit sets? Suppose, for the sake of argument, that the NTDRB system $\{\{\mathbf{a}_h\}, \{\mathbf{c}^h\}\}$ did indeed converge to the invariant system $\{\{E^h\}, \{F^h\}\}$, which is therefore forcibly thought as biorthonormal. Because we know that $\mathbf{g}(\mathbf{x})$ spans the central invariant subspace at all times (associated with the zero Liapunov exponent) then the projections $f^h(\mathbf{x}) = \mathbf{c}^h(\mathbf{x})\mathbf{g}(\mathbf{x}) = 0$ for $h \neq h^*$ where h^* is the number of non-negative Liapunov exponents and therefore E^{h^*} is the central subspace. On the other hand we would expect $f^{h^*}(\mathbf{x}) = \mathbf{c}^{h^*}(\mathbf{x})\mathbf{g}(\mathbf{x}) = \|\mathbf{g}(\mathbf{x})\|$. Therefore the scaling $f^h(\mathbf{x}) = e^{t\langle (J\mathbf{a}_h, \mathbf{a}_h) \rangle(t)} f^h(\mathbf{x}_o)$ would not translate into $f^h \sim e^{t\Lambda_h}$ because $f^h(\mathbf{x}_o) = 0$ for $h \neq h^*$. In reality we know that the NTDRB system $\{\{\mathbf{a}_h\}, \{\mathbf{c}^h\}\}$ does not coincide with $\{\{E^h\}, \{F^h\}\}$ but rather with the system of invariant filtrations. For the projections, therefore, we observe a different behavior than the one

described, which may be summarized as follows:

$$f^h(t)/\|\mathbf{g}(t)\| = \begin{cases} \mathcal{O}(1) & h \leq h^* \\ e^{t\Lambda_h} & h > h^*. \end{cases} \quad (3.122)$$

This behavior can be observed in Fig.(3.10) for the Lorenz system in its chaotic attractor. In this

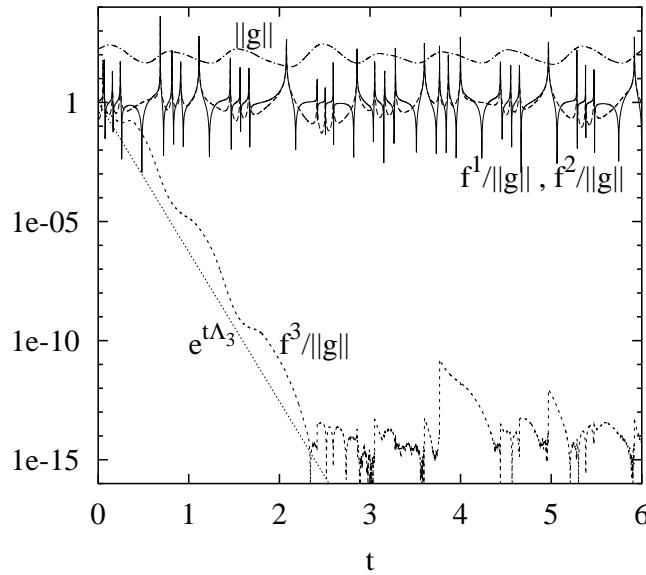


Fig. 3.15: Lorenz system. Projections $f^h(\mathbf{x})/\|\mathbf{g}(\mathbf{x})\| = \mathbf{c}^h(\mathbf{x})\mathbf{g}(\mathbf{x})/\|\mathbf{g}(\mathbf{x})\|$.

case the time averaged vector norm growth rates $\langle \mu_i \rangle = \langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle$ tend to the approximate (as we have seen) Liapunov spectrum $\Lambda_1 = 0.91$, $\Lambda_2 = 0.19$ and $\Lambda_3 = -14.45$ and hence $h^* = 2$ (i.e. there exists one unstable subspace and a central subspace, which NTDRB, with the value 0.19, still views as unstable). So we observe $f^h(t)/\|\mathbf{g}(t)\| = \mathcal{O}(1)$ for $h = 1, 2$ and $f^3(t)/\|\mathbf{g}(t)\| \sim e^{t\Lambda_3}$. Although we will not attempt a rigorous proof of Eq.(3.122) we will try nonetheless to understand why NTDRB yields such scaling for the projections. We know NTDRB reconstructs the system of

invariant filtrations in the following way:

$$\begin{aligned}
 \overline{E}_{\mathbf{x}}^1 &= \text{span}\{\mathbf{a}_1\} = E_{\mathbf{x}}^1 \\
 &\vdots \\
 \overline{E}_{\mathbf{x}}^{h^*} &= \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{h^*}\} = E_{\mathbf{x}}^1 \oplus \dots \oplus E_{\mathbf{x}}^{h^*} \\
 &\vdots \\
 \overline{E}_{\mathbf{x}}^n &= \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{h^*}\} \oplus \text{span}\{\mathbf{a}_{h^*+1}, \dots, \mathbf{a}_n\} = \overline{E}_{\mathbf{x}}^{h^*} \oplus (E_{\mathbf{x}}^{h^*+1} \oplus \dots \oplus E_{\mathbf{x}}^n).
 \end{aligned} \tag{3.123}$$

We know that $\mathbf{g}(t) \in E_{\mathbf{x}}^{h^*}$ where as mentioned $E_{\mathbf{x}}^{h^*}$ is the central subspace. However NTDRB has no perception of $E_{\mathbf{x}}^{h^*}$, and what we can say at best is that asymptotically $\mathbf{g}(t) \in \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{h^*}\}$ and not $\text{span}\{\mathbf{a}_{h^*+1}, \dots, \mathbf{a}_n\}$ because $E_{\mathbf{x}}^{h^*} \in \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{h^*}\}$. Because $\|\mathbf{g}\|$ is bounded and because there is no reason to expect $\mathbf{g}(t)$ to have a greater mean projection on any of the vectors constituting $\text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{h^*}\}$ we should therefore expect (on average) $f^h/\|\mathbf{g}\| = \mathcal{O}(1)$ for $h = 1, \dots, h^*$. The remaining ones are projections onto vectors constituting $\text{span}\{\mathbf{a}_{h^*+1}, \dots, \mathbf{a}_n\}$ which represents the collection of all the stable invariant subspaces characterized by negative Liapunov exponents. Because initially $f^h(\mathbf{x}_o) \neq 0$ with $h = h^* + 1, \dots, n$ almost everywhere we should still expect the scaling $f^h(\mathbf{x}) = e^{t\langle(J\mathbf{a}_h, \mathbf{a}_h)\rangle(t)} f^h(\mathbf{x}_o)$ where $\langle(J\mathbf{a}_h, \mathbf{a}_h)\rangle(t) \rightarrow \Lambda_h < 0$ with Λ_h an approximation of the corresponding Liapunov exponent. In reality we observe the scaling $f^h(\mathbf{x}) = e^{t\langle\mu_h^G\rangle(t)} f^h(\mathbf{x}_o)$ where μ_h^G is the geometric growth rate coefficient. We won't attempt a rigorous proof for this result which, however, underlines the fact that projections onto the collection of stable subspaces scale in fact with the exact negative Liapunov exponents.

As a further example let us consider the five-dimensional Franceschini Tebaldi model [40] (which will be referred to as the FT model):

$$\begin{aligned}
 \dot{x}_1 &= -2x_1 + 4x_2 x_3 + 4x_4 x_5 \\
 \dot{x}_2 &= -9x_2 + 3x_1 x_3 \\
 \dot{x}_3 &= -5x_3 - 7x_1 x_2 + r \\
 \dot{x}_4 &= -5x_4 - x_1 x_5 \\
 \dot{x}_5 &= -x_5 - 3x_1 x_4
 \end{aligned} \tag{3.124}$$

The FT model is a five-mode truncation of the Navier-Stokes equations for a two-dimensional incompressible fluid on a torus. By varying the parameter r , different qualitative asymptotic behaviors can be observed: from limit-cycle oscillations to chaotic attractors obtained from a sequence of period-doubling bifurcations: for $r = 33.6$ we observe a limit cycle (Fig. 3.16 (a)) and for $r = 30$ a chaotic attractor (Fig. 3.16 (b)). The table below is a summary of the Liapunov exponents Λ_h

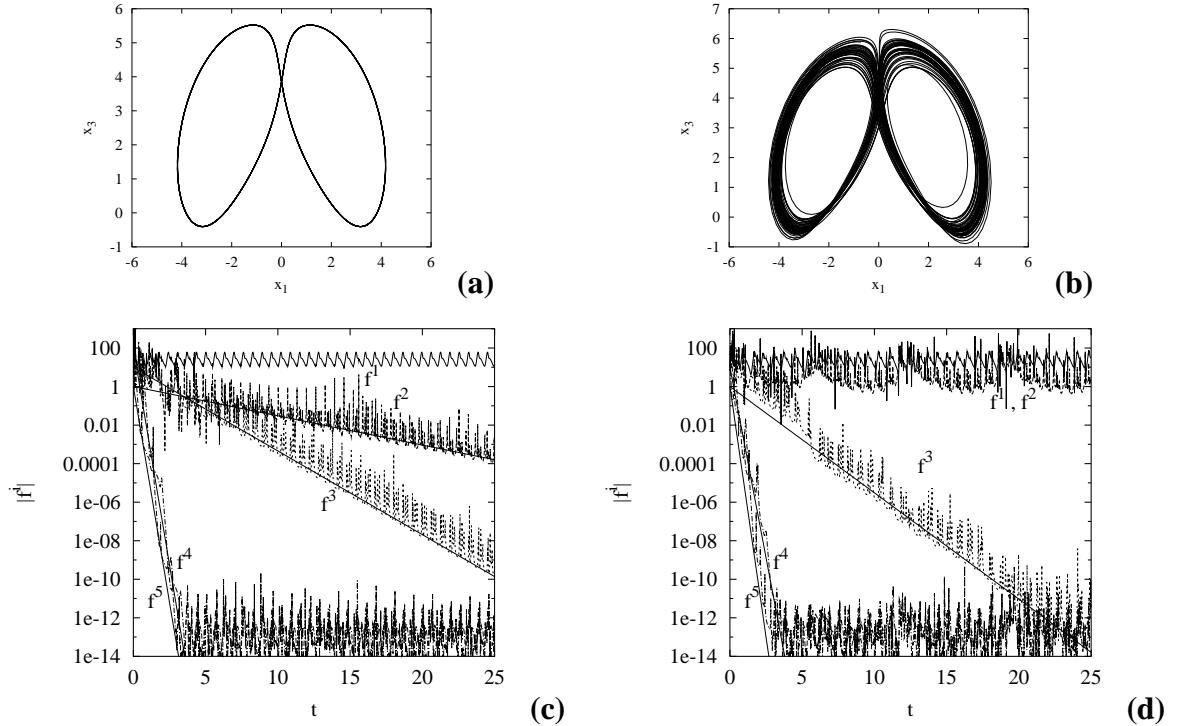


Fig. 3.16: Franceschini-Tebaldi system. (a) and (b) projections of the limit set on the (x_1, x_3) plane for the periodic and chaotic cases. (c) and (d) corresponding projections $f^h(\mathbf{x})$.

calculated with the exterior algebra technique for the limit cycle and chaotic attractor cases. Also shown for comparison the limit values of the vector norm rates $\langle \mu_h \rangle$ as calculated with NTDRB. Figs. 3.16 (c)(d) show the projections $f^h(t) = \mathbf{c}^h(t)\mathbf{g}(t)$ as calculated with NTDRB. The continuous lines represent the scalings $\sim e^{\Lambda_h}$ for the corresponding negative Liapunov exponents thus showing that the projections do scale like them. Note that projections relative to positive or zero Liapunov exponents do not increase nor decrease. The analysis of the projections $f^h(t)$ constitutes yet another technique to estimate the negative part of the Liapunov spectrum.

3.11 Model reduction

The natural ordering of NTDRB vectors and covectors in a decreasing Liapunov exponent order induces a natural ordering of the system's source term projections or modes as seen in section 3.9.

$(r = 33.6) \Lambda_h$	0	-0.35	-1	-10.3	-10.35
$(r = 33.6) \langle \mu_h \rangle$	0	-0.06	-1.56	-9.84	-10.34
$(r = 30.0) \Lambda_h$	0.34	0	-1.27	-9.1	-11.97
$(r = 30.0) \langle \mu_h \rangle$	0.34	0.26	-1.57	-8.69	-11.96

 Table 3.1: Liapunov exponents Λ_h calculated with the exterior algebra technique for the FT system.

This natural ordering of modes (which makes any sorting procedure useless) makes it especially easy to perform a reduction of the original system. We may think of monitoring the quantities $\langle (J\mathbf{a}_i, \mathbf{a}_i) \rangle(t)$ for a sufficiently large time, at which point they will be a reasonable estimate of the Liapunov spectrum, until the following invariant decomposition of $TC_{\mathbf{x}}$ is established: $TC_{\mathbf{x}} = \overline{E}^{h^*} \oplus E^s$ where

$$\overline{E}^{h^*} = E^1 \oplus \cdots \oplus E^{h^*} = \text{span}\{\mathbf{a}_1, \dots, \mathbf{a}_{h^*}\} \quad (3.125)$$

corresponds to the direct sum of all the unstable (slow) invariant subspaces, relative to all the positive Liapunov exponents, and the central invariant subspace, corresponding to the zero exponent, (and \overline{E}^{h^*} is the unstable/central invariant filtration), and

$$E^s = E^{h^*+1} \oplus \cdots \oplus E^n = \text{span}\{\mathbf{a}_{h^*+1}, \dots, \mathbf{a}_n\} = \overline{E}^n \setminus \overline{E}^{h^*} = TC \setminus \overline{E}^{h^*} \quad (3.126)$$

corresponds to all the stable (fast) invariant subspaces and h^* is the number of non-negative Liapunov exponents¹⁷. When for all the stable modes we can apply $f^i(\mathbf{x}(t)) = \mathbf{c}^i(t)\mathbf{g}(t) \approx 0$ with $i = h^* + 1, \dots, n$ (and this will happen within the slowest of the stable intrinsic timescales, i.e. $1/|\Lambda_{h^*+1}|$) then the reduced system will be

$$\dot{\mathbf{x}} = \sum_{i=1}^{h^*} \mathbf{a}_i(t) f^i(t) \quad (3.127)$$

which will be defined in the set $\{\mathbf{x} \in C | \mathbf{g}(\mathbf{x}) \in \overline{E}^{h^*}(\mathbf{x})\} \subseteq C$ or $\{\mathbf{x} \in C | f^i(\mathbf{x}(t)) = 0, \quad h = 1, h^*\} \subseteq C$. The same reasoning may be applied to singularly perturbed systems possessing a trivial limit set (equilibrium point). In this case no positive or zero Liapunov exponents are present, but we may set a threshold value separating the set of slow stable timescales from the set of fast stable timescales. Clearly $\mathcal{M} = \{\mathbf{x} \in C | \mathbf{g}(\mathbf{x}) \in \overline{E}^{h^*}(\mathbf{x})\} \subseteq C$ in this case defines a slow manifold.

As we integrate, as soon as the most stable projection, say $f^{M(t)}$, is sufficiently small, with $M(t) \in$

¹⁷The number of non-negative Liapunov exponents is sometimes used in the literature to estimate the Liapunov dimension $d_L = h^* + \frac{\sum_{k=1}^{h^*} \Lambda_k}{|\Lambda_{h^*+1}|}$. The asymptotic behavior of the system takes place within the so called global attractor C whose embedding dimension can be estimated by d_L .

$[h^*+1, \dots, n]$, there is no need to keep integrating the dynamics of the corresponding (and more stable yet) vectors/covectors $\{\mathbf{a}_i\}_{i=M}^n, \{\mathbf{c}^i\}_{i=M}^n$. This is because the dynamics of each vector/covector is uncoupled from the dynamics of the remaining vectors/covectors. At the same time the system's source term may be updated so that $\dot{\mathbf{x}} = \sum_{i=1}^{M-1} \mathbf{a}_i(t) f^i(t)$.

3.12 Applications of NTDRB to reactive systems.

This section is devoted to the application of the NTDRB technique to spatially homogeneous reactive systems. As a test, the homogeneous isobaric H_2 combustion in air was chosen, using a simplified 8 species ($\text{H}_2, \text{O}_2, \text{O}, \text{OH}, \text{H}_2\text{O}, \text{H}, \text{HO}_2, \text{N}_2$ [dilutant]), 6 reactions mechanism, displayed in chemkin format on Table 3.2. The mechanism was obtained using the CSP-based simplification algorithm described in chapter 2 section 2.5, using as input CSP data from the stoichiometric combustion of hydrogen in air at $p = 1\text{Atm}$ and an initial temperature $T = 1000K$ governed by a detailed mechanism comprised of 9 species and 19 reactions.

Reaction	A	β	E (cal/mole)
$\text{H}+\text{O}_2=\text{O}+\text{OH}$	1.915×10^{14}	0.00	1.644×10^4
$\text{O}+\text{H}_2=\text{H}+\text{OH}$	5.080×10^4	2.67	6.290×10^3
$\text{H}_2+\text{OH}=\text{H}_2\text{O}+\text{H}$	2.160×10^8	1.51	3.430×10^3
$\text{H}+\text{O}_2+\text{M}=\text{HO}_2+\text{M}$	6.170×10^{19}	-1.42	0.0
$\text{HO}_2+\text{H}=\text{H}_2+\text{O}_2$	6.630×10^{13}	0.0	2.130×10^3
$\text{HO}_2+\text{OH}=\text{H}+\text{H}_2\text{O}$	1.690×10^{14}	0.0	8.740×10^2

Table 3.2: Chemkin format H_2 combustion simplified mechanism. For each reaction, A is the pre-exponential factor, β the temperature exponent and E the activation energy in the Arrhenius forward rate constant $k = AT^\beta e^{-E/(RT)}$.

The NTDRB technique was used on such a reactive system as described in section 3.8, solving the skew-product system (3.99)-(3.101) and performing the biorthonormalization procedure. Note that Eq.(3.99) is composed of the species conservation equations and the energy equation as described in chapter 2 section 2.2. The 9×9 Jacobian matrix of the chemical source term was determined analytically to improve the overall accuracy of the implementation. Note that for mechanisms involving a greater number of species, the analytical Jacobian can be too cumbersome to be realistically used. The biorthonormal system of vectors $\{\{\tilde{\mathbf{a}}_i(t)\}, \{\tilde{\mathbf{c}}^i(t)\}\}$, available at every biorthonormalization step, is used as a projection basis for the chemical source term $\mathbf{g}(\mathbf{x})$:

$$\mathbf{g}(\mathbf{x}(t)) = \sum_i \tilde{\mathbf{a}}_i(t) f^i(t) ; \quad f^i(t) = \tilde{\mathbf{c}}^i(t) \mathbf{g}(\mathbf{x}(t)). \quad (3.128)$$

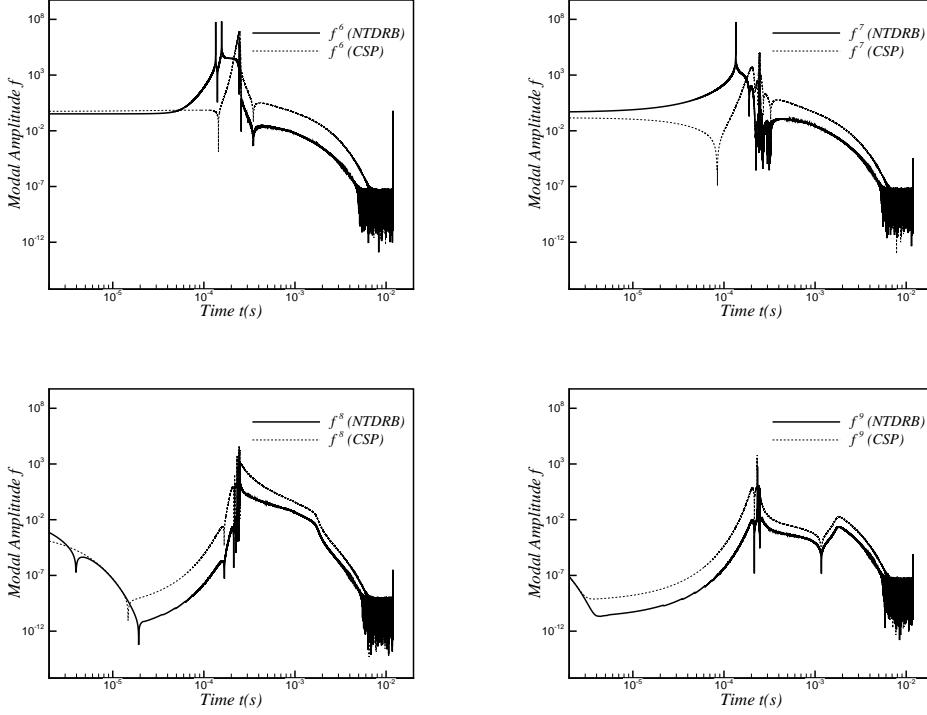


Fig. 3.17: Modal amplitudes $f^6(t)$ through $f^9(t)$ calculated with NTDRB vectors and CSP vectors (eigen-vectors).

The reactive system's trajectory is attracted to an asymptotically stable equilibrium point after the combustion process has taken place. This implies the presence of a set of lower dimensional invariant structures or manifolds, locally tangent at the equilibrium point to the eigenpaces of the Jacobian, and related to the spectrum of intrinsic timescales of the system, which are defined as the phase space loci where the system evolves according to an individual intrinsic timescale. In particular, in the presence of M stable and fast intrinsic timescales, a generic trajectory will be attracted and constrained to the $(N - M)$ -dimensional surface defined by the M relations $f^i(\mathbf{x}) = \tilde{\mathbf{c}}^i(\mathbf{x})\mathbf{g}(\mathbf{x}) = 0$ for $i = (N - M), \dots, N$. The NTDRB modal amplitudes $f^i(t)$ were monitored along a system's trajectory and compared with the CSP modal amplitudes resulting from a projection onto the local eigenvectors of the Jacobian matrix. Because of the lack of invariance of the CSP basis and the resulting coupling between the modal amplitudes, one can see that the CSP amplitudes fail to saturate to zero according to individual timescales, and are constantly greater than the NTDRB amplitudes, as can be seen from Fig.3.17.

Fig.3.18 displays one of the instantaneous timescales, τ_8 , as calculated with NTDRB and CSP. In the NTDRB case, such timescales were taken as the numbers $\tau_i = 1/|(J(t)\mathbf{a}_i(t), \mathbf{a}_i(t))|$, whereas in

the CSP case simply as the reciprocal of the absolute values of the real part of the local eigenvalues. It is instructive to notice that the two timescales converge towards the equilibrium point. This is trivially a consequence of the fact that, at the equilibrium point, the invariant subspaces coincide with the eigenvectors.

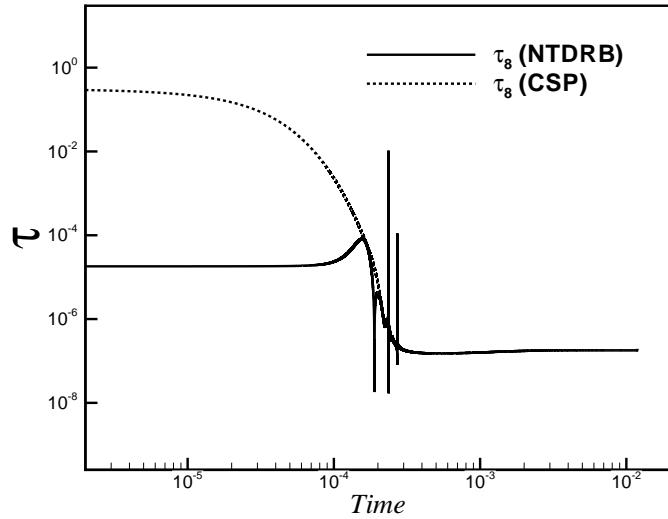


Fig. 3.18: NTDRB and CSP timescale $\tau_8(t)$.

Conclusions

The Computational Singular Perturbation method for chemical kinetic systems reduction was revised in the context of Fenichel’s Geometric Singular Perturbation Theory. The ILDM and Fraser-Roussel methods were also revised and compared to CSP. The CSP basis vector refinement procedure was described by introducing a block power method for the block-diagonalization of a square matrix. The general implementation of the refinement procedure was then discussed for higher dimensional systems. It was shown that for a singularly perturbed system, with ε as the small parameter measuring time scale separation, the CSP method generates term by term the asymptotic expansion of the system’s slow invariant manifold \mathcal{M}_ε up to and including $\mathcal{O}(\varepsilon^q)$ terms, where q is the number of CSP refinements performed, consistently with the findings of Zagaris et. al [19].

A CSP-based algorithm for the simplification of chemical kinetics mechanisms was also introduced. Starting from a solution of the detailed kinetics, the CSP analysis generates a CSP database comprising importance indices and pointers to CSP radicals. The simplification algorithm analyzes the CSP database and, by discarding reactions and species deemed unimportant, generates simplified (skeletal) mechanisms on the basis of (*i*) a tolerance threshold on importance indices and (*ii*) a set of ‘kernel’ species whose production/depletion pathways are to be reconstructed. Dimensionality reduction is achieved when species are completely eliminated from a simplified mechanism. The algorithm was extensively tested on the Gri Mech.3.0 detailed mechanism used for the autoignition of methane in air. The diagnostic features of the CSP databases was also illustrated by analyzing the different pathways concurring to NO formation.

The extension of CSP to convective-diffusive-reactive systems and the limits of applicability of such an extension were also addressed. The simplification algorithm was then tested on methane/air premixed laminar flames. Several comprehensive simplified mechanisms were obtained, each stemming from the concurrent analysis of several CSP databases. Comprehensiveness was confirmed by the validity of the simplified mechanisms persisting over a range of equivalence ratios.

Both for the autoignition and premixed flame problems, the simplified mechanisms achieved a maximum CPU speed-up factor of 3 before accuracy is compromised beyond acceptable levels.

Finally the Natural Tangent Dynamics with Recurrent Biorthogonalization (NTDRB) technique was introduced. NTDRB unifies into a single technique the analysis of dynamical systems possessing either a slow manifold (associated with an equilibrium point), for which the main focus relies in the identification of the slow/fast timescales, or systems exhibiting non trivial asymptotic oscillatory or chaotic behaviors, for which the central issue is the temporal dichotomy between unstable/stable components of the dynamics described by the Lyapunov spectrum associated to Oseledec's invariant subspaces. The linear non-autonomous variational equation for natural vector/covector evolution (induced by the system) in the tangent/cotangent bundles is used, together with a recurrent biorthonormalization procedure, to estimate the system's Lyapunov spectrum and the associated invariant subspaces. The Lyapunov spectrum is introduced as the intrinsic timescale spectrum of non linear systems as opposed to the local eigenvalue spectrum. The inadequacy of local eigenvalues as timescales for the class of systems possessing non trivial asymptotic dynamics is addressed. NTDRB is applied to systems exhibiting both periodic or chaotic behaviors for which CSP or ILDM fail. A practical application of NTDRB is also given for a simplified H₂/O₂ system, comparing it with corresponding CSP results.

Appendix A

The block-power method

Given the constant $n \times n$ matrix C and its right and left eigenvector matrices V and W , then

$$W C V = \text{diag}(\mu_i) \quad (\text{A.1})$$

where $\mu_i, i = 1, n$ are the eigenvalues of C . Suppose the eigenvalues are ordered as

$$|\mu_1| > \cdots > |\mu_M| >> |\mu_{M+1}| > \cdots > |\mu_n| \quad (\text{A.2})$$

and let us define the small number $\varepsilon = |\mu_{M+1}|/|\mu_M| \ll 1$. The ordering of eigenvalues establishes an ordering of eigenvectors which determines a decomposition of R^n into a subspace spanned by the eigenvectors with largest eigenvalues and a subspace spanned by the eigenvectors with the smallest eigenvalues, so that:

$$V = (V_r, V_s) \quad ; \quad W = (W^r, W^s)^T \quad (\text{A.3})$$

where V_r is the $n \times M$ matrix of column eigenvectors $\{\mathbf{v}_r\}_{r=1}^M$, V_s is the $n \times (n - M)$ matrix of column eigenvectors $\{\mathbf{v}_s\}_{s=M+1}^n$, W^r is the $M \times n$ matrix of row eigenvectors $\{\mathbf{w}^r\}_{r=1}^M$ and W^s is the $(n - M) \times n$ matrix of row eigenvectors $\{\mathbf{w}^s\}_{s=M+1}^n$, and where the orthonormality condition $\mathbf{w}^i \mathbf{v}_j = \delta_j^i$ holds.

Any set of n column vectors $\{\mathbf{a}_i\}_{i=1}^n$ or row vectors $\{\mathbf{b}^i\}_{i=1}^n$ may be represented in the left or right

eigenbasis as

$$(A_r, A_s) = (V_r, V_s) \begin{pmatrix} Q_r^r & Q_s^r \\ Q_r^s & Q_s^s \end{pmatrix} \quad (\text{A.4})$$

$$\begin{pmatrix} B^r \\ B^s \end{pmatrix} = \begin{pmatrix} P_r^r & P_s^r \\ P_r^s & P_s^s \end{pmatrix} \begin{pmatrix} W^r \\ W^s \end{pmatrix} \quad (\text{A.5})$$

where $A_r = (\mathbf{a}_1, \dots, \mathbf{a}_M)$, $A_s = (\mathbf{a}_{M+1}, \dots, \mathbf{a}_n)$, $B^r = (\mathbf{b}^1, \dots, \mathbf{b}^M)^T$ and $B^s = (\mathbf{b}^{M+1}, \dots, \mathbf{b}^n)^T$.

In contracted form the above relations read $A = VQ$ and $B = PW$.

The general idea behind the block-power method is to consider an iterative procedure of the kind:

$$A_r(m+1) = CA_r(m) = C^m A_r(0) \quad (\text{A.6})$$

$$B^r(k+1) = B^r(k)C = B^r(0)C^k \quad (\text{A.7})$$

where C^m and C^k denotes exponentiation. This is re-written as

$$A_r(m+1) = (\mu_r^r)^m V_r Q_r^r + (\mu_s^s)^m V_s Q_s^s = \quad (\text{A.8})$$

$$= \|(\mu_r^r)^m\| \left(\frac{(\mu_r^r)^m}{\|(\mu_r^r)^m\|} V_r Q_r^r + \mathcal{O}(\varepsilon^m) \right); \quad (\text{A.9})$$

$$B^r(k+1) = P_r^r (\mu_r^r)^k W^r + P_s^r (\mu_s^s)^k W^s = \quad (\text{A.10})$$

$$= \|(\mu_r^r)^k\| \left(P_r^r \frac{(\mu_r^r)^k}{\|(\mu_r^r)^k\|} W^r + \mathcal{O}(\varepsilon^k) \right). \quad (\text{A.11})$$

Thus for a sufficient number of iterations the system $\{\mathbf{a}_r\}_{r=1}^M$ will lie in the subspace spanned by $\{\mathbf{v}_r\}_{r=1}^M$ and $\{\mathbf{b}^r\}_{r=1}^M$ will lie in the subspace spanned by $\{\mathbf{w}^r\}_{r=1}^M$.

For $A(\ell)$ and $B(\ell)$ to represent an orthonormal system after any number ℓ of iterations, matrices $Q(\ell)$ and $P(\ell)$ should be such that $P(\ell)Q(\ell) = Q(\ell)P(\ell) = I$. If this is the case then we can define a matrix Λ similar to C as

$$\Lambda(\ell) = B(\ell)CA(\ell) = \begin{pmatrix} \Lambda_r^r(\ell) & \Lambda_s^r(\ell) \\ \Lambda_r^s(\ell) & \Lambda_s^s(\ell) \end{pmatrix} \quad (\text{A.12})$$

where we assumed C constant.

The iterative block-diagonalization procedure aims at reducing $\Lambda(\ell)$ in block-diagonal form for $\ell \rightarrow \infty$. The procedure is divided into two distinct steps, a so called A_r -refinement which alters A_r (and

B^s to preserve orthonormality) leaving A_s and B^r unaltered. Such a step puts Λ in a upper block-triangular form by diminishing the norm of submatrix Λ_r^s . The second step is referred to as the B^r -refinement altering B^r (and A_s to preserve orthonormality). This diminishes the norm of submatrix Λ_s^r putting Λ in lower block-triangular form. The two steps may be performed independently as each preserves biorthogonality.

The A_r -refinement.

The first step is controlled by the two operators

$$\Gamma(m) = \begin{pmatrix} I_r^r & 0_s^r \\ \Lambda_r^s(m)(\Lambda_r^r(m))^{-1} & I_s^s \end{pmatrix} ; \quad \Delta(m) = \begin{pmatrix} I_r^r & 0_s^r \\ -\Lambda_r^s(m)(\Lambda_r^r(m))^{-1} & I_s^s \end{pmatrix}, \quad (\text{A.13})$$

where m is the iteration index for this step. These operators are equivalent to $(I + L)$ and $(I - L)$ seen in chapt. 1. Thus the A_r -refinement is defined as:

$$A(m+1) = A(m)\Gamma(m) \quad (\text{A.14})$$

$$B(m+1) = \Delta(m)B(m). \quad (\text{A.15})$$

Noting that $\Gamma(m)\Delta(m) = \Delta(m)\Gamma(m) = I$, then the A_r -refinement preserves biorthogonality as $A(m+1)B(m+1) = A(m)\Gamma(m)\Delta(m)B(m) = A(m)B(m)$. Note also that B^r and A_s are unchanged while A_r and B^s are subject to:

$$A_r(m+1) = A_r(m) + A_s(m)\Lambda_r^s(m)(\Lambda_r^r(m))^{-1} = CA_r(m)(\Lambda_r^r(m))^{-1} \quad (\text{A.16})$$

$$\begin{aligned} B^s(m+1) &= B^s(m) - \Lambda_r^s(m)(\Lambda_r^r(m))^{-1}B^r(m) = \\ &= B^s(m)(I_r^r - A_r(m+1)B^r(m+1)) \end{aligned} \quad (\text{A.17})$$

Eq.A.16 is a result of Eq.A.12 which is equivalent to $CA = A\Lambda(m)$. From the latter we have $CA_r = A_r\Lambda_r^r + A_s\Lambda_r^s$ which reduces to Eq.A.16. Eq.A.17 is obtained by direct substitution of Eq.A.16. It is also immediate to verify that matrix Λ is transformed as

$$\begin{aligned} \Lambda(m+1) &= \Delta(m)\Lambda(m)\Gamma(m) = \\ &= \begin{pmatrix} \Lambda_r^r(m) + \Lambda_r^r(m)\Lambda_r^s(m)(\Lambda_r^r(m))^{-1} & \Lambda_r^r(m) \\ \Lambda_s^s(m)\Lambda_r^s(m)(\Lambda_r^r(m))^{-1} & \Lambda_s^s(m) - \Lambda_r^s(m)(\Lambda_r^r(m))^{-1}\Lambda_s^r(m) \end{pmatrix}. \end{aligned} \quad (\text{A.18})$$

Therefore the A_r -refinement does not affect the upper right block of matrix Λ . Note also that $\Lambda(m+1)$ is similar to $\Lambda(m)$ because $\Delta(m)\Gamma(m) = \Gamma(m)\Delta(m) = I$.

The newly refined bases $A(m+1)$ and $B(m+1)$ can be represented in the right and left eigenvector bases V and W :

$$A(m+1) = A(m)\Gamma(m) = VQ(m)\Gamma(m) = VQ(m+1) \quad (\text{A.19})$$

$$B(m+1) = \Delta(m)B(m) = \Delta(m)P(m)W = P(m+1)W \quad (\text{A.20})$$

where

$$Q(m+1) = Q(m)\Gamma(m) \quad (\text{A.21})$$

$$P(m+1) = \Delta(m)P(m). \quad (\text{A.22})$$

It readily follows from the definition Eq.A.13 of Γ and Δ that the upper right and lower right blocks of Q are unchanged. Similarly the upper left and upper right blocks of P are left unchanged. On the other hand using Eq.A.19 and Eq.A.16 we have

$$A_r(m+1) = V_r Q_r^r(m+1) + V_s Q_r^s(m+1) = \quad (\text{A.23})$$

$$= CA_r(m)(\Lambda_r^r(m))^{-1} = \quad (\text{A.24})$$

$$= C(V_r Q_r^r(m) + V_s Q_r^s(m))(\Lambda_r^r(m))^{-1} = \quad (\text{A.25})$$

$$= \mu_r^r V_r Q_r^r(m)(\Lambda_r^r(m))^{-1} + \mu_s^s V_s Q_r^s(m)(\Lambda_r^r(m))^{-1} \quad (\text{A.26})$$

from which we conclude that

$$Q_r^r(m+1) = \mu_r^r Q_r^r(m)(\Lambda_r^r(m))^{-1} \quad (\text{A.27})$$

$$Q_r^s(m+1) = \mu_s^s Q_r^s(m)(\Lambda_r^r(m))^{-1} = \|\mu_r^r\| \left(\frac{\mu_s^s}{\|\mu_r^r\|} Q_r^s(m)(\Lambda_r^r(m))^{-1} \right) \quad (\text{A.28})$$

Let us now establish the order of magnitude of $\|Q_r^r(m+1)\|$ and $\|Q_r^s(m+1)\|$. From the definition of Λ :

$$\begin{aligned}
\Lambda_r^r(m) &= B^r C A_r = \\
&= (P^r(m)W^r + P_s^r(m)W^s)C(V_r Q_r^r(m) + V_s Q_r^s(m)) = \\
&= P_r^r(m)\mu_r^r Q_r^r(m) + P_s^r(m)\mu_s^s Q_r^s(m).
\end{aligned} \tag{A.29}$$

From the above relation it can be shown that

$$\|(\Lambda_r^r(m))^{-1}\| \leq \frac{1}{\|\mu_r^r\|} \frac{1}{\|P_r^r(m)\| \|Q_r^r(m)\|} (1 + \mathcal{O}(\varepsilon)). \tag{A.30}$$

From Eqs.A.27, A.28 it follows that

$$\|Q_r^r(m+1)\| = \mathcal{O}\left(\|Q_r^r(m)\| \frac{1}{\|P_r^r(m)\| \|Q_r^r(m)\|} (1 + \mathcal{O}(\varepsilon))\right) \tag{A.31}$$

$$\|Q_r^s(m+1)\| = \mathcal{O}\left(\varepsilon \|Q_r^s(m)\| \frac{1}{\|P_r^r(m)\| \|Q_r^r(m)\|} (1 + \mathcal{O}(\varepsilon))\right). \tag{A.32}$$

We therefore conclude that if $\|Q_r^r(0)\| = \mathcal{O}(1)$, $\|Q_r^s(0)\| = \mathcal{O}(1)$ and $\|P_r^r(0)\| = \mathcal{O}(1)$, then being $P_r^r(m) = P_r^r(0)$, we have

$$\|Q_r^r(m)\| = \mathcal{O}(1) \tag{A.33}$$

$$\|Q_r^s(m)\| = \mathcal{O}(\varepsilon^m). \tag{A.34}$$

In fact if for $m = 0$ all submatrices of Q and P are $\mathcal{O}(1)$ it is straight forward to show from $P(m)Q(m) = I$ that

$$\|P_r^s(m)\| = \mathcal{O}(\varepsilon^m) \tag{A.35}$$

$$\|P_s^s(m)\| = \mathcal{O}(1). \tag{A.36}$$

At every iteration step we can therefore formally express A and B as

$$(A_r(m), A_s(m)) = (V_r, V_s) \begin{pmatrix} Q_r^r(m) & Q_r^s(m) \\ \varepsilon^m Q_r^s(m) & Q_s^s(m) \end{pmatrix} \tag{A.37}$$

$$\begin{pmatrix} B^r(m) \\ B^s(m) \end{pmatrix} = \begin{pmatrix} P_r^r(m) & P_s^r(m) \\ \varepsilon^m P_r^s(m) & P_s^s(m) \end{pmatrix} \begin{pmatrix} W^r \\ W^s \end{pmatrix} \tag{A.38}$$

where all the submatrices of $Q_{r,s}^{r,s}(m)$ and $P_{r,s}^{r,s}(m)$ are $\mathcal{O}(1)$. From the above we conclude that for a sufficient number of iterations, column vectors $\{\mathbf{a}_i(m)\}_{i=1}^M$ contained in $A_r(m)$ will tend to span the subspace represented by V_r , i.e. the subspace spanned by $\{\mathbf{v}_i\}_{i=1}^M$. At the same time row vectors $\{\mathbf{b}^i(m)\}_{i=M+1}^n$ contained in $B^s(m)$ will tend to span the subspace represented by W^s , i.e. the subspace spanned by $\{\mathbf{w}^i\}_{i=M+1}^n$.

As far as matrix $\Lambda(m)$ is concerned we have

$$\Lambda(m) = B(m)CA(m) = P(m)WCVQ(m) = P(m) \begin{pmatrix} \mu_r^r & 0_s^r \\ 0_r^s & \mu_s^s \end{pmatrix} Q(m) = \quad (\text{A.39})$$

$$= \begin{pmatrix} P_r^r(m)\mu_r^rQ_r^r(m) + \mathcal{O}(\varepsilon^m) & P_r^r(m)\mu_r^rQ_s^r(m) + \mathcal{O}(\varepsilon^m) \\ \varepsilon^m(P_r^s(m)\mu_r^rQ_r^r(m) + P_s^s(m)\mu_s^sQ_r^s(m)) & P_s^s(m)\mu_s^sQ_s^s(m) + \mathcal{O}(\varepsilon^m) \end{pmatrix} \quad (\text{A.40})$$

which clearly shows that after a sufficient number of refinements Λ reduces to upper-block triangular form. It can also be shown that the eigenvalues of the diagonal block $\Lambda_r^r(m)$ coincide with the M largest eigenvalues of C , μ_r^r , to $\mathcal{O}(\varepsilon^m)$. Similarly the eigenvalues of $\Lambda_s^s(m)$ coincide with the $n - M$ smallest eigenvalues of C , μ_s^s , to $\mathcal{O}(\varepsilon^m)$.

The B^r -refinement.

Similar considerations apply to the B^r -refinement which constitutes the second independent step of a complete refinement cycle. For completeness we will write the operators governing this step:

$$\Theta(k) = \begin{pmatrix} I_r^r & -(\Lambda_r^r(k))^{-1}\Lambda_s^r(k) \\ 0_r^s & I_s^s \end{pmatrix} \quad ; \quad \Pi(k) = \begin{pmatrix} I_r^r & (\Lambda_r^r(k))^{-1}\Lambda_s^r(k) \\ 0_r^s & I_s^s \end{pmatrix}, \quad (\text{A.41})$$

where k is the iteration index. These operators are equivalent to $(I - U)$ and $(I + U)$ seen in chapt. 1. Thus the B^r -refinement is defined as:

$$A(k+1) = A(k)\Theta(k) \quad (\text{A.42})$$

$$B(k+1) = \Pi(k)B(k). \quad (\text{A.43})$$

This implies that B^r and A_s are modified while A_r and B^s are left unchanged:

$$B^r(k+1) = B^r(k) + (\Lambda_r^r(k))^{-1} \Lambda_s^r(k) B^s(k) = (\Lambda_r^r(k))^{-1} B^r(k) C \quad (\text{A.44})$$

$$\begin{aligned} A_s(k+1) &= A_s(k) - A_r(k)(\Lambda_r^r(k))^{-1} \Lambda_s^r(k) = \\ &= (I_r^r - A_r(k+1)B^r(k+1)) A_s(k). \end{aligned} \quad (\text{A.45})$$

Appendix B

CSP Refinements: Evaluation of derivatives

The basis vector time derivatives are obtained by means of chain-rule differentiation of the relations illustrated in Chap.1 Sec.1.8.2. In the following we summarize the main results. To lighten the notation we substituted $(\Lambda_r^r)^{-1} = \tau_r^r$ throughout this appendix.

Given that:

$$\begin{aligned} \frac{d\Lambda_r^r(0,0)}{dt} &= \frac{d}{dt} \left[\left(\frac{dB^r(0,0)}{dt} + B^r(0,0)J \right) A_r(0,0) \right] &= \frac{d}{dt} [B^r(0,0)JA_r(0,0)] = \\ &= B^r(0,0) \frac{dJ}{dt} A_r(0,0) \end{aligned} \quad (\text{B.1})$$

$$\begin{aligned} \frac{d\tau_r^r(0,0)}{dt} &= -\tau_r^r(0,0) \frac{d\Lambda_r^r(0,0)}{dt} \tau_r^r(0,0) = \\ &= -\tau_r^r(0,0) B^r(0,0) \frac{dJ}{dt} A_r(0,0) \tau_r^r(0,0) \end{aligned} \quad (\text{B.2})$$

then the time derivative of $B^r(1,0)$ and $B^r(1,1)$ can be evaluated as:

$$\begin{aligned} \frac{dB^r(1,1)}{dt} &= \frac{dB^r(1,0)}{dt} = \\ &= \frac{d\tau_r^r(0,0)}{dt} B^r(0,0) J + \tau_r^r(0,0) B^r(0,0) \frac{dJ}{dt} \\ &= -\tau_r^r(0,0) B^r(0,0) \frac{dJ}{dt} A_r(0,0) \tau_r^r(0,0) B^r(0,0) J + \tau_r^r(0,0) B^r(0,0) \frac{dJ}{dt} \\ &= -\tau_r^r(0,0) B^r(0,0) \frac{dJ}{dt} A_r(0,0) B^r(1,0) + \tau_r^r(0,0) B^r(0,0) \frac{dJ}{dt} \\ &= \tau_r^r(0,0) B^r(0,0) \frac{dJ}{dt} [I - A_r(0,0) B^r(1,0)] \end{aligned} \quad (\text{B.3})$$

Moreover, given that:

$$\begin{aligned}
 \frac{d\Lambda_r^r(1,0)}{dt} &= \frac{d}{dt} [B^r(1,0)JA_r(0,0)] = \frac{dB^r(1,0)}{dt}JA_r(0,0) + B^r(1,0)\frac{dJ}{dt}A_r(0,0) \\
 \frac{d\tau_r^r(1,0)}{dt} &= -\tau_r^r(1,0)\frac{d\Lambda_r^r(1,0)}{dt}\tau_r^r(1,0) \\
 &= -\tau_r^r(1,0)\frac{dB^r(1,0)}{dt}JA_r(0,0)\tau_r^r(1,0) + \\
 &\quad -\tau_r^r(1,0)B^r(1,0)\frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) \\
 &= -\tau_r^r(1,0)\frac{dB^r(1,0)}{dt}A_r(1,1) - \tau_r^r(1,0)B^r(1,0)\frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) \quad (\text{B.4})
 \end{aligned}$$

then the time derivative of $A_r(1,1)$ and $A_r(2,1)$ can be evaluated as:

$$\begin{aligned}
 \frac{dA_r(1,1)}{dt} &= \frac{dA_r(2,1)}{dt} = \\
 &= \frac{d}{dt} [JA_r(0,0)\tau_r^r(1,0)] = \frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) + \\
 &\quad + J\frac{dA_r(0,0)}{dt}\tau_r^r(1,0) + JA_r(0,0)\frac{d\tau_r^r(1,0)}{dt} \\
 &= \frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) - JA_r(0,0)\tau_r^r(1,0)\frac{dB^r(1,0)}{dt}A_r(1,1) + \\
 &\quad - JA_r(0,0)\tau_r^r(1,0)B^r(1,0)\frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) \\
 &= \frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) - A_r(1,1)\frac{dB^r(1,0)}{dt}A_r(1,1) + \\
 &\quad - A_r(1,0)B^r(1,0)\frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) \\
 &= [I - A_r(1,1)B^r(1,0)] \frac{dJ}{dt}A_r(0,0)\tau_r^r(1,0) - A_r(1,1)\frac{dB^r(1,0)}{dt}A_r(1,1)
 \end{aligned} \quad (\text{B.5})$$

It follows that the time derivative of $B^r(2, 1)$ can be evaluated as:

$$\begin{aligned}
\frac{dB^r(2, 1)}{dt} &= \frac{d}{dt} \left[\tau_r^r(1, 1) \left(\frac{dB^r(1, 0)}{dt} + B^r(1, 0)J \right) \right] \\
&= \frac{d\tau_r^r(1, 1)}{dt} \frac{dB^r(1, 0)}{dt} + \tau_r^r(1, 1) \frac{d^2B^r(1, 0)}{dt^2} + \\
&\quad + \frac{d\tau_r^r(1, 1)}{dt} B^r(1, 1)J + \tau_r^r(1, 1) \frac{dB^r(1, 0)}{dt} J + \tau_r^r(1, 1)B^r(1, 0) \frac{dJ}{dt} \\
&= \frac{d\tau_r^r(1, 1)}{dt} \left[\frac{dB^r(1, 0)}{dt} + B^r(1, 1)J \right] + \tau_r^r(1, 1) \frac{d^2B^r(1, 0)}{dt^2} + \\
&\quad + \tau_r^r(1, 1) \frac{dB^r(1, 0)}{dt} J + \tau_r^r(1, 1)B^r(1, 1) \frac{dJ}{dt} \\
&= \frac{d\tau_r^r(1, 1)}{dt} \Lambda_r^r(1, 1)B^r(2, 1) + \tau_r^r(1, 1) \frac{d^2B^r(1, 0)}{dt^2} + \\
&\quad + \tau_r^r(1, 1) \frac{dB^r(1, 0)}{dt} J + \tau_r^r(1, 1)B^r(1, 1) \frac{dJ}{dt}
\end{aligned} \tag{B.6}$$

where

$$\begin{aligned}
\frac{d\tau_r^r(1, 1)}{dt} &= -\tau_r^r(1, 1) \frac{d\Lambda_r^r(1, 1)}{dt} \tau_r^r(1, 1) \\
&= -\tau_r^r(1, 1) \frac{d}{dt} \left[\left(\frac{dB^r(1, 0)}{dt} + B^r(1, 0)J \right) A_r(1, 1) \right] \tau_r^r(1, 1) \\
&= -\tau_r^r(1, 1) \left[\frac{d^2B^r(1, 0)}{dt^2} A_r(1, 1) + \frac{dB^r(1, 0)}{dt} \frac{dA_r(1, 1)}{dt} + \right. \\
&\quad \left. + B^r(1, 0) \frac{dJ}{dt} A_r(1, 1) + B^r(1, 0)J \frac{dA_r(1, 1)}{dt} + \frac{dB^r(1, 0)}{dt} J A_r(1, 1) \right] \tau_r^r(1, 1)
\end{aligned} \tag{B.7}$$

By replacing Eq. (B.6) in Eq. (B.7) yields:

$$\begin{aligned}
\frac{dB^r(2,1)}{dt} = & -\tau_r^r(1,1) \frac{d^2B^r(1,0)}{dt^2} A_r(1,1) B^r(2,1) \\
& -\tau_r^r(1,1) \frac{dB^r(1,0)}{dt} \frac{dA_r(1,1)}{dt} B^r(2,1) \\
& -\tau_r^r(1,1) B^r(1,0) \frac{dJ}{dt} A_r(1,1) B^r(2,1) \\
& -\tau_r^r(1,1) B^r(1,0) J \frac{dA_r(1,1)}{dt} B^r(2,1) \\
& -\tau_r^r(1,1) \frac{dB^r(1,0)}{dt} J A_r(1,1) B^r(2,1) \\
& +\tau_r^r(1,1) \frac{d^2B^r(1,0)}{dt^2} \\
& +\tau_r^r(1,1) \frac{dB^r(1,0)}{dt} J \\
& +\tau_r^r(1,1) B^r(1,1) \frac{dJ}{dt}
\end{aligned} \tag{B.8}$$

After regrouping, one obtains:

$$\begin{aligned}
\frac{dB^r(2,1)}{dt} = & \tau_r^r(1,1) \left[\frac{dB^r(1,0)}{dt} J + B^r(1,0) \frac{dJ}{dt} + \frac{d^2B^r(1,0)}{dt^2} \right] \times \\
& \times [I - A_r(1,1) B^r(2,1)] - \tau_r^r(1,1) \left[\frac{dB^r(1,0)}{dt} + B^r(1,0) J \right] \frac{dA_r(1,1)}{dt} B^r(2,1)
\end{aligned} \tag{B.9}$$

From:

$$\frac{dB^r(1,0)}{dt} + B^r(1,0) J = \Lambda_r^r(1,1) B^r(2,1) \tag{B.10}$$

one obtains:

$$\begin{aligned}
\frac{dB^r(2,1)}{dt} = & \tau_r^r(1,1) \left[\frac{dB^r(1,0)}{dt} J + B^r(1,0) \frac{dJ}{dt} + \frac{d^2B^r(1,0)}{dt^2} \right] \times \\
& \times [I - A_r(1,1) B^r(2,1)] - B^r(2,1) \frac{dA_r(1,1)}{dt} B^r(2,1)
\end{aligned} \tag{B.11}$$

The second order time derivative of $B^r(1, 0)$ can be obtained as:

$$\begin{aligned}
 \frac{d^2 B^r(1, 0)}{dt^2} &= \frac{d}{dt} \left[\tau_r^r(0, 0) B^r(0, 0) \frac{dJ}{dt} (I - A_r(0, 0) B^r(1, 0)) \right] \\
 &= \frac{d\tau_r^r(0, 0)}{dt} B^r(0, 0) \frac{dJ}{dt} \\
 &+ \tau_r^r(0, 0) B^r(0, 0) \frac{d^2 J}{dt^2} \\
 &- \frac{d\tau_r^r(0, 0)}{dt} B^r(0, 0) \frac{dJ}{dt} A_r(0, 0) B^r(1, 0) \\
 &- \tau_r^r(0, 0) B^r(0, 0) \frac{d^2 J}{dt^2} A_r(0, 0) B^r(1, 0) \\
 &- \tau_r^r(0, 0) B^r(0, 0) \frac{dJ}{dt} A_r(0, 0) \frac{dB^r(1, 0)}{dt}
 \end{aligned} \tag{B.12}$$

After regrouping, one obtains:

$$\begin{aligned}
 \frac{d^2 B^r(1, 0)}{dt^2} &= \left[\frac{d\tau_r^r(0, 0)}{dt} B^r(0, 0) \frac{dJ}{dt} + \tau_r^r(0, 0) B^r(0, 0) \frac{d^2 J}{dt^2} \right] [I - A_r(0, 0) B^r(1, 0)] + \\
 &\quad - \tau_r^r(0, 0) B^r(0, 0) \frac{dJ}{dt} A_r(0, 0) \frac{dB^r(1, 0)}{dt}
 \end{aligned} \tag{B.13}$$

Additional formulas for terms not appearing in the refinement procedure but required for diagnostic purposes, e.g. in the evaluation of the off diagonal blocks Λ_s^r and Λ_r^s are reported in the following:

$$\begin{aligned}
 \frac{dA_s(1, 0)}{dt} = \frac{dA_s(1, 1)}{dt} &= \frac{d}{dt} [(I - A_r(0, 0) B^r(1, 0)) A_s(0, 0)] \\
 &= -A_r(0, 0) \frac{dB^r(1, 0)}{dt} A_s(0, 0)
 \end{aligned}$$

$$\begin{aligned}
 \frac{dB^s(1, 1)}{dt} = \frac{dB^s(2, 1)}{dt} &= \frac{d}{dt} [B^s(1, 0) (I - A_r(1, 1) B^r(1, 1))] \\
 &= \frac{d}{dt} [B^s(0, 0) (I - A_r(1, 1) B^r(1, 0))] \\
 &= -B^s(0, 0) \left[\frac{dA_r(1, 1)}{dt} B^r(1, 0) + A_r(1, 1) \frac{dB^r(1, 0)}{dt} \right]
 \end{aligned}$$

$$\begin{aligned}\frac{dA_s(2,1)}{dt} &= \frac{d}{dt} [(I - A_r(2,1)B^r(2,1)) A_s(1,1)] = \frac{d}{dt} [(I - A_r(1,1)B^r(2,1)) A_s(1,0)] \\ &= - \left[\frac{dA_r(1,1)}{dt} B^r(2,1) + A_r(1,1) \frac{dB^r(2,1)}{dt} \right] A_s(1,0) + \\ &\quad + (I - A_r(1,1)B^r(2,1)) \frac{dA_s(1,0)}{dt}.\end{aligned}$$

Bibliography

- [1] D. A. Goussis M. Valorani, F. Creta, “Reactive and Reactive/Diffusive Time Scales in Stiff Reaction-Diffusion Systems,” *Progress In Computational Fluid Dynamics, An International Journal (PCFD)*, *in print*, April 2004.
- [2] A. Adrover, F. Creta, M. Giona, M. Valorani, and V. Vitacolonna, “Natural Tangent Dynamics with Recurrent Biorthogonalizations: a geometric computational approach to dynamical systems exhibiting slow manifolds and periodic/chaotic limit sets,” *Physica D*, *submitted*, 2004.
- [3] A. Adrover, F. Creta, M. Giona, and M. Valorani, “Biorthogonalization, invariant filtrations and Lyapunov spectra of chaotic dynamical systems,” *Physics Letters A*, *submitted*, 2004.
- [4] M. Valorani, D. A. Goussis, F. Creta, and H.N. Najim, “Higher Order Corrections in the Approximation of Low Dimensional Manifolds and the Construction of Simplified Problems with the CSP Method.,” *J. Comput. Phys.*, *in print*, October 2004.
- [5] M. Valorani, D. A. Goussis, F. Creta, and H.N. Najm, “An Automatic Procedure for the Simplification of Chemical Kinetics Mechanisms based on CSP.” .
- [6] A. Adrover, F. Creta, M. Giona, M. Valorani, and S. Cerbelli, “Slow Manifold Structure and Bifurcation in Prototypical Models of Thermal Explosions.,” .
- [7] J. C. Lee. F. Creta H.N. Habib D.A. Goussis, M. Valorani, “Chemical Kinetics Mechanism Simplification via CSP.”, in *3rd MIT Conference on Computational Fluid and Solid Mechanics 2003*, 2005.
- [8] M. Valorani and F. Creta, “Systematic Investigation of CSP Generated Simplified Kinetics Mechanisms for GRI3.0 Methane/Air Mixtures,” in *Joint Meeting of the Italian and Greek Sections of The Combustion Institute*, Corfu, Greece, June 2004.
- [9] F. Creta and M. Valorani, “Automatic Generation of Skeletal Kinetics for HC Fuels in LRE.,” in *European Conference for Aero-Space Sciences (EUCASS)*, Moscow, Russia, July 2005.

- [10] M. Valorani, F. Creta H.N. Habib D.A. Goussis, “Inertial manifolds with csp.,” in *2nd MIT Conference on Computational Fluid and Solid Mechanics 2003*, June 2003, vol. 2, pp. 1951–1954.
- [11] D. A. Goussis M. Valorani, F. Creta, “Local and Global manifols in stiff reaction-diffusion systems,” *2nd MIT Conference on Computationa Fluid and Solid Mechanics*, vol. 2, pp. 1548–1551, June 2003.
- [12] D.A. Goussis and S.H. Lam, “A Study of Homogeneous Methanol Oxidation Kinetic Using CSP,” in *Proc. Comb. Inst.*, 1992, vol. 24, pp. 113–120.
- [13] M. Hadjinicolaou and D.A. Goussis, “Asymptotic Solution of Stiff PDEs with the CSP Method - The Reaction Diffusion Equation,” *SIAM J. Sci. Comp.*, vol. 20, pp. 781–810, 1999.
- [14] S.H. Lam, “Using CSP to Understand Complex Chemical Kinetics,” *Combustion Science and Technology*, vol. 89, pp. 375–404, 1993.
- [15] S.H. Lam and D.A. Goussis, “Understanding Complex Chemical Kinetics with Computational Singular Perturbation,” in *Proc. Comb. Inst.*, 1988, vol. 22, p. 931.
- [16] S.H. Lam and D.A. Goussis, “Conventional Asymptotics and Computational Singular Perturbation for Simplified Kinetics Modelling,” in *Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames*, M.O. Smooke, Ed., number 384 in Springer Lecture Notes. Springer Verlag, 1991.
- [17] S.H. Lam and D.A. Goussis, “The CSP Method for Simplifying Kinetics,” *International Journal of Chemical Kinetics*, vol. 26, pp. 461–486, 1994.
- [18] A. Massias, D. Diamantis, E. Mastorakos, and D.A. Goussis, “An Algorithm for the Construction of Global Reduced Mechanisms with CSP Data,” *Combustion and Flame*, vol. 117, pp. 685–708, 1999.
- [19] A. Zagaris, H. G. Kaper, and T. J. Kaper, “Analysis of the Computational Singular Perturbation Reduction Method for Chemical Kinetics.,” *Nonlinear Science*, vol. 14, pp. 59–91, 2004.
- [20] U. Maas, R. W. Dibble, J. Warnatz, and E. Zwicker, “*Combustion: Physical and Chemical Fundamentals, Modeling and Simulation, Experiments, Pollutant Formation*”, Springer-Verlag, Berlin, 1999.

- [21] J. H. Seinfeld and S. N Pandis, “*Atmospheric Chemistry and Physics: From Air Pollution to Climate change*”, Wiley, New York, 1998.
- [22] J. Monod, J. Wyman, and J. P. Changeux, “On the Nature of Allosteric Transitions: A Plausible Model,” *J. Mol. Biol.*, vol. 12, 1965.
- [23] L. A. Segel and M. Slemrod, “The quasi steady state assumption: a case study in perturbation.,” *SIAM Rev.*, vol. 31, pp. 446–447, 1989.
- [24] A. N. Yannacopoulos, A. S. Tomlin, J. Brindley, J. H. Merkin, and M. J. Pilling., “The error of the quasi steady state approximation in spatially distributed systems.,” *Chem. Phys. Lett.*, , no. 248, pp. 63–70, 1996.
- [25] M. D. Smooke (Ed.),, “Reduced Kinetic Mechanisms and Asymptotic Approximations for Methane-Air Flames.,” *Lectur Notes in Physics.*, vol. 348, 1991.
- [26] Maas U. and Pope S.B., “Implementation of simplified chemical kinetics based on intrinsic low-dimesional manifolds,” in *Proc. of the 24th International Symposium on Combustion.*, 1992, pp. 103–112.
- [27] Maas U. and Pope S.B., “Simplifying chemical kinetics: intrinsic low-dimensional manifolds in composition space.,” *Combustion and Flame*, vol. 88, no. 11, 1992.
- [28] M. R. Roussel and S. J. Fraser, “Invariant Manifold Methods for Metabolic Model Reduction.,” *Chaos*, vol. 11, no. 1, 2001.
- [29] M. R. Roussel and S. J. Fraser, “Geometry of the steady-state approximation. Perturbation and accelerated convergence method.,” *J. Chem. Phys.*, vol. 93, pp. 1072, 1990.
- [30] M. R. Roussel and S. J. Fraser, “On the geometry of transient relaxation.,” *J. Chem. Phys.*, vol. 94, pp. 7106, 1991.
- [31] M. J. Davis and R. T. Skodje, “Geometric Approach to Multiple-Time-Scale Kinetics: A Nonlinear Master Equation Describing Vibration-to-Vibration Relaxation.,” *Zeitschrift fur Physikalische Chemie*, , no. 215, pp. 233–252, 2001.
- [32] T. F. Lu and C. K. Law, “A Directed Relation Graph Method for Mechanism Reduction.,” *Preprint*, 2004.

- [33] N. Fenichel., “Geometric Singular Perturbation Theory for Ordinary Differential Equations.,” *J. Diff. Eq.*, vol. 31, pp. 53–98, 1979.
- [34] C. K. R. T. Jones., “Geometric Singular Perturbation Theory.,” *Dynamical Systems, Montecatini Terme.*, vol. 1609, 1994.
- [35] H. G. Kaper and T. J. Kaper, “Asymptotic analysis of two reduction methods for systems of chemical reactions,” *Physica D*, vol. 165, no. 66, 2002.
- [36] M. J. Davis and R. T. Skodje, “Geometric investigation of low-dimensional manifolds in systems approaching equilibrium,” *Chemical Physics*, vol. 111, no. 859, 1999.
- [37] L. Michaelis and M. L. Menten, ,” *Biochem. Z.*, vol. 49, 1913.
- [38] V. Henri, “Lois generales de l'action des diastases.,” *C. R. Acad. Sci.*, vol. 135, 1902.
- [39] S. J. Fraser, “The steady state and equilibrium approximations: a geometrical picture.,” *J. Chem. Phys.*, , no. 88, pp. 4732–4738, 1988.
- [40] V. Franceschini and C. Tebaldi, “Sequences of Infinite Bifurcations and Turbulence in Five-Mode Truncation of the Navier-Stokes Equations,” *J. Stat. Phys.*, vol. 21, no. 707, 1979.
- [41] A. C. Hindmarsh, “Lois generales de l'action des diastases.,” *Scientific Computing, R. S. Stepleman et al.*, 1983.
- [42] S.H.Lam, “Reduced chemistry modeling and sensitivity analysis.” *Lecture Notes for Aerothermochemistry for Hypersonic Technology, 1994-1995 Lecture Series Programme, Von Karman Institute for Fluid Dynamics*, p. 3, Apr. 1995.
- [43] M. Valorani and D. A. Goussis H. N. Najim, “CSP analysis of a transient flame-vortex interaction: time scales and manifolds,” *Combustion and Flame*, , no. 134, pp. 35–53, 2003.
- [44] H. Wang C.K. Law, C.J. Sung and T.F. Lu, “Development of Comprehensive Detailed and Reduced Reaction Mechanisms for Combustion Modeling,” *AIAA Journal*, vol. 41, no. 9, September 2003.
- [45] Gregory P. Smith, David M. Golden, Michael Frenklach, Nigel W. Moriarty, Boris Eiteneer, Mikhail Goldenberg, C. Thomas Bowman, Ronald K. Hanson, Soonho Song, William C. Gardiner Jr., Vitali V. Lissianski, and Zhiwei Q., “Gri mechanism for methane/air, version 3.0.,” www.me.berkeley.edu/gri_mech/

- [46] D. Lentini and I.K. Puri, "Radiation and NO_x pathways in nonpremixed turbulent flames," *Journal of Propulsion and Power*, , no. 17, pp. 222–224, 2001.
- [47] D. Lentini, "Prediction of NO_x emissions in gas turbines combustors inclusive of the N₂O contribution," *Journal of Power and Energy, Proc. Instn. Mech. Engrs. Part A*, , no. 217, pp. 83–90, 2003.
- [48] J.A. Miller and C.T. Bowman., "Mechanism and modeling of nitrogen chemistry in combustion.," *Prog. Energy Combust. Sci.*, vol. 113, pp. 287–338, 1989.
- [49] P. H. Paul C. J. Mueller H. N. Najm and P. S. Wyckoff., "On the adequacy of certain experimental observables as measurements of flame burning rate.," *Combustion and Flame*, vol. 113, pp. 312–332, 1998.
- [50] B. S. Brewster S. M. Cannon and L.D. Smoot., "Stochastic modeling of co and no in premixed methane combustion.," *Combustion and Flame*, vol. 113, pp. 135–146, 1998.
- [51] P. C. Malte D. G. Nicol R. C. Steele and J. C. Kramlich., "No_x and n₂0 in lean-premixed jet-stirred flames.," *Combustion and Flame*, vol. 110, pp. 440–449, 1995.
- [52] S. C. Hill and L. D. Smoot., "Modeling of nitrogen oxides formation and destruction in combustion systems.," *Prog. Energy Combust. Sci.*, vol. 26, pp. 417–458, 2000.
- [53] J. Tomeczek and B. Gradon., "The role of n₂o and nnh in the formation of no via hcn in hydrocarbon flames.," *Combustion and Flame*, vol. 133, pp. 311–322, 2003.
- [54] A. A. Konnov and J. De Ruyck., "Temperature-dependent rate constant for the reaction nnh + o → nh + no.," *Combustion and Flame*, vol. 125(4), pp. 1258–1264, June 2001.
- [55] O. Gicquel, N. Darabiha, and D. Thevenin, , " in *Proc. Comb. Inst.*, 2000, pp. 1901–1908.
- [56] S. Singh, J. M. Powers, and S. Paolucci, "On slow manifolds of chemically reactive systems," *Journal of Chemical Physics*, vol. 117, no. 4, pp. 1482–1496, July 2002.
- [57] S. S. Girimaji and C. Brau, , " *Theoret. Comput. Fluid Dynamics*, vol. 17, pp. 171–188, 2002.
- [58] J. A. vanOijen, F. A. Lammers, and L. P. H. deGoey, , " *Combustion and Flame*, vol. 127, pp. 2124–2134, 2001.
- [59] H. Bongers, J. A. vanOijen, and P. H. deGoey, , " in *Proc. Comb. Inst.*, 2002, number 29, pp. 1371–1378.

- [60] Habib N. Najm, Peter S. Wyckoff, and Omar M. Knio, “A Semi-implicit Numerical Scheme for Reactive Flow,” *J. Comp. Phys.*, , no. 143, pp. 381–402, 1998.
- [61] Frenklach M., Wang H., Goldenberg M, Smith G. P., Golden D. M., Bowman T, Hanson R. K., Gardiner W. C., and Lissianski V., “GRI-Mech – An Optimized Detailed Chemical Reaction Mechanism for Methane Combustion,” *Gas Research Inst., Chicago.*, November 1995.
- [62] A. E. Lutz, R. J. Kee, J. F Grcar, and F. M. Rupley, “Oppdiff:A Fortran program for computing opposed-flow diffusion flames,” *Sandia Report SAND96-8243 UC-1409*, May 1997.
- [63] B. R. Johnsn and S. K. Scott, “Period doubling and chaos during the oscillatory ignition of the CO+O₂ reaction,” *J. Chem Soc. Faraday Trans.*, , no. 86, pp. 3701, 1990.
- [64] B. Hess and A. Boiteux, “Oscillatory phenomena in biochemistry,” *Ann. Rev. Biochem.*, , no. 40, pp. 237, 1971.
- [65] L. F. Olsen and H. Degn, “Chaos in enzyme reactions,” *Nature*, , no. 267, pp. 177, 1977.
- [66] A. Katok and L. Mendoza, “*Dynamical systems with nonuniform hyperbolic behavior*” (*Introduction to the modern theory of dynamical systems by A. Katok and B. Hasselblatt*), Cambridge University Press, Cambridge, 1995.
- [67] P. Holmes J. Guckenheimer, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Springer-Verlag New York Inc., New York, Berlin, Heidelberg, 7th Edition, 1983.
- [68] L. Arnold, *Random Dynamical Systems*, Springer-Verlag, New York, Berlin, Heidelberg, 1st Edition, 1998.
- [69] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag New York Inc., New York, Berlin, Heidelberg, 2nd Edition, 2003.
- [70] Y. B. Pesin L. Barreira, *Lyapunov Exponents and Smooth Ergotic Theory*, vol. 23, University Lecture Series, American Mathematical Society, Providence, RI, 2002.
- [71] W. Magnus, “On the exponential solution of differential equations for a linear operator,” *Comm. Pure and Appl. Math.*, vol. 7, no. 649, 1954.
- [72] S. P. Norsett A. Iserles, “On the Solution of Linear Differential Equations in Lie Groups,” *University of Cambridge Numerical Analysis Reports*, Mar. 1997.

- [73] S. P. Norsett A. Iserles, “Linear ODEs in Lie Groups,” *University of Cambridge Numerical Analysis Reports*, May 1997.
- [74] S. Smale M. W. Hirsch, *Differential Equations, Dynamical Systems and Linear Algebra*, Academic Press Inc., Orlando, FL, 1974.
- [75] Rex T. Skodje and Michael J. Davis, “Geometric Simplification of Complex Kinetic Systems,” *J. Phys. Chem.*, , no. 105, pp. 10356–10365, 2001.
- [76] K. D. Mease, S. Bharadwaj, and S. Iravanchy, “Timescale Analysis for Nonlinear Dynamical Systems,” *Journal of Guidance and Control*, vol. 26, no. 2, pp. 318–330, 2003.
- [77] M. Giona, A. Adrover, F. J. Muzzio, S. Cerbelli, and M. M. Alvarez, “The geometry of mixing in time-periodic chaotic flows. I. Asymptotic directionality in physically realizable flows and global invariant properties,” *Physica D*, vol. 132, pp. 298–324, 1999.
- [78] V.I. Oseledec, “A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems,” *Trans. Moscow Math. Soc.*, vol. 19, no. 197, 1968.
- [79] M. Giona A. Adrover, S. Cerbelli, “Exterior Algebra-based Algorithms to estimate Liapunov spectra and stretching statistics in high-dimensional distributed systems,” *Int. J. Bifurcation and Chaos*, vol. 12, no. 353, 2002.