ACADEMIC YEAR 2012/2013

Corso di Laurea Magistrale in Ingegneria Nucleare
Scuola di Ingegneria Industriale e dell'Informazione

Entropy-Gradient Dynamical Models for a Thermodynamic System and Their Realisation in Kinetic Theory

Advisor:	Prof. Alfonso Niro	(Politecnico di Milano)
Co-Advisor:	Prof. Gian Paolo Beretta	(Università degli Studi di Brescia)
Co-Advisor:	Prof. Romano Borchiellini	(Politecnico di Torino)

Master of Science Thesis by

Francesco Leonardo Consonni	(Matricola: 784132)
Alberto Montefusco	(Matricola: 783936)

Contents

List of Figures

Abstract in English

Thermodynamics is divided into the branches of *kinematics*, which deals with the description of the possible states of a system, and *dynamics*, which aims at describing the causes and the effects of the motion of the system. Regarding the second aspect, a dynamical law of evolution for thermodynamic systems does not exist. The resolution of this problem would not only be a milestone in the field of Physics, but it would also allow a more precise description of the evolution of non-equilibrium systems, a need felt in the most different sectors of engineering. In particular, through its realization in Kinetic Theory, it would help in modelling the time-evolution of rarefied systems for which the Navier-Stokes equations do not apply. Among the dynamical principles, one of those that have most frequently been proposed is the Maximum-Entropy Production Principle.

The present work is focused on the dynamical modelling of thermodynamic systems and is articulated in three parts. The first is a systematic review of some dynamic principles proposed during the history of thermodynamics. The second aims at understanding similarities and differences between the Steepest Entropy Ascent (SEA) dynamical model proposed by Beretta and the GENERIC dynamic formalism developed, among others, by Öttinger e Grmela. In order to accomplish this task, a reformulation of SEA dynamics using Differential Geometry formalism has been considered necessary and constitutes one of the most innovative outputs of the present thesis. It is shown that both dynamic models are of the *entropy-gradient* type, the main difference being that GENERIC is built in a more structured manifold. In the third part, the realization of both these dynamical models in Kinetic Theory is illustrated: the Boltzmann equation is *interpreted* differently using the *building blocks* of the two models. Moreover, as SEA aims at proposing new model equations for its resolution, numerical results of the application of SEA methods for the relaxation from non-equilibrium states are presented: good agreement with the *exact* solution is shown for near-equilibrium situations, while poorer results are obtained farther from equilibrium. This means that improvements, with particular regard to the choice of the *metric*, are needed.

Keywords: non-equilibrium thermodynamics, dynamical models, Steepest Entropy Ascent, GENERIC, Boltzmann Equation, kinetic models.

Abstract in Italiano

La termodinamica è divisa in *cinematica*, che si occupa della descrizione dei possibili stati di un sistema, e *dinamica*, che mira a descrivere le cause e gli effetti del moto del sistema. In merito al secondo aspetto, una legge di evoluzione dinamica per un sistema termodinamico non esiste. La soluzione a questo problema non solo rappresenterebbe una pietra miliare nel campo della Fisica, ma permetterebbe anche una descrizione più accurata dell'evoluzione dei sistemi di non-equilibrio, un'esigenza sentita nei più diversi settori dell'ingegneria. In particolare, attraverso la sua realizzazione nella Teoria Cinetica, aiuterebbe a modellare l'evoluzione temporale di sistemi rarefatti, per cui le equazioni di Navier-Stokes non sono applicabili. Tra i vari principi dinamici, uno di quelli più frequentemente proposti è il Principio di Massima Produzione di Entropia.

Il presente lavoro di tesi si focalizza sulla modellazione dinamica di sistemi termodinamici ed è articolato in tre parti. La prima è una review sistematica di alcuni principi dinamici proposti nella storia della termodinamica. La seconda parte ha come obiettivo la comprensione delle analogie e delle differenze tra il modello dinamico Steepest Entropy Ascent, proposto da Beretta, e il formalismo GENERIC, sviluppato, tra gli altri, da Öttinger e Grmela. Per realizzare questo, si è resa necessaria una riformulazione della dinamica SEA attraverso il formalismo della Geometria Differenziale. Questa parte costituisce uno degli output più innovativi della presente tesi. Si mostra che entrambi i modelli dinamici sono del tipo *gradiente-di-entropia*, con la principale differenza che GENERIC è sviluppato in una *varietà* più strutturata. Nella terza parte del lavoro, si illustra la realizzazione di entrambi i modelli dinamici nella Teoria Cinetica: l'equazione di Boltzmann è *interpretata* in modo differente utilizzando i *blocchi costitutivi* dei due modelli. Inoltre, dal momento che SEA mira a proporre nuovi modelli per la risoluzione dell'equazione, si presentano i risultati numerici dell'applicazione di metodi SEA per il rilassamento da uno stato di non-equilibrio: un buon accordo con la soluzione *esatta* è evidenziato per situazioni vicine all'equilibrio, mentre lontano dall'equilibrio i risultati sono meno soddisfacenti. Questo significa che sono necessari miglioramenti, con particolare riferimento alla scelta della *metrica*.

Parole Chiave: termodinamica del non-equilibrio, modelli dinamici, Steepest Entropy Ascent, GENERIC, Equazione di Boltzmann, modelli cinetici.

Estratto in Italiano

La Termodinamica nella Fisica e nell'Ingegneria

Beretta e Gyftopoulos definiscono la Termodinamica come *lo studio delle osservabili fisiche dei moti di costituenti fisici (particelle e radiazioni), dovuti a forze applicate esternamente, o da forze interne.* Nonostante i concetti base della Termodinamica siano stati utilizzati per la realizzazione di strumenti pratici già dal XVII Secolo e nonostante i primi tentativi di sistemazione teorica risalgano all'Ottocento (con la formulazione di Carnot, secondo il concetto di *ciclo*), esistono ancora oggi notevoli divergenze interpretative sui concetti base e sulla struttura dell'impianto teorico. Una visione non comune, ma a nostro avviso chiara e rigorosa, della disciplina è quella fornita dalla *Keenan School* del Massachusetts Institute of Technology, cui appartegono, oltre al fondatore Joseph Keenan e George Hatsopoulous, anche i già citati Beretta e Gyftopoulos. Essi propongono una visione della Termodinamica come estensione della Meccanica, un impianto teorico che mira a evitare i 'loop' logici e le ambiguità delle esposizioni tradizionali, una definizione di entropia valida per tutti gli stati, inclusi quelli di non-equilibrio, e, infine, l'idea che l'irreversibilità sia intrinsecamente contenuta nella natura microscopica dei fenomeni. Questa visione si oppone decisamente ad una delle formulazioni più note, quella che interpreta i principi della termodinamica in senso statistico, e alla visione per cui l'irreversibilità emerge nel passaggio dal livello microscopico al livello macroscopico.

Indipendentemente dalla visione di insieme, essa è, come tutte le discipline della fisica, divisa nella branca della *cinematica*, relativa alla descrizione dei possibili stati di un sistema, e della *dinamica*, relativa allo studio delle cause e degli effetti del moto del sistema. Sotto quest'ultimo punto di vista, una legge di evoluzione dinamica per un sistema termodinamico, analoga alla Legge di Newton per la Meccanica Classica e all'Equazione di Schrödinger per la Meccanica Quantistica, non esiste. Essa non sarebbe solamente una pietra miliare nella storia della fisica, ma aiuterebbe anche a modellare meglio l'evoluzione temporale dei sistemi fuori equilibrio, un'esigenza sentita nei più svariati settori dell'ingegneria. Ad esempio, un semplice processo di ossidazione può avvenire in diversi modi: mediante una combustione con fiamma, processo caratterizzato da maggiori irreversibilità e minore produzione di lavoro utile, oppure attraverso una cella a combustibile, processo caratterizzato da minori irreversibilità e maggiore produzione di lavoro utile. Una migliore conoscenza dell'evoluzione temporale dei sistemi termodinamici permetterebbe una migliore gestione di questi fenomeni. In questo ambito, gli scopi della tesi sono:

- effettuare una review di alcuni principi dinamici per sistemi termodinamici che sono stati proposti nel tempo, con particolare riferimento al Principio di Massima Produzione di Entropia;

- confrontare la teoria dinamica Steepest Entropy Ascent, proposta da Beretta, con il formalismo GENERIC, sviluppato, tra gli altri, da Öttinger e Grmela; entrambe le strutture sono state motivate dalla ricerca di formulazioni della Termodinamica del Non-Equilibrio pienamente compatibili con la Seconda Legge della Termodinamica;

- verificare come queste due teorie si applicano alla Teoria Cinetica e all'equazione di Boltzmann, al fine di proporre nuovi modelli cinetici che possano essere utili nello studio dell'ampia varietà di fenomeni fisici che vengono solitamente modellati attraverso l'equazione di Boltzmann stessa, quali il moto di elettroni in un conduttore, il comportamento dei fononi in un isolante, il trasporto dei neutroni in un reattore, il comportamento di un plasma e il comportamento di gas rarefatti o, più in generale, la modellazione di tutte quelle situazioni fisiche in cui, a causa del numero di Knudsen significativamente diverso da zero, la validità delle equazioni di Navier-Stokes cade.

La validazione del modello

Il secondo capitolo del lavoro si propone di illustrare brevemente, a causa della loro vastità, i due *impianti teorici* che sono utilizzati per *validare* i modelli dinamici successivamente illustrati: la Termodinamica Classica del Non-Equilibrio e la Teoria Cinetica. Da un lato, infatti, la Termodinamica Classica del Non-Equilibrio, soprattutto nella sua approssimazione lineare, è utilizzata per modellizzare fenomeni che non si discostano troppo da una situazione di equilibrio: pertanto, qualsiasi modello dinamico valido deve contemplarne gli aspetti peculiari nelle situazioni vicine all'equilibrio. Dall'altro lato, la Teoria Cinetica e, soprattutto, l'Equazione di Boltzmann, sono *utilizzate* sia nell'ambito del GENERIC, sia nell'ambito di SEA, come *banco di prova* per la loro struttura: i vari termini dell'Equazione di Boltzmann sono *interpretati* in maniera differente attraverso i *blocchi costitutivi* delle due teorie.

Per quanto riguarda la Termodinamica Classica del Non-Equilibrio, si illustrano le ipotesi che vi sono alla base, il modo in cui la produzione entropica locale emerge come prodotto tra forze e flussi, come questi ultimi possono essere considerati una funzione lineare delle forze, nonché il principio di Curie e le relazioni di reciprocità di Onsager, la cui dimostrazione è basata sull'ipotesi della *reversibilità microscopica* delle equazioni del moto e su altri punti che sono illustrati. Si illustra poi il Principio di Minima Entropia di Prigogine, chiarendone il legame con il Principio di Massima Entropia: mentre il primo è un principio *globale* (si applica a tutto il corpo preso in considerazione), legato all'andamento nel tempo della produzione entropica, il secondo è generalmente interpretato come un principio *locale*, legato alle possibili direzioni di evoluzione di un sistema ad un istante temporale fissato. Inoltre, mentro il secondo è del tutto generale, il primo è valido sotto ipotesi restrittive. Si mostrano infine alcuni ambiti di applicazione dell'impianto della Termodinamica Classica di Non-Equilibrio, come gli effetti *termoelettrici*, gli effetti *termomeccanici* e l'effetto Righi-Leduc.

Nella seconda parte del capitolo si illustra la derivazione dell'Equazione di Boltzmann, preceduta dall'illustrazione dell'Equazione di Liouville, valida in assenza di collisioni, e introducendo l'ipotesi della *Stosszahlansatz*, detta anche *ipotesi del caos molecolare*. Si espone il concetto di *invariante collisionale* e si vede come essi possano essere espressi come combinazione lineare di massa, momento ed energia. Infine si spiega come si ricava l'espressione per la soluzione di equilibrio dell'equazione, ossia la Maxwelliana, come la risoluzione dell'equazione possa essere semplificata attraverso opportuni *modelli cinetici* o *metodi approssimati* e quali siano le applicazioni dell'equazione.

Alla ricerca di un modello dinamico

Il terzo capitolo del lavoro si propone come review di alcune significative teorie elaborate nel corso del XX Secolo nell'ambito della Termodinamica dei Processi Irreversibili. Tra queste, si illustra la prima sistematica esposizione del Principio di Massima Produzione di Entropia, proposto da Hans Ziegler. L'esposizione di Ziegler costituisce anche una possibile *geometrizzazione* della Termodinamica del Non-Equilibrio in quanto il principio di massimo vincolato da lui proposto è anche conosciuto come *principio di ortogonalità*, in quanto ha come conseguenza il fatto che la derivata della produzione entropica rispetto ai flussi debba essere parallela alle forze (la ragione dell'*ortogonalità* è illustrata nel Capitolo). Inoltre, partendo dal principio di Ziegler e assumendo un legame lineare tra forze e flussi, è possibile dimostrare le relazioni di reciprocità di Onsager. Dopo aver brevemente introdotto il modello SEA, si illustra la teoria proposta da Edelen, basata sulla scomposizione dei flussi in due componenti, una *dissipativa* e una *non dissipativa*; questo approccio può essere visto come *precursore* del formalismo GENERIC, illustrato nel Capitolo successivo.

Termodinamica geometrica

Il quarto capitolo del lavoro si concentra sul rapporto tra la Termodinamica e la Geometria: probabilmente unica tra le discipline della Fisica, infatti, come sostiene Mrugała, la Termodinamica non è ancora stata sistematicamente *geometrizzata*. La Meccanica Classica, ad esempio, è stata invece razionalizzata da un punto di vista geometrico utilizzando le strutture delle *varietà simplettiche* e delle *varietà di Poisson*, la cui illustrazione occupa la prima parte del capitolo e costituisce la base per la successiva illustrazione della struttura del GENERIC. L'esigenza di *geometrizzare* la Termodinamica è stata recepita, sul versante delle situazioni di equilibrio da Carathéodory ed altri, mentre, sul versante della Termodinamica di Non-Equilibrio, il tentativo più compiuto è la cosiddetta *dinamica metriplettica* che ha una delle formulazioni più compiute nel GENERIC. L'obiettivo del capitolo è pertanto quello di confrontare questo tentativo di *geometrizzazione* della termodinamica con l'approccio Steepest Entropy Ascent proposto da Beretta, inizialmente in un framework quantistico e successivamente applicato anche a sistemi meso- e macroscopici. Per effettuare il confronto, è stato necessario rielaborare il modello SEA originariamente proposto per fornirne una versione più matematica e astratta, utilizzando il formalismo della Geometria Differenziale. A valle di questa riformulazione, le analogie e le differenze tra le due teorie emergono in maniera molto chiara e possono essere sinteticamente elencate come segue:

- Steepest Entropy Ascent si focalizza solo sulla modellazione della parte *dissipativa* della dinamica, utilizzando una metrica non-degenere, mentre GENERIC modellizza esplicitamente sia la parte *non-dissipativa* (*Hamiltoniana*), utilizzando una struttura geometrica tipica della Meccanica Classica, sia la parte *dissipativa* (*irreversibile*), utilizzando una *cometrica* degenere;

- essendo l'approccio Steepest Entropy Ascent meno strutturato in partenza, l'imposizione della costanza delle quantità conservate lungo la traiettoria del processo termodinamico è effettuata *a posteriori*, mentre GENERIC la considera già imponendo condizioni di degenerazione sulle strutture che modellano i due tipi di dinamica;

- al fine di poter effettuare un vero parallelismo tra i due costrutti, è necessario poter definire i gradienti anche mediante la cometrica degenere che caratterizza la parte dissipativa del GENERIC, cosa che può essere fatta mediante l'aggiunta di un'ulteriore condizione sulla struttura stessa;
- a valle di questa definizione, è possibile evidenziare che il GENERIC altro non è se non uno Steepest Entropy Ascent, ossia un moto nella direzione del gradiente dell'entropia, su *foglie metriche*, ossia superfici dello spazio caratterizzate dalla costanza dei valori dell'energia e delle quantità conservate;
- è possibile concludere che l'approccio Steepest Entropy Ascent è più generale, in quanto meno vincolato, e che qualunque dinamica di tipo GENERIC, che soddisfi l'ulteriore condizione dell'*identità di Leibniz*, è automaticamente Steepest Entropy Ascent.

Nella parte finale del capitolo si illustrano il legame tra la originaria formulazione dinamica con *doppio generatore* proposta da Edelen e la formulazione del GENERIC, e il modo in cui, dalla formulazione GENERIC, emerge l'approssimazione della Termodinamica Classica del Non-Equilibrio.

Applicazioni

Quest'ultimo capitolo si occupa dell'illustrazione delle applicazioni del GENERIC e dello Steepest Entropy Ascent a equazioni e modelli correntemente utilizzati in Fisica, permettendo così di comprendere anche quali sono le differenze *filosofiche* di fondo tra i due modelli dinamici, oltre a quelle *geometriche* illustrate nel Capitolo precedente. In primo luogo si mostra come le equazioni dell'idrodinamica classica possano essere inquadrate nell'ambito del formalismo GENERIC che separa esplicitamente le componenti di evoluzione temporale delle variabili di stato (che sono responsabili dell'avanzamento dello stato nello spazio di interesse), le componenti avvettive delle equazioni, che rappresentano la parte *Hamiltoniana*, e le componenti irreversibili, che costituiscono la parte *dissipativa.* Si illustra poi un analogo inquadramento delle equazioni della magnetoidrodinamica mediante il formalismo GENERIC.

La parte principale del capitolo si concentra, tuttavia, sulla differente interpretazione dell'equazione di Boltzmann. I vari termini dell'equazione vengono *associati* ai differenti *blocchi costitutivi* dei due modelli. In particolare, per quanto riguarda GENERIC, si individua un *operatore di Poisson*, definito in ogni punto dello spazio, che permette di ricostruire la parte avvettiva dell'Equazione di Boltzmann. Dall'altro lato, si individua un *operatore dissipativo* che permette di ricostruire la parte collisionale. Analogamente a quanto fatto per il GENERIC, anche SEA separa le due parti reversibile e dissipativa dell'equazione, ma non esplicita la scelta della metrica. In questo ambito è possibile comprendere la differenza fondamentale nell'interpretazione dell'equazione di Boltzmann tra i due modelli dinamici: GENERIC mira a riprodurre le equazioni tali e quali, riscrivendole semplicemente utilizzando il proprio formalismo, mentre SEA mira a cercare una metrica adatta a creare un nuovo modello cinetico, ossia una metrica *semplificata* rispetto a quella *esatta* del GENERIC, che però possa essere più utile per la risoluzione numerica dell'equazione.

Nell'ultima parte del capitolo si illustrano i risultati numerici del rilassamento da uno stato di non-equilibrio spazialmente omogeneo mediante l'equazione di Boltzmann, recentemente ottenuti da Beretta e Hadjiconstantinou. Si effettua il confronto tra la soluzione ritenuta *esatta*, ottenuta mediante simulazione Montecarlo, e differenti modelli cinetici, ossia espressioni semplificate per l'integrale collisionale. I modelli utilizzati sono il BGK standard, il BGK a frequenza di collisione variabile e due modelli SEA che differiscono proprio per la scelta della metrica: il primo è caratterizzato da una metrica uniforme (di Fisher), mentre il secondo è caratterizzato da una metrica modificata dalla presenza di una *funzione peso* non unitaria. Si mostra come i modelli cinetici SEA soddisfino i requisiti fondamentali di un modello cinetico, ossia la conservazione degli invarianti collisionali e il teorema H, nonché il fatto che, per piccoli scostamenti dall'equilibrio, essi convergano ai corrispondenti metodi BGK. I risultati numerici mostrano tuttavia una scarsa compatibilità dell'evoluzione temporale proposta dai modelli SEA con la soluzione *esatta*, in particolare modo per una situazione iniziale lontana dall'equilibrio. Per situazioni iniziali più vicine all'equilibrio, invece, i modelli SEA risultano effettivamente avere andamenti molto più vicini a quelli dei corrispondenti modelli BGK e, conseguentemente, riprodurre meglio l'andamento temporale della soluzione *esatta*.

Conclusioni e sviluppi futuri

I tre obiettivi, che erano stati fissati all'inizio del lavoro e che sono stati precedentemente elencati, sono stati raggiunti.

Per quanto riguarda il primo obiettivo, sono state chiarite le differenze tra il Principio di Massima Produzione di Entropia e il Principio di Minima Produzione di Entropia di Prigogine, la struttura geometrica della teoria elaborata da Ziegler, con le sue implicazioni riguardanti le relazioni di reciprocità, e, infine, si è individuato il formalismo di Edelen come precursore della teoria dei *due generatori* su cui si basa il GENERIC.

Per quanto riguarda il secondo obiettivo, dopo aver individuato la Geometria Differenziale come lo scenario ideale per il confronto tra i due modelli termodinamici cha hanno l'ambizione di *geometrizzarne* l'evoluzione temporale, il modello Steepest Entropy Ascent è stato riscritto in termini più matematici e astratti in quella che è, a nostro avviso, la parte più innovativa del lavoro. Si mostra come entrambi i modelli dinamici siano del tipo *gradiente di entropia* e come SEA sia, in quanto meno strutturato, più generale della dinamica sviluppata, tra gli altri, da Öttinger e Grmela, la quale, con l'aggiunta di una semplice condizione, si può considerare uno Steepest Entropy Ascent su *foglie metriche*. Si è inoltre verificata la compatibilità dei due modelli con l'approssimazione lineare della Termodinamica Classica del Non-Equilibrio.

Infine, per quanto riguarda il terzo obiettivo, si è compreso il differente approccio delle due modellazioni dinamiche all'equazione di Boltzmann: da un lato, GENERIC mira a *riscrivere* l'equazione nella sua forma originale individuando però i *blocchi costitutivi* fondamentali, nello specifico, l'operatore di Poisson e l'operatore che regola la dinamica irreversibile. Dall'altro lato, SEA mira a proporre un modello cinetico che, sulla linea degli altri modelli cinetici che sono stati proposti nel corso del tempo, possa semplificare la risoluzione numerica dell'equazione. Si mostra come i due modelli SEA proposti soddisfino i requisiti fondamentali richiesti ad un modello cinetico, ma diano risultati numericamente peggiori rispetto a quelli ottenuti dai modelli BGK corrisponenti, a cui tendono per situazioni di basso scostamento dall'equilibrio.

Le conclusioni ottenute forniscono lo spunto per ulteriori sviluppi:

- rimane aperto il problema, chiave nella Cinematica, dell'identificazione sistematica delle quantità conservate;

- rimane aperto il problema della modellazione delle interazioni, ossia della *geometrizzazione* dei sistemi aperti, dal momento che ciò che è stato ottenuto (in modo particolare il GENERIC) è legato a sistemi chiusi;
- si suggerisce di sviluppare algoritmi numerici che permettano di verificare la validità dell'identità di Leibniz nei casi pratici più complicati, analogamente a quanto fatto da Kröger, Hütter e Öttinger per l'identità di Jacobi;
- la questione della metrica da utilizzare nel modello cinetico SEA rimane aperta e, dal momento che dallo studio del GENERIC non si sono ricavate risposte convincenti, possibili suggerimenti potrebbero essere trovati nel campo emergente dell'Information Geometry.

> *“[...] the traditional meaning of the term thermodynamics needs to be reconsidered. Physics is the science that attempts to describe all aspects of all phenomena pertaining to the perceivable universe. It can be viewed as a large tree with many branches, such as mechanics, electromagnetism, gravitation, and chemistry, each specialized in the description of a particular class of phenomena. Thermodynamics is not a branch. It pervades the entire tree. To emphasize this conception, we often use the words physics and thermodynamics as synonyms.”*
>
> Elias P. Gyftopoulos and Gian Paolo Beretta *in*
> Thermodynamics: Foundations and Applications

1

Thermodynamics in Physics and Engineering

The aim of the present chapter is, first of all, to give the reader a brief historical introduction and a general, not universally accepted, overview of the subject of thermodynamics, which is the topic that frames the present work. Then, a basic categorization of the areas of study in the subject will be introduced, identifying the location of the main topic of the thesis in the wider picture. Finally, the structure of the work and its scope will be illustrated.

1.1 A brief history of Thermodynamics

1.1.1 Etymology

The word *thermodynamics* has its roots in the Greek ϑέρμη (thermē), meaning *heat*, and δύναμις (dynamis), meaning *power*. Its etymology explains that the word has historically been intended to designate the discipline that aims at explaining the relationship between the properties of bodies of being *hot* and *cold*, the natural phenomenon of balancing these *non-equilibria* by transferring *heat* and the ability to move objects, that is, to do *work*. Words referring to commonly used concepts in the previous sentence have been highlighted in order to underline the fact that they are not as immediate as it is usually thought to be: in fact, profound differences exist in their exact definition and in their positioning in the more general building of thermodynamics.

The date of birth of the word *thermodynamics* is controversial: the most widespread assumption, supported, for example, by Bolton and by Berger, is that it was coined by William Thomson, Lord Kelvin, in 1854; others (Cengel and Turner and Sebastian, for example) claim that the word was used by the same Lord Kelvin before that date, around 1850-1852, while a less plausible hypothesis is that the word had been coined even before, around 1840, as it is stated by American biophysicist Haynie.

1.1.2 Pioneering experiments on engines

The history of thermodynamics, however, dates back to previous centuries, when, on one hand, various experiments had been conducted in order to develop machines and devices of practical interest and, on the other hand, theories regarding the still unknown nature of heat and its modalities of propagation had been developed. Between these two approaches, it has undoubtedly been the first one that has contributed the most to the increase in the knowledge of thermodynamics and the laws that underlie it. In particular, the development of the discipline has been strongly linked to the improvements in the knowledge and realization of engines, aimed at satisfying elementary needs such as those related to transportation, cooking or early industrial applications.

One of the pioneers in the work on engines has been Otto von Guericke, who invented the first vacuum pump in 1650, in order to contradict the theory, developed by Aristotle, of *horror vacui*, stating that nature abhors vacuum and thus tends to fill every possible space. Its work was followed by Robert Hooke and Robert Boyle that, six years later, developed an air pump, exploiting it to study the relationship between the thermodynamic concepts of *pressure*, *temperature* and *volume*. Commercial realizations of thermodynamic devices were developed at the end of the century and in the following one: in 1698, Thomas Savery patented a device that, with the use of steam, could pump water from a lower to a higher level in order to solve water flooding problems in mines. Successively, around 1710, Thomas Newcomen developed Savery's invention by adding a piston and a cylinder to better exploit the condensation of steam in order to pump water. The following significant step came in the second half of the XVIII century when James Watt introduced two improvements on Newcomen's machine: the external condenser, that increased the efficiency of the machine, and rotary motion, replacing the alternating one of the previous devices, thus reducing the stresses on the components [Th02]. For a detailed history of the progress in steam engines and a characterization of the great minds that have made improvements in this field possible, we refer to the book written by Thurston, the first professor of Mechanical Engineering at Stevens Institute of Technology and, successively, Director of Sibley College at Cornell University [Th02].

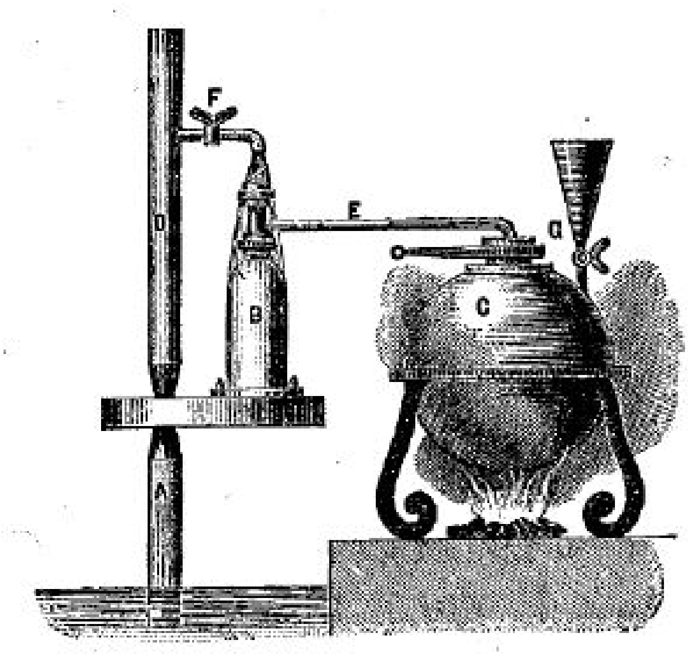

Fig. 1.1: Illustration of Thomas Savery's Engine of 1698.

1.1.3 Theoretical evolution of the subject

Even though practical devices, starting from the XVII Century, had been constantly improved, the theoretical knowledge regarding these processes was still anchored to the concepts of *phlogiston*, a mysterious substance released during combustion, and *caloric*, a fluid transferring from hotter bodies to colder bodies. Starting from the '700s, however, some scientists began to suggest that the concept of *heat* was related to the movement of particles inside a body, thus assuming that it was a form of energy.

The first systematic study of the causes underlying the previously cited phenomena was conducted by Sadi Carnot in 1824 in his *Réflexions sur la Puissance Motrice du Feu et sur les Machines Propres à Développer Cette Puissance (Reflections on the Motive Power of Fire).* As underlined in the preface by the editor, Robert H. Thurston, in the 1897 publication of Carnot's work, the young French scientist first introduced many ideas that are at the base of modern thermodynamics. The goal of his book, very simple from a mathematical standpoint, was, on one hand, to understand what was the maximum possible amount of work that could be extracted from a given quantity of heat and, secondly, if this amount of work was independent of the particular substance used in the machine. In answering these questions, he introduced the idea of a sort of *conservation principle* for the *motive power* of heat, thus giving an embryonic idea of the Second Law of Thermodynamics. Then, he introduced the concepts of *cycle*, stating that, in order to actually evaluate the effects of a process on the environment, the process itself must return to the initial point, and, finally, he introduced the notion of *reversibility*, associated to the perfection of the cycle [Car97]. The answers to the questions that had given birth to the book are the principles that underlie the statement of the Second Law of Thermodynamics in its traditional formulation: there is a maximum amount of work that can be extracted from a given quantity of heat, depending on the temperatures of the hot and cold source, and this amount is independent of the nature of the particular substance that is used in the machine.

From the investigation of a problem of 'engineering economics', thermodynamics has grown into a *body of doctrine of profound philosophical significance* [W52], even though, as Callen evidences in the preface to his monumental book, thermodynamics was the last branch of classical physics to be reformulated from a theoretical standpoint. This was caused by the fact that thermodynamics has always been strongly linked to macroscopic observations, as the deep nature of the phenomena that it investigated, related to the molecular theory of matter, remained rather obscure for a long time. This is also the reason why its original formulation had been developed in terms of cycles and transformations and only successively its structure was overhauled and restated in terms of state functions and equilibrium states, generating a simplification from a mathematical standpoint [Cal60]. However, we feel that still nowadays the theoretical structure of thermodynamics is very much debated and a general view shared by the majority of scholars does not exist.

What we claim is particularly important to understand from the development of the history of thermodynamics is that, with respect to other disciplines, such as electromagnetism, it has always been more tightly linked with practical applications and results. This fact, united with the absence of a universally shared theoretical structure and with the fact that the chemical and mechanical industries, two sectors that strongly contributed to technological and economic progress in the XX Century, heavily rely on thermodynamics, has caused a wider engineering presence in a field that should be a prerogative of physicists. With the present work, as engineers that are developing a topic with strongly theoretical features, we are thus just joining what has been a mainstream tendency in the history of the discipline.

1.2 Main problems in Thermodynamics

After having explained the etymology of the word *thermodynamics*, briefly illustrated its history and, most important, pointed out why we, as engineers, are working in this field, in the present section we aim at clarifying some basic concepts in thermodynamics, which will be helpful to understand the exact place of our thesis topic in the more general picture of the discipline.

Recalling Albert Einstein in [Ei79], we introduce the subject by stating:

> *"A theory is the more impressive the greater the simplicity of its premises, the more different kinds of things it relates, and the more extended its area of applicability. Therefore the deep impression that classical thermodynamics made upon me. It is the only physical theory of universal content which I am convinced will never be overthrown, within the framework of applicability of its basic concepts."*

The quote by the man who was probably the most prominent scientist of the XX Century introduces to the problem of the exact positioning of thermodynamics among the various *branches* of physics and its relationship with Mechanics. The issue is still very much controversial and our aim is just to sketch the prevalent positions and to point out the view through which we were introduced to the subject, which we found illuminating because of its clarity, sequentiality and rigour, even though it is not one of the mainstream ways to introduce students to thermodynamics. After that, we will briefly illustrate *kinematics* and *dynamics* as branches of thermodynamics: with the first word, we refer to the *description of the possible or allowed states of a system* [GB05], while, with the second one, we refer to a *causal description of the time evolution of a state* [GB05].

1.2.1 Epistemology

As it has been evidenced in the previous paragraph, thermodynamics was born from a theoretical point of view thanks to the work of Carnot at the beginning of the XIX Century. Many other scientists, most notably James Clerk Maxwell, Ludwig Boltzmann and Josiah Willard Gibbs, contributed to changing the interpretation of the discipline and, most important, linking it with microscopic phenomena. In particular, one of the most common views of thermodynamics, based on the ideas of Statistical Mechanics, is that of thermodynamics as a **statistical science**. It is based on the realization that the description of a macroscopic system cannot be conducted by describing the evolution of all its constituents because it would lead to a practically unsolvable system of equations and variables. As a consequence of this, thermodynamics provides a sort of effective 'synthesis' of the behaviour of a huge number of microscopic particles at the macroscopic level. As Glansdorff and Prigogine have stated, thermodynamics provides a *reduced description* or *simplified language* to describe macroscopic systems [GP71]. It can thus be said that the *statistical* interpretation arises because of the impossibility to describe the system deterministically starting from the equations of motion of the single particles.

However, statistical foundations of thermodynamics intended in this sense have given birth to more *radical* interpretations, such as those suggested by its merging with Information Theory. The information-theoretic interpretation of the discipline suggests that the thermodynamic state of a system does not depend only on the system itself, but also on the knowledge that the observer has: based on this knowledge, the observer may assign probability values to the possible states of the system and the actual configuration of the system (i.e. the actual probability values assigned to each state) is the one that maximizes a certain function, the *entropy*, as defined by Shannon ([Sha48; Ja57; Ka67]).

However, this interpretation would suggest that the actual state of the system is not univocally determined as the maximization process is constrained by the actual knowledge of the observer. This goes against common intuition as macroscopic phenomena are characterized by the same patterns independently of the particular observer. Moreover, the *statistical* interpretation of thermodynamics poses another significant problem: how is it possible that, in its macroscopic description, nature has a precise *direction* (irreversible phenomena exist), while the microscopic equations of motion are perfectly invariant under time-reversal? Many scientists have tried to give an explanation to this: one of the most common is the one that claims that macroscopic laws are valid *on average* and the fact that nature does not *go back* in time is due to the reason that it is highly improbable (it would happen probably one time during the life of the Universe), but not impossible.

The approach of the *Keenan School* is radically different. With this name, we refer to the School of Thermodynamics of the Massachusetts Institute of Technology (MIT), started with the pioneering work of Joseph Keenan, which brings an unconventional and often challenged view to the subject. The work of the founder has then been carried on by several authors during the second half of the XX Century; among these, there are Elias P. Gyftopoulos and Gian Paolo Beretta, whose work is frequently cited in the present thesis. According to the scholars of the Keenan School, Thermodynamics is a **non-statistical science** that applies to both micro and macro systems. An observer has no role in the interpretation of macroscopic phenomena and irreversibility is not a feature that depends on the particular *level of description* that is adopted (i.e., it exists at the macroscopic level, but does not exist at the microscopic level), but exists on all levels. The general results of this new approach to the discipline may be summarized in the following main points:

- Thermodynamics is an extension of Mechanics, as the states of Mechanics are *zero-entropy states* for thermodynamics; this is witnessed by the following quotation from Gyftopoulos and Beretta:

  > *"Were we to assume that a system is subject only to the laws of mechanics, we would conclude that all the energy of the system in excess of the ground-state energy can be used to lift a weight. [...] But this conclusion is not consistent with all experimental results. [...] To account for these experiences, the laws of thermodynamics entail a greater variety of states than contemplated by the laws of mechanics.*
  >
  > *The contrast between the values of the adiabatic availability in mechanics and in general is yet another way to present the fun-*

> *damental differences between the domain of validity of the laws of mechanics and that of the laws of thermodynamics, which includes the domain of mechanics as a special and limiting case." [GB05, p. 81]*

- the consideration is valid also on the microscopic level, as more general quantum states than those considered by Quantum Mechanics are said to exist; this is related to the fact that the *existence* of entropy implies a new *paradigm*[1], as illustrated in the following quotation:

  > *"The possibility for entropy to be created by irreversibility reflects a physical phenomenon that is sharply distinct from the great conservation principles that underlie the description of physical phenomena in mechanics. It brings forth the need to consider not only properties that are conserved, such as energy, mass, momentum, and electric charge, but also properties that may be spontaneously created, such as entropy.*
  >
  > *In fact, even the requirement of entropy conservation in reversible processes of isolated systems introduces a radical departure from the description of physical phenomena in mechanics, a departure that would persist even if no process in nature were irreversible. The reason is that this requirement brings forth the need to describe not only states with zero entropy, such as the states encountered in mechanics, but also states with various nonzero values of entropy." [GB05, p. 106]*

- as irreversibility is a *built-in* characteristic of nature, the microscopic equations of evolution are not complete because they are time-reversible; a new equation of motion, that has Schrodinger's equation as a particular case, is thus needed and has indeed been proposed [Bere81].

In the following **Fig.** 1.2, a typical *energy-entropy graph* showing all the states of thermodynamics is illustrated. It may be understood in which sense Thermodynamics is considered as an extension of Mechanics: the states of Mechanics are

---

[1] The word *paradigm*, in this context, has the same meaning that Kuhn gave it in his 1962 masterpiece *The Structure of Scientific Revolutions*, where he challenged the consolidated view according to which science proceeds by *accumulation*, replacing it with the idea that periods of *normal science* are alternated with periods of *revolutionary science*, characterized by the presence of new *paradigms*. These are "*the set of practices that define a scientific discipline at any particular period of time*" [Ku62].

zero-entropy states in which all the energy of the state in excess of the state of minimum energy may be 'used to lift a weight', i.e., may be converted to work. This would not happen starting from an *isolated, non zero-entropy state* because reaching the state of minimum energy would be impossible without decreasing the value of the entropy. It may thus be understood that evolution in Mechanics takes place only vertically on the line of *zero-entropy states*, while evolution in thermodynamics takes place in the whole graph; moreover, for each value of the energy, Thermodynamics considers one stable equilibrium state and infinite non-equilibrium states, while Mechanics is characterized by the existence of only one state for each value of the energy and this state is a non-equilibrium state, unless it is the one of lowest energy.

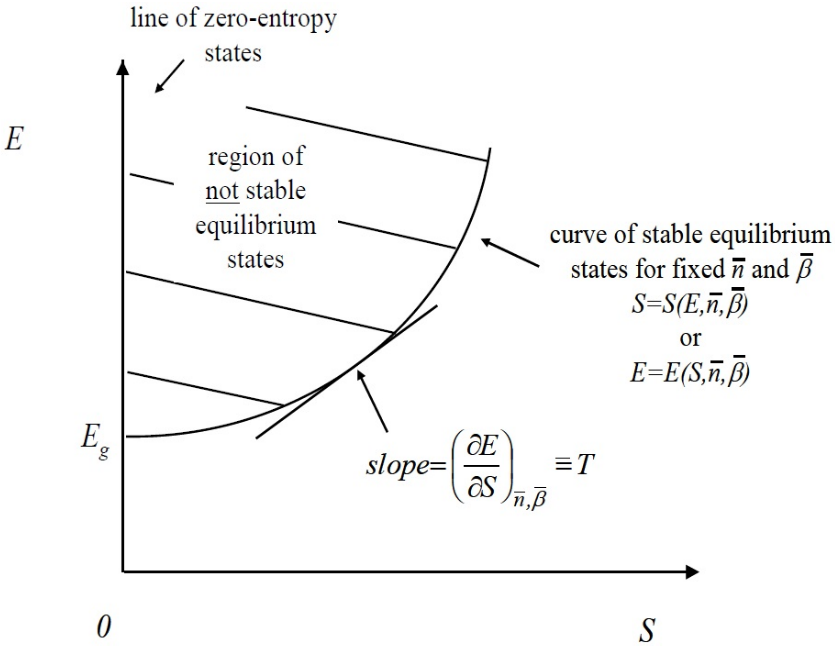

Fig. 1.2: Energy versus entropy graph: states on the continuous line are thermodynamic stable equilibrium states, while states inside the curve are non-equilibrium states.

1.2.2 Kinematics

As it has been stated by Gyftopoulos and Beretta, **kinematics** is the branch of physics that has the *description of the states of a system* as the object of its study [GB05]. This applies in general to all the disciplines included in the broad range of *physics*, thus also to thermodynamics. Fundamental in the previous sentence is the concept of *state*: among the different definitions, we chose the concise and qualitative one given by the same authors, who define the *state* as the *set that specifies everything about a system at one instant of time* [GB05]. For a wider and more rigorous definition of *state*, we refer to Beretta and Zanchini who, in presenting a logical scheme that rigorously defines entropy, also give careful operative definitions of many of the basic concepts needed to face the study of thermodynamics; in addition to *state*, also *system, property, environment, process, isolated system* and other concepts are given [BZ].

The key aspect thus becomes the choice of all the variables that may allow the specification of all the information about a system at a precise instant of time. The choice of the variables depends on the particular *level of description* that is chosen for the system under consideration. This is related to the fact that different systems may be described with different degrees of *coarseness*. The various levels of description have a hierarchical structure and are characterized by different conservation laws: in particular, a level L_1 is called *deeper* than a level L_2 if all the *constituents* of L_1 are conserved in the physical process under examination, while this is not true for the *constituents* of L_2 [BZ][2]. The switch from a *lower* level of description to a *higher* level of description is done through **coarse-graining** procedures, which may be seen as a sort of *averaging over microscopic states* [Ö05]. *Coarse-graining* procedures influence both the kinematics and the dynamics of a physical system because, in the passage between different levels of description, both the variables used to describe the system and its evolution features change. Regarding these procedures, an interesting epistemological question, currently debated by scholars, arises: is it true that, by the successive application of *coarse-graining* schemes, the researcher distances himself more and more from the actual *functioning of nature*, that is from *natural laws*, and widens the space allocated to the *modelling* part, that he himself provides?

The topic of the **choice of variables** is thus particularly significant. In *Equilibrium Thermodynamics*, the different frameworks used to build the theoretical construct have all agreed upon the fact that, under some restrictive hypotheses, the *intensive* thermodynamic stable equilibrium states may be determined through two intensive, independent variables. That is, all the other intensive variables of the system are univocally determined once two of them have been chosen. A third variable is needed to scale the dimensions of the system. In general, however, if particular hypotheses on the system are not assumed, the number of independent

---

[2] *Constituents* may be seen as the *elementary building blocks* of matter [BZ]

properties in a *basis* used to describe the state is infinite and, thus, the choice of the variables is much harder [GB05]. Particularly hard is thus the choice of the variables for non-equilibrium systems.

It is not in our intention to review all the solutions that have been adopted in the past decades in the study of non-equilibrium phenomena. However, we would like to cite some of the most significant ones in order to offer to the reader ideas for possible answers to the problem and to highlight the spirit that underlies these solutions. Classical Non-Equilibrium Thermodynamics, which will be extendedly exposed in the second chapter of the work, considers the same variables that characterize equilibrium thermodynamics. The additional variables that may be considered for non-equilibrium systems may be of several different types; among these, there are time derivatives of equilibrium variables, spatial fluxes and internal variables related to the structure of the system. The choice may depend, on one hand, on the time scales that are considered and, on the other hand, on the particular interests of the observer [Jo13]. As for the time scale, for example, the interest is usually on degrees of freedom of the system whose relaxation times are comparable to the rates of external perturbations. Indeed, *slower* degrees of freedom will be considered *frozen*, while *faster* degrees of freedom will be assumed as instantaneously reaching equilibrium ([JR11; LJC92; LVR01]). On the other hand, for example, if the interest of the observer is on steady states, rates of change would be of little use and fluxes would be much more appropriate variables.

Regarding this last aspect, the class of non-equilibrium theories that considers the fluxes as independent variables for the definition of a state is called *Extended Irreversible Thermodynamics*. Indeed, EIT adds the fluxes of the conserved densities of Classical Non-equilibrium Thermodynamics (CNET), i.e. mass and energy flux, to the set of basic independent variables in order to better describe high-frequency or short-wavelength phenomena (hyperbolic equations of evolution, with finite propagation speeds, are obtained in this way, oppositely to parabolic equations obtained through classical constitutive laws). CNET equations are obtained in the limit of slow phenomena. Further generalization and wider domains of validity are then obtained by adding higher-order fluxes (see, e.g., [JCL88; LVCJ98]).

1.2.3 Dynamics

Dynamics is defined generally as the branch of physics that has "the causes of motion and the analysis of their effects" as the objects of study [GB05]. *Motion* is intended as the change in state of a system in time and is due to the presence of external or internal forces [GB05]. In order to describe the dynamical evolution of a system, *equations of motion* are needed: for example, in Classical Mechanics, Newton's Law, relating the force acting on a body and its acceleration, describes the dynamical features of the system, while in Quantum Mechanics, the time evolution is depicted by the Schrödinger's Equation. In thermodynamics, such a law is still lacking and the topic is the subject of research [GB05]. According to Gyftopoulos and Beretta, whose view we would like to adopt in the present thesis, even though a precise equation of motion has not been found yet, general features that characterize this equation have been discovered and are represented by the First and Second Law of Thermodynamics. As a consequence of this, if an equation were discovered which describes the motion of the system in state space (without discussing if and to what extent could the equation be used in practical terms to calculate motions of complex systems), the two Laws could be straightforwardly derived from it as theorems. Many attempts have been made towards this goal, but no unquestionable result has been obtained yet.

Among these attempts, one of the most recent and, probably, most ambitious is the one proposed by Adrian Bejan, professor of Mechanical Engineering at Duke University [CLaw]. He has proposed the **Constructal Law of Evolution** as an "additional self-standing law" to the already existing Laws of Thermodynamics. He claim that the *Constructal Law* is a general law of physics that applies to all flow systems, both animate and inanimate and is stated as follows:

> *"For a finite-size flow system to persist in time (to live), its configuration must evolve in such a way that provides greater and greater access to the currents that flow through it" [CLaw]*

Bejan claims that neither of the two Laws of Thermodynamics take into account *design* or *optimization* phenomena, thus the *Constructal Law* is needed to consider the time evolution of a thermodynamic system (even though its application is asserted to be much more general). The narrower optimality principles and patterns that are found both in thermodynamics and in other fields of human experience, such as biology, technology or society, are seen as particular realization of the more general law.

Among the other principles that have been proposed to satisfy the need of an equation of motion for a thermodynamic system, one of the most studied is the

Maximum Entropy Production Principle (MEPP). The principle has been proposed by several authors working independently and may be stated as follows:

> *"By this principle, a nonequilibrium system develops so as to maximize its entropy production under present constraint." [MS06, p. 3]*

The reader must be warned about the fact that great confusion upon the hypotheses at the base of this principle and upon its actual statement and meaning is still present nowadays in the scientific community: this general expression is often used in different contexts where different constraints and different optimized variables are present. One of the purposes of the present work is the one of rationalizing part of the enormous amount of work that has been done on this topic in order to understand the actual significance of the principle. Moreover, the MEPP must not be confused nor assumed to be in contrast with another principle that has been stated to rule the evolution of non-equilibrium thermodynamic systems: Prigogine's Minimum Entropy Production Principle [GP54]. Indeed, as it will be illustrated during the course of the work, Prigogine's principle is, on one hand, a global principle related to the asymptotic time evolution of the entropy production and, on the other hand, it is valid under more restrictive hypotheses. The Maximum Entropy Production Principle is instead a principle related to the evolution of a system at a fixed instant of time and may be considered much more general: indeed, Prigogine's principle may be considered as a consequence of MEPP under particular restrictions. The difference will be clarified during the course of the work.

Closely related to the principle of Maximum Entropy Production is the Steepest Entropy Ascent (SEA) dynamical model proposed by Beretta, initially in a quantum framework and successively adapted to meso- and macroscopic systems [Bere13]. Also Grmela and Öttinger have proposed a GENERIC (General Equation for Non-Equilibrium Reversible Irreversible Coupling), that is, an equation that describes the dynamical evolution of non-equilibrium systems at various scales [GÖ97]. The study of this dynamical model (and its relationship with other *maximum-entropy-producing* principles) is included in the goals of the thesis. Many other models and theories of dynamical evolution for a thermodynamic system have been developed and might be cited, but we will omit them, as the purpose of the present paragraph is just sketching the borders of the topics that will be developed in the thesis work, that is, dynamical models for thermodynamic systems, and not reviewing the scientific literature in the field.

1.3 Non-Equilibrium Thermodynamics

1.3.1 Non-equilibrium Thermodynamics and Engineering

As it was said in the previous paragraph, the general focus of the present work will be on the dynamical modelling of thermodynamic systems. The subject is a frontier topic in the field of physics and the reader is perfectly legitimated to wonder the reason why the present topic is debated in an engineering thesis. A partial answer resides in what was exposed in the paragraph regarding the history of thermodynamics: for various reasons, developments in this discipline, much more than in other branches of physics, have been due to the work of engineers, even from a *theoretical* standpoint. However, a much deeper reason is indeed present and has a typical engineering feature: usefulness. The research for the patterns or, more ambitiously, the equation of dynamical evolution of thermodynamic systems is not only a challenging conceptual problem whose answer would be a milestone for Theoretical Physics, but it would also be an extremely advantageous device to better comprehend, govern and exploit practical phenomena. Among these, there are thermodynamic phenomena of relaxation towards equilibrium from non-equilibrium states, which characterize an enormous number of engineering systems of practical use. These phenomena are indeed the object of study of a wide branch of physics, called **Non-Equilibrium Thermodynamics** or **Thermodynamics of Irreversible Processes**, as processes of evolution from non-equilibrium states are usually characterized by entropy production. *Non-equilibrium thermodynamics* differs from *equilibrium thermodynamics* or *thermostatics* because of the need of modelling systems that change in time. Indeed, the latter focuses on systems having a fixed state, independent of time, or systems undergoing very slow (*quasi-static*) transformations [Cal60]. As it has been stated by de Groot and Mazur, "the field of non-equilibrium thermodynamics provides us with a general framework for the macroscopic description of irreversible processes" [dGM84]. However, it must be evidenced that most of the consolidated work in this field has been done on the linear approximation for near-equilibrium states. Further details on the meaning of *linear* approximation will be given in Chapter 2. It may thus be said that far-equilibrium states and nonlinear behaviour have been excluded from traditional treatise, even though they "are key to many technological applications" [Ö05]; as a consequence of this, there is the need to have a "unified approach to nonlinear irreversible thermodynamics, the existence, necessity, and usefulness of which has currently not been fully realized by large fractions of the science and engineering communities" [Ö05]. On the other hand, from a more theoretical standpoint, the same author claims that a unified approach would provide a deeper understanding of *beyond-equilibrium* systems by offering a general framework for different empirical applications, consistency with experimental results and recipes for going from more detailed to less detailed levels of description [Ö05].

1.3.2 Potential applications

As it has been hinted at in the previous subsection, the spectrum of *non-equilibrium situations* is extremely wide and very broad is also the range of practical situations modelled through Classical Non-Equilibrium Thermodynamics and Kinetic Theory. Moreover, dynamical principles such as the Maximum Entropy Production Principle have been adapted and used in a variety of disciplines, distant from *pure* thermodynamics. Below is a list of the practical situations related to the topics that have been cited: some of the them will be further analysed in the following Chapters.

- A typical non-equilibrium process is the relaxation towards a stable equilibrium state of a system that has undergone a chemical reaction of oxidation, as it happens in the vast majority of power system used nowadays by man: a deeper knowledge of this process of evolution towards equilibrium may thus allow to reduce the entropy generated by irreversibility during the process and to improve the efficiency of thermal machines. In **Fig.** 1.3, an illustration of different processes of evolution towards equilibrium from the initial non-equilibrium condition are shown in an enthalpy-entropy graph: fuel-cell oxidation generates less entropy than flame combustion, thus allowing the extraction of a greater amount of work.

- Dynamical evolution models for far-equilibrium situations are also needed for the description of natural phenomena for which Navier-Stokes description fails: in these situations, Kinetic Theory and, in particular, the Boltzmann Equation describe the evolution of the system, even though, in order to simplify computations, *Model Equations*, usually based on strong assumptions, are used in calculation programmes. A dynamical evolution equation for a thermodynamic system might improve the description of relaxation towards equilibrium in these contexts: one of the purposes of Steepest-Entropy-Ascent dynamics is exactly this one, as it will be shown in the last part of the

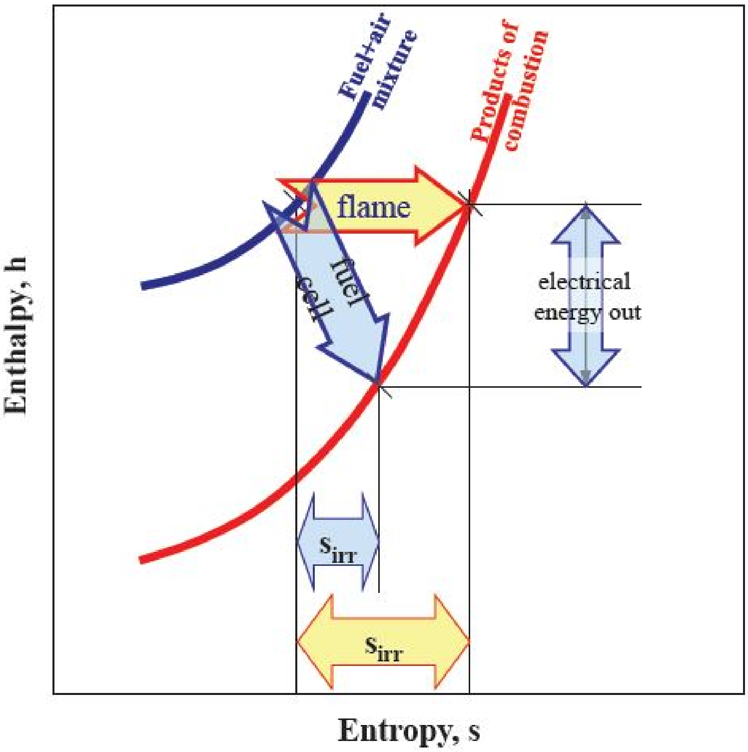

Fig. 1.3: Different evolutions towards equilibrium in an enthalpy-entropy graph from the same non-equilibrium initial condition: oxidation in a fuel cell compared with flame combustion.

thesis. As illustrated in **Fig.** 1.4, various situations of practical interest are characterized by the failure of Navier-Stokes modelling: this happens when the internal scale of the fluid λ approaches the characteristic scale of the flow L, that is, when the *Knudsen number* $Kn = \lambda/L$ is significantly different from zero. Indeed, when the Knudsen number is close to zero, transport is collision-dominated and Navier-Stokes equations are adopted; as the fluid *internal scale* approaches the *characteristic lengthscale* of the flow, Navier-Stokes equations remain valid only in the bulk of the flow and, eventually, as the Knudsen number gets even higher, they have to be replaced by ballistic motion and kinetic models. Recent work on a possible application was done by Diab and Lakkis, who studied the behaviour of a rarefied gas film under an oscillating micro-cantilever RF (Radio Frequency) switch [DL12].

- Typical applications of non-equilibrium thermodynamics in its classical linear approximation is the modelling of thermoelectric and thermodiffusive

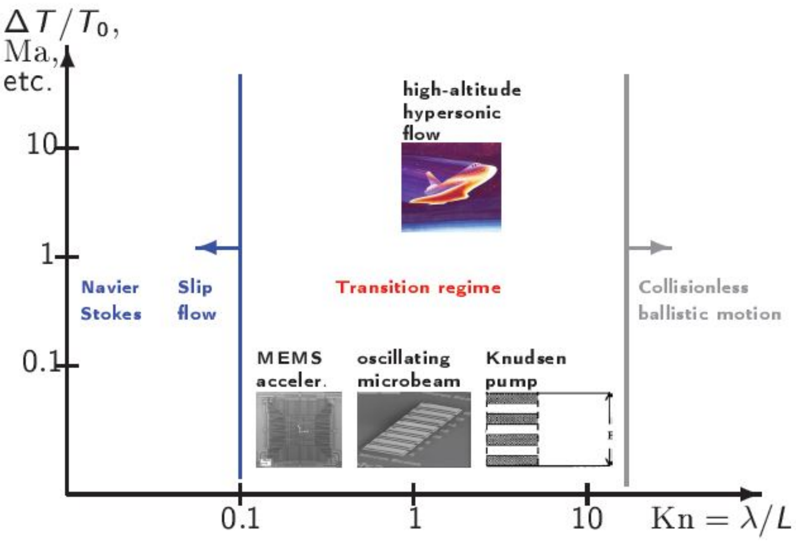

Fig. 1.4: Classification of physical phenomena according to the Mach number and the Knudsen number, with domains of validity of the different equations.

processes: further details will be given in the exposition of the Classical Non-Equilibrium theory in Chapter 2.

- Maximum-entropy production evolution principles are applied not only in the strictly thermodynamic field, but also in many other scientific problems and even in the social studies sphere. Citing a few examples, maximum-entropy production models have been applied to climate studies by Paltridge, who showed that the Earth's atmosphere is seemingly characterized by maximum dissipation processes related to horizontal energy flows [Pa79]; Juretic and Zupanovic have modelled steady-state bacterial photosynthesis concluding that "photosynthetic proton pumps operate close to the maximum-entropy production mode" [JZ03]; Shizawa and Zbib developed a *thermodynamic theory of elastoplasticity*, introducing a dislocation density tensor on the basis of the principle of maximal entropy production rate [SZ99].

1.4 Structure and scope of the Thesis

After having defined the general topic that frames the thesis, the illustration of its scopes and its structure may be more clear.

1.4.1 Goals of the Work

The main goals of the present work are the following:

- first of all, some Maximum-Entropy-Production dynamical models for Non-Equilibrium Thermodynamics, elaborated during the course of the last decades, will be presented and systematically reviewed, introducing then the Steepest Entropy Ascent dynamical model, recently adapted by Beretta to meso- and macroscopic systems [Bere13];

- the Steepest Entropy Ascent (SEA) dynamical model, which aims at *geometrizing* the dynamic evolution of a thermodynamic system, is then compared with one of the most renowned presentations of *metriplectic dynamics*, aimed at a *geometrization* of Non-Equilibrium Thermodynamics, that is, the GENERIC (General Equation for Non-Equilibrium Reversible-Irreversible Coupling) formalism, developed, among others, by Grmela and Öttinger in the last two decades of the XX Century [GÖ97]: a comparison between the two models, both motivated by the search for Non-Equilibrium Thermodynamics formulations that are fully compatible with the Second Law, with the identification of similar and different features, will be conducted;

- finally, in order to have a clearer comprehension of the two dynamic models, their compatibility with Classical Non-Equilibrium Thermodynamics will be evaluated and their realization in Kinetic Theory will be illustrated, with a numerical example of the use of a Steepest Entropy Ascent kinetic model for the Boltzmann Equation to reproduce the relaxation from a highly non-equilibrium state, as recently studied by Beretta and Hadjiconstantinou [BH13].

1.4.2 Organization of the Work

The structure of the thesis is listed as follows.

- Chapter 2 will be devoted to the explanation of the classical consolidated theories used to test the successively illustrated models; firstly, Classical Non-Equilibrium Thermodynamics (CNET), formalized by de Groot and Mazur in their best-selling physics book [dGM84], and linear approximations of non-equilibrium phenomena, with particular reference to Onsager's theory, will be explained; then, a general introduction to Kinetic Theory and to the Boltzmann Equation will be given.

- Chapter 3 will be dedicated to the illustration of some of the Maximum-Entropy Production Principles (MEPP) that have been developed in the history of thermodynamics, with particular reference to the theory developed by Ziegler [Zi57], and to the explanation of their relationship with CNET; the Steepest Entropy Ascent (SEA) model will then be introduced; in the second part of the chapter, theoretical studies conducted by Bataille, Edelen, and Kestin [BEK79] will be illustrated, paying particular attention to the introduction of the idea of the *two-generator* formalism, which is also used in the GENERIC structure.

- Chapter 4 will be devoted to the illustration of the Classical Mechanics equations in the Differential Geometry framework; indeed, as it has been exposed by Marsden and Ratiu [MR03], Differential Geometry is the most natural setting of the classical Lagrange and Hamilton equations and some peculiar structures used in this modelling are then at the base of the GENERIC formalism developed, among others, by Grmela and Öttinger [GÖ97], which is successively illustrated in the chapter. The second part of the Chapter constitutes the most innovative output of our thesis work: Differential Geometry has been considered to be the most appropriate setting in which a comparison of the Steepest Entropy Ascent dynamical model and the GENERIC formalism may be conducted. Because of this, SEA has been "translated" into Differential Geometry terms and, following this reformulation, it becomes clear that the GENERIC formalism adopts an *entropy-gradient* dynamics, such as the one proposed in SEA, with the difference that it is built in a more structured environment. Considerations on the relationship among SEA, GENERIC and the models explained in the previous chapters will then be illustrated.

- Chapter 5 will be devoted to the illustration of the applications of the two models. First of all, the *interpretation* of the equations of Hydrodynamics and Magnetohydrodynamics (MHD) in the GENERIC formalism will be shown. Then, the realization of SEA and GENERIC in Kinetic Theory will be illustrated. The different terms of the Boltzmann equation may be associated to the different *building blocks* of the two theories. In particular, the metrics that are used will be discussed. Finally, a numerical example of Steepest Entropy Ascent relaxation from a highly non-equilibrium state, recently developed by Beretta and Hadjiconstantinou, will be illustrated and comments will be made on the effectiveness of the model equation thus obtained, compared with those that are normally used [BH13].

- Chapter 6 will be dedicated to the conclusions and future developments and it will be followed by two Appendices. The first one contains a *Compendium* of Differential Geometry; we indeed felt that this Appendix was necessary in order to introduce the reader to the basic concepts and formalism used to reformulate SEA dynamics as they are not usually found in a standard engineering Master of Science course. The second Appendix is much shorter and aims at explaining the mathematics underlying Edelen's theory, exposed in Chapter 3, related to the so-called Helmholtz-Hodge decomposition.

We warn the reader used to a typical engineering treatise that the second part of the thesis may, at first sight, be found particularly heavy and hard to understand, especially because of the Differential Geometry mathematical formalism that is usually not taught in regular curricula. Because of this, we made a strong effort of always trying to anticipate through physical or geometric intuition the general meaning of the mathematical expressions that we would have successively introduced. However, we felt that this formalism was strictly necessary: if the reader deeply comprehends it, he will clearly understand that it is the most natural and limpid setting of the dynamical modelling of thermodynamic systems.

References

[BEK79] J. Bataille, D. G. B. Edelen, and J. Kestin. “On the structuring of thermodynamic fluxes: a direct implementation of the dissipation inequality”. In: *International Journal of Engineering Science* 17.5 (1979), pp. 563–572 (cit. on p. 19).

[Bere13] G. P. Beretta. “Steepest-Entropy-Ascent and Maximal-Entropy-Production Dynamical Models of Irreversible Relaxation to Stable Equilibrium from Any Non-Equilibrium State. Unified Treatment for Six Non-Equilibrium Frameworks”. In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013*. Ed. by M. Pilotelli and G. P. Beretta. Brescia: Cartolibreria Snoopy, 2013, pp. 100–109 (cit. on pp. 13, 18).

[Bere81] G. P. Beretta. “On the general equation of motion of quantum thermodynamics and the distinction between quantal and nonquantal uncertainties”. thesis. 1981 (cit. on p. 8).

[Berg06] S. Berger. *A Companion to Nineteenth-century Europe*. Blackwell, 2006 (cit. on p. 2).

[BH13] G. P. Beretta and N. G. Hadjiconstantinou. “Steepest Entropy Ascent Models of the Boltzmann Equation. Comparisons with Hard-Sphere Dynamics and Relaxation-Time Models for Homogeneous Relaxation from Highly Non-Equilibrium States”. In: *Proceedings of the ASME 2013 International Mechanical Engineering Congress and Exposition, IMECE 2013*. 2013 (cit. on pp. 18, 20).

[Bo00] J. Bolton. *Classical Physics of Matter*. CRC Press, 2000 (cit. on p. 2).

[BZ] G. P. Beretta and E. Zanchini. “Rigorous and General Definition of Thermodynamic Entropy”. In: *Thermodynamics*. Ed. by M. Tadashi. InTech, 2011, pp. 23–50 (cit. on p. 10).

E. Zanchini and G. P. Beretta. “Rigorous operational definition of entropy not based on the concepts of heat and of empirical temperature”. In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013*. Ed. by M. Pilotelli and G. P. Beretta. Brescia: Cartolibreria Snoopy, June 2013, pp. 152–159. Cit. on p. 10.

[Cal60] H. B. Callen. *Thermodynamics: an Introduction to the Physical Theories of Equilibrium Thermostatics and Irreversible Thermodynamics*. New York: John Wiley & Sons, 1960 (cit. on pp. 4, 14).

[Car97] N. L. S. Carnot. *Réflexions sur la Puissance Motrice du Feu et sur les Machines Propres à Développer Cette Puissance (Reflections on the Motive Power of Fire)*. Ed. by R. H. Thurston. 2nd ed. New York: John Wiley & Sons, 1897 (cit. on pp. 4, 6).

[CLaw] A. Bejan. "Constructal-theory network of conducting paths for cooling a heat generating volume". In: *International Journal of Heat and Mass Transfer* 40.4 (1997), pp. 799–816 (cit. on p. 12).

A. Bejan and S. Lorente. "The constructal law and the thermodynamics of flow systems with configuration". In: *International Journal of Heat and Mass Transfer* 47.14-16 (2004), pp. 3203–3214 .

A. Bejan and S. Lorente. "Constructal law of design and evolution: Physics, biology, technology, and society". In: *Journal of Applied Physics* 113.15 (2013). Cit. on p. 12.

[CT04] Y. A. Cengel and R. H. Turner. *Fundamentals of Thermal-Fluid Sciences*. McGraw-Hill, 2004 (cit. on p. 2).

[dGM84] S. R. de Groot and P. Mazur. *Non-equilibrium Thermodynamics*. 2nd ed. (1st ed., North-Holland, 1962). Mineola (New York): Dover Publications, 1984 (cit. on pp. 14, 18).

[DL12] N. A. Diab and I. A. Lakkis. "DSMC Simulations of Squeeze Film Under a Harmonically Oscillating Micro RF Switch With Large Tip Displacements". In: *ASME 2012 International Mechanical Engineering Congress and Exposition*. Houston, 2012, pp. 187–197 (cit. on p. 16).

[Ei79] A. Einstein. *Autobiographical notes. A centennial edition*. Ed. by P. A. Schilpp. Chicago, Illinois: Open Court Publishing Company, 1979 (cit. on p. 5).

[GB05] E. P. Gyftopoulos and G. P. Beretta. *Thermodynamics: Foundations and Applications*. 2nd ed. (1st ed., Macmillan, 1991). Mineola (New York): Dover Publications, 2005 (cit. on pp. 6–12).

[GÖ97] M. Grmela and H. C. Öttinger. “Dynamics and thermodynamics of complex fluids. I. Development of a general formalism”. In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 56.6 (1997), pp. 6620–6632 (cit. on pp. 13, 18–19).

[GP54] P. Glansdorff and I. Prigogine. “Sur les proprietes differentielles de la production d’entropie”. In: *Physica* 20 (1954), pp. 773–780 (cit. on p. 13).

[GP71] P. Glansdorff and I. Prigogine. *Thermodynamic Theory of Structure, Stability and Fluctuations*. New York: Wiley-Interscience, 1971 (cit. on p. 6).

[Hay01] D. T. Haynie. *Biological Thermodynamics*. 2nd ed. Cambridge: Cambridge University Press, 2001 (cit. on p. 2).

[Ja57] E. T. Jaynes. “Information Theory and Statistical Mechanics”. In: *Physical Review* 106.4 (1957), pp. 620–630 (cit. on p. 6).

[JCL88] D. Jou, J. Casas-Vázquez, and G. Lebon. “Extended irreversible thermodynamics”. In: *Reports on Progress in Physics* 51 (1988), pp. 1105–1179 (cit. on p. 11).

[Jo13] D. Jou. “Entropy, entropy flux, temperature, and second law in extended irreversible thermodynamics”. In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013*. Ed. by M. Pilotelli and G. P. Beretta. Brescia: Cartolibreria Snoopy, 2013, pp. 211–216 (cit. on p. 11).

[JR11] D. Jou and L. Restuccia. “Mesoscopic transport equations and contemporary thermodynamics”. In: *Contemporary Physics* 52 (2011), pp. 465–474 (cit. on p. 11).

[JZ03] D. Juretic and P. Zupanovic. “Photosynthetic models with maximum entropy production in irrevrsible charge transfer steps”. In: *Computational Biology and Chemistry* 27 (2003), pp. 541–553 (cit. on p. 17).

[Ka67] A. Katz. *Principles of Statistical Mechanics. The Information Theory Approach*. San Francisco: Freeman, 1967 (cit. on p. 6).

[Ku62] T. S. Kuhn. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press, 1962 (cit. on p. 8).

[LJC92] G. Lebon, D. Jou, and J. Casas-Vázquez. “Questions and answers about a thermo dynamic theory of the third type”. In: *Contemporary Physics* 33 (1992), pp. 41–51 (cit. on p. 11).

[LVCJ98] R. Luzzi, Á. R. Vasconcellos, J. Casas-Vázquez, and D. Jou. “On the selection of the state space in nonequilibrium thermodynamics”. In: *Physica A* 248 (1998), pp. 111–137 (cit. on p. 11).

[LVR01] R. Luzzi, Á. R. Vasconcellos, and J. G. d. P. Ramos. *Statistical Foundations of Irreversible Thermodynamics*. Leipzig: Teubner, 2001 (cit. on p. 11).

[MR03] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry. A Basic Exposition of Classical Mechanical Systems*. 2nd ed. (1st ed., 1994). New York: Springer-Verlag, 2003 (cit. on p. 19).

[MS06] L. M. Martyushev and V. D. Seleznev. “Maximum entropy production principle in physics, chemistry and biology”. In: *Physics Reports* 426.1 (2006), pp. 1–45 (cit. on p. 13).

[Ö05] H. C. Öttinger. *Beyond Equilibrium Thermodynamics*. 1st ed. Hoboken (New Jersey): Wiley, 2005 (cit. on pp. 10, 14–15).

[Pa79] G. W. Paltridge. “Climate and thermodynamic systems of maximum dissipation”. In: *Nature* 279 (1979), pp. 630–631 (cit. on p. 17).

[Seb01] A. Sebastian. *A Dictionary of the History of Science*. Taylor and Francis, 2001 (cit. on p. 2).

[Sha48] C. E. Shannon. “A Mathematical Theory of Communication”. In: *The Bell System Technical Journal* 27 (1948), pp. 379–423, 623–656 (cit. on p. 6).

[SZ99] K. Shizawa and H. M. Zbib. “A Thermodynamical Theory of Plastic Spin and Internal Stress with Dislocation Density Tensor”. In: *Journal of Engineering Materials and Technology* 121.2 (1999), pp. 247–253 (cit. on p. 17).

[Th02] R. H. Thurston. *A History of the Growth of the Steam Engine*. New York: D. Appleton and Company, 1902 (cit. on p. 3).

[W52] L. P. Wheeler. *Josiah Willard Gibbs*. New Haven: Yale University Press, 1952 (cit. on p. 4).

[Zi57] H. Ziegler. “Thermodynamik und rheologische Probleme”. In: *Ingenieur-Archiv* 25.1 (1957), pp. 58–70 (cit. on p. 19).

2

Validating the model

The purpose of the present chapter is to briefly present the two ‘theories’ that will be used to ‘validate’ the dynamical models illustrated in the following chapters, with particular regard to Steepest Entropy Ascent and the GENERIC formalism. The two ‘theories’ are **Classical Non-Equilibrium Thermodynamics** (CNET) in its linear approximation and **Kinetic Theory** (KT). As these two ‘theories’ are used to model very different types of phenomena, we have to specify the sense in which we aim at validating the dynamical models through them.

CNET in its linear approximation, which will be described in its main features in the following section, has been proved to be effective in modelling so-called *near equilibrium* situation, that is, situations in which the system under observation is situated close to thermodynamic equilibrium. Its main features, which are the linear relationship between forces and fluxes and the symmetry of the matrix of coefficients that relates the first to the second ones, must then be reproduced by all dynamical theories, also those that aim at modelling far-equilibrium situations, when equilibrium is approached. Thus, in the following chapters it will be shown how these features may or may not naturally emerge in the context of SEA and GENERIC for near-equilibrium approximations.

On the other hand, Kinetic Theory models systems characterized by a high number of particles in random motion, allowing the computation of macroscopic properties from a statistical microscopic description. On the basis of the more or less restrictive hypotheses that may be stated to 'build' an equation governing the time evolution of the system, different differential equations for the main variable, that is the distribution function f, may be obtained. Mathematical challenges related to the resolution of these equations are linked to the number and type of hypotheses that were made to obtain them. One of the most used among these equations is the Boltzmann equation. The procedure to obtain it, its physical meaning and the possible ways to solve it will be shown in the second part of the chapter. The equation will then be 'used' to understand how it is framed in the SEA and GENERIC dynamical models, that is, how the different terms of the equation are associated to the different 'blocks' of the two models and how the equation is shown to emerge naturally in the two settings. Finally, a numerical example of relaxation from a highly non-equilibrium state, using a Steepest-Entropy-Ascent model of the Boltzmann equation will be shown, in order to understand the effectiveness of the dynamic theory in modelling actual situations.

The premise for the present Chapter was necessary in order to correctly frame the following paragraphs in the more general picture of the thesis and in order to understand that the exposition of CNET and KT is not an end in itself, but is functional to the following parts of the work. In these paragraphs, we have made the deliberate choice of favouring a more qualitative description of the topics, rather than reporting numerous equations, in order to convey the main ideas to the reader more rapidly.

2.1 Non-Equilibrium Thermodynamics

As it has been stated in the introduction, *Non-Equilibrium Thermodynamics* is the branch of physics that studies systems that are not in a condition of thermodynamic equilibrium, that is, systems that cannot be described through the relationships peculiar to thermostatics. The discipline is also called *thermodynamics of irreversible processes*, as the processes of evolution towards equilibrium from non-equilibrium situations are (almost) always characterized by the generation of entropy. Non-equilibrium thermodynamics may be considered an extension of thermostatics towards the *continuum* disciplines, such as fluid dynamics and electromagnetism. Indeed, while thermostatics considers one value of the state variables for the whole system, on the other hand, non-equilibrium thermodynamics considers state variables as fields that are defined for every point in space and time [dGM84].

2.1.1 The hypothesis of local equilibrium

The fundamental assumption that resides at the base of Classical Non-Equilibrium Thermodynamics is the hypothesis of *local equilibrium*, also known as hypothesis of *evolutive equilibrium*. According to this assumption, the system can be divided in many subsystems of small dimensions, each of which can be considered in a thermodynamic equilibrium state, thus allowing the use of the fundamental relations of thermodynamics, with particular reference to the Gibbs equation. The fundamental justification of this hypothesis lies in the validity of the results that are obtained from its application.

However, there is a criterion that may be used to verify if the hypothesis of evolutive equilibrium for a generic system is acceptable: the comparison between the internal time-to-equilibrium of the system and the characteristic time of the perturbation applied to the system. If the former is much smaller than the latter, it means that the system reaches quickly equilibrium after a modification of the external conditions. It is thus a good approximation to use thermodynamic relations at equilibrium. On the other hand, if the characteristic time of the perturbation is comparable to the time-to-equilibrium of the system, the assumption loses its physical validity because the frequency of excitation is sufficiently high to never let the system reach a near-equilibrium condition.

In general, as the time-to-equilibrium of a system depends on its dimensions (the smaller the system, the shorter the time-to-equilibrium), progressively dividing a system into many subsystems eventually leads to a time-to-equilibrium that is much smaller than the characteristic time of the perturbation, thus making the *local equilibrium* assumption reasonable. However, there is a constraint in this process: the subsystems considered at *local equilibrium* should contain a sufficiently high number of particles. If the system were too rarefied, the equations of thermostatics would not be valid. Numerically, the applicability of these equations is verified upon the value of the Knudsen number, defined as

$$\mathrm{Kn} = \frac{\lambda}{D}, \tag{2.1}$$

where λ is the mean free path of the particles and D is the characteristic length of the system under consideration. If the Knudsen number is sufficiently close to zero, the system may be considered as sufficiently *dense*.

2.1.2 The entropy production term

Traditional non-equilibrium thermodynamics and most of the successive nonequilibrium theories develop starting from an observation on the form that the entropy production term assumes after combining the various balance equations.

The fundamental balance equations in continuum physics are the mass balance, momentum balance and energy balance equations. In non-equilibrium thermodynamics, a relevant role is attributed to a fourth equation: the entropy balance equation. The entropy balance equation in its local formulation can be written as follows:

$$\frac{\partial}{\partial t}(\rho s) = -\boldsymbol{\nabla}\cdot\boldsymbol{j}_s + \sigma_s, \tag{2.2}$$

where ρ is the mass density of the medium, s is the entropy per unit mass, $\boldsymbol{j_s}$ is the entropy flux and σ_s is the entropy production per unit volume and per unit time. The equation appears as a regular balance equation with a time-variation term, a flux term and a generation term. The peculiarity is that the second law of thermodynamics imposes that the latter must be non-negative as entropy may only be generated and cannot be destroyed.

If the first three, fundamental, balance equations are inserted into the Gibbs relation, that is, the expression for the differential of the entropy as a function of the differentials of internal energy, volume and concentrations, valid at thermodynamic equilibrium, an expression for both the entropy flux and the entropy production term may be recognized. Indeed, the resulting expression may be shown to have the form of a balance equation:

$$\begin{aligned}\rho\frac{\mathrm{D}s}{\mathrm{D}t} = &-\boldsymbol{\nabla}\cdot\left(\frac{\boldsymbol{j}_q - \sum_k \mu_k \boldsymbol{j}_k}{T}\right) - \frac{\boldsymbol{j}_q\cdot\boldsymbol{\nabla}T}{T^2} + \\ &-\frac{1}{T}\sum_{k=1}^{n}\boldsymbol{j}_k\cdot\left(T\,\boldsymbol{\nabla}\frac{\mu_k}{T} - \boldsymbol{F}_k\right) - \frac{1}{T}\Pi:\boldsymbol{\nabla}\boldsymbol{v} - \frac{1}{T}\sum_{j=1}^{r}J_jA_j,\end{aligned} \tag{2.3}$$

where $\boldsymbol{j}_q$ is the heat flow, $\boldsymbol{j}_k$ is the *diffusion flow* of component k, μ_k is the chemical potential of component k, $\boldsymbol{F}_k$ is the force per unit mass exerted on component k, Π is the non-isotropic part of the pressure tensor and, finally, J_j and A_j are the chemical reaction rate and the chemical affinity of reaction j, respectively.

Once the equation has been cast in a balance form, the flux term and the generation term may be easily identified simply by comparison with the original form of the entropy balance equation:

$$\boldsymbol{j}_s = \frac{\boldsymbol{j}_q - \sum_k \mu_k \boldsymbol{j}_k}{T}, \tag{2.4}$$

and

$$\sigma = -\frac{\boldsymbol{j}_q \cdot \boldsymbol{\nabla} T}{T^2} - \frac{1}{T}\sum_{k=1}^{n} \boldsymbol{j}_k \cdot \left(T\, \boldsymbol{\nabla} \frac{\mu_k}{T} - \boldsymbol{F}_k \right) - \frac{1}{T}\Pi : \boldsymbol{\nabla}\, \boldsymbol{v} - \frac{1}{T}\sum_{j=1}^{r} J_j A_j. \tag{2.5}$$

The entropy flow term is characterized by two contributions: a term due to the heat flow and a term due to the flow of matter. This expression is often referred to as the *heat-diffusion interaction.* It may also be seen from the previous expressions that the entropy production term is given by a bilinear form: it is the sum of the products of a *flow* with a *conjugated force.* The force, also called *affinity*, is either the gradient of an intensive state variable, an external force or the difference between thermodynamic state variables. From a physical standpoint, a thermodynamic force quantifies the gap of a certain variable from its equilibrium value. Indeed, at equilibrium, all thermodynamic forces vanish, because there are no gradients, nor gaps from equilibrium values of state variables. As a consequence of this, the entropy production term goes to zero, which is its minimum value, as it is always non-negative because of the second law of thermodynamics.

2.1.3 Linear relationships and Curie's Principle

The simplest assumption on the functional dependence of the fluxes states that a generic flux depends on the thermodynamic forces and on the thermodynamic state variables. It can be said that this assumption is based on "common sense" and reasonably simplifies successive analytical developments. However, other theories that extend the range of functional dependency to other variables have been developed. A significant example that has already been cited in the introduction is the one of non-local theories[1], such as the one developed by Jou, Casas-Vázquez, and Lebon, which claim that fluxes also depend on the gradients of the thermodynamic forces [JCL10].

---

[1] Non-local theories are theories characterized by the fact that the variables at one point in space are not determined from *local-equilibrium relationships.*

Once the functional dependence has been set, it is the particular form of the force-flux relation that characterizes a specific non-equilibrium thermodynamic theory. The simplest version is a linear relationship. Linear relations between fluxes and their conjugate thermodynamic forces had already been developed, from a strictly phenomenological point of view, in the XIX century during the study of various physical phenomena. Among these, there are Fourier's Law for heat conduction, Ohm's Law in electrodynamics and Fick's Law in mass transfer. The peculiarity of these phenomenological laws is related to the fact that each flux only depends on its corresponding force, i.e. the heat flux only on the temperature gradient, the electric current only on the gradient of the electric potential, the diffusive flow only on the concentration gradient.

However, a thermodynamic flux may, more generally, depend on a linear combination of all thermodynamic forces. The linear relationships may thus be generalized as

$$J_i = \sum_k L_{ik} X_k, \tag{2.6}$$

However, fluxes and forces have different dimensions: they can be scalars, vectors or second order tensors. In the more general case, each flux depends on all types of forces and, as a consequence of this, the linear coefficients must have different dimensions. The presence of particular symmetries for the medium under consideration allows a simplification of the more general dependence of each flux from all types of forces: this is called *Curie's symmetry principle* [Cu08]. This proof is based on the invariance of the phenomenological coefficients, that is, the coefficients relating fluxes and forces (which have different tensorial orders), under suitable transformations, related to the type of symmetry that is assumed for the medium. For example, if the medium is isotropic, the phenomenological coefficients should be invariant under inversion and under all types of rotations. This leads to the fact that, for an isotropic medium, phenomenological coefficients relating fluxes and forces of different orders are zero and the remaining ones all depend on a sole scalar value. This simplifies the expression for the entropy production term in the isotropic case. In the same way, simplified forms of the phenomenological coefficients may be found when the medium is supposed to be characterized by other types of symmetries (for example, cubic symmetry imposes invariance under inversions, 90-degrees rotations around coordinate axes and 120-degrees rotations around body diagonals) (see Chapter 6 of [dGM84]). In the isotropic case, we have thus the following relationships between forces and fluxes:

$$J^s = L^{ss} X^s, \tag{2.7}$$

$$\boldsymbol{J^v} = L^{vv} \boldsymbol{X^v}, \tag{2.8}$$

$$J^t = L^{tt} X^t, \tag{2.9}$$

where J are the fluxes, X are the forces, the letter s stands for *scalar*, the letter v stands for *vector* and the letter t stands for *tensor*. Implicit in the previous notation are the sum signs over the different forces of the same tensorial order.

2.1.4 Onsager's theory

A further simplification of the system of equations that has been obtained is possible thanks to Onsager reciprocal relations [O31], which may be stated as

$$L_{kl} = L_{lk}. \tag{2.10}$$

Considering the matrix L of phenomenological coefficients, the symmetry relationships may be written as

$$L = L^T. \tag{2.11}$$

The following paragraph will illustrate the original formulation of the proof of the validity of Onsager's symmetry relationships, which is based on the idea of microscopic reversibility, that is, the fact that a particle at the microscopic level always traces back its path when its velocity is inverted.

Proof from Statistical Mechanics

The proof of the reciprocal relationships was given by Onsager in the framework of statistical mechanics. In particular, a system of definite energy E and definite number of particles N, obeying the micro-canonical distribution (all states comprised between energy E and energy $E+dE$ are equally probable), is considered in the phase space. A macroscopic description of the system can be given on the base of n extensive variables, which compose the vector $\boldsymbol{A} = \boldsymbol{A}(\boldsymbol{r}_N, \boldsymbol{p}_N)$ that depends on the position and momentum of each particle of the system. Because of the motion of particles in the phase space, the values of $\boldsymbol{A}$ are subject to fluctuations; thus, their instantaneous value may be different from their average value and the following relevant quantity may be defined:

$$\alpha_i = A_i - \langle A_i \rangle \tag{2.12}$$

Thus, the vector $\boldsymbol{\alpha}$ is the vector containing the shift of the value of the macroscopic extensive variables from their average value (averages must be interpreted either as time averages or as averages over a certain *ensemble* of systems).

Onsager's proof, which will not be illustrated, is based on several assumptions, listed as follows.

- First of all, the probability density function $f(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is assumed to have a Gaussian distribution:

  $$f(\alpha_1, \alpha_2, \ldots, \alpha_n) = c e^{-\frac{1}{2k}\sum_{ij} g_{ij}\alpha_i\alpha_j}, \tag{2.13}$$

  where c is a normalization constant, k is Boltzmann's constant and $\boldsymbol{g}$ is a symmetric positive definite matrix.

- Secondly, the average value of $\boldsymbol{\alpha}$ is assumed to obey the linear differential equation of the first order

  $$\frac{\partial \bar{\boldsymbol{\alpha}}^{\boldsymbol{\alpha_0}}}{\partial t} = -M \bar{\boldsymbol{\alpha}}^{\boldsymbol{\alpha_0}}, \tag{2.14}$$

  where the $\boldsymbol{\alpha_0}$ indicates the imposed initial conditions and M is the matrix of phenomenological coefficients, independent of time. It has been experimentally proved for a wide range of conditions that this linear differential equation, called *regression law* is actually verified.

- Third, *time reversal invariance* of the equation of motion is considered: that is, if the momenta of all the particles at a certain point in time and space are inverted, the particles trace back their paths with reversed momenta. This hypothesis leads to the property of *microscopic reversibility.* Even though these properties have been proved in the framework of statistical mechanics, they could be obtained also considering the quantum behaviour of particles. For the sake of brevity, we omit the mathematical treatise regarding this aspect of the proof. It must also be noted that we are adopting the simplifying assumptions of the absence of magnetic fields and the consideration of variables that are even with respect to velocity. With the presence of magnetic fields and odd variables with respect to velocity, the proof would be a little more burdensome.

- Finally, Boltzmann's entropy postulate is adopted:

  $$f(\boldsymbol{\alpha}) = f(0)e^{\Delta S/k}. \tag{2.15}$$

  This postulate relates the entropy of a generic state described by the vector $\boldsymbol{\alpha}$ to the entropy of the most probable state, called *equilibrium state* (i.e. the state characterized by a null gap from the average values of the macroscopic variables), through the values of the microscopic probability density distribution of the two states.

The vector $\boldsymbol{X}$ is then defined as

$$\boldsymbol{X} \overset{\text{def}}{=} k\frac{\partial}{\partial \boldsymbol{\alpha}} \ln f, \tag{2.16}$$

and, because of **Eq.** 2.15, is related to the entropy as follows:

$$\boldsymbol{X} = \frac{\partial}{\partial \boldsymbol{\alpha}}\Delta S. \tag{2.17}$$

Thus, the vector $\boldsymbol{X}$ can be interpreted as the vector of intensive thermodynamic variables conjugated to the vector $\boldsymbol{\alpha}$ of the extensive ones, in the same way as in equilibrium thermodynamics the intensive state variables are the partial derivatives of the entropy with respect to the extensive ones (for example, $1/T = \frac{\partial s}{\partial u}$).

Based on the previous hypothesis, it can be proved that the matrix L, defined as follows, is symmetric:

$$\frac{\partial}{\partial t}\boldsymbol{\alpha}^{\boldsymbol{\alpha}_0} = L\boldsymbol{X}^{\boldsymbol{\alpha}_0}. \tag{2.18}$$

This conclusion was the one obtained by Onsager in his papers and it shows the validity of its reciprocal relations if we agree in considering the time derivatives of the state variables as *fluxes* and the quantity $\boldsymbol{X}$ as an *affinity*. The problem of the *transfer* of these relationships, obtained on a microscopical point of view, to the phenomenological equations, whose fluxes are not only time-derivatives, but usually spatial derivatives, will be faced in the following paragraph.

Under the same hypotheses, it can be proved that the quadratic forms associated to L are positive definite; that is, for a generic vector $\boldsymbol{\xi}$, it can be proved that

$$\boldsymbol{\xi}^T L \boldsymbol{\xi} \geq 0 \tag{2.19}$$

In this way, the positive definite character of the entropy production, which is a bilinear form in L, if linear relationships between forces and fluxes are assumed, is proved to hold microscopically under the previously illustrated assumptions. On the other hand, from a macroscopic point of view, the positive definite character of the entropy production derives from the second principle of thermodynamics and the positive definite character of the matrix L would be a consequence of this. The fact that, under the previously stated hypotheses, the non-negativeness of the entropy generation is derived as a consequence may be a hint on the fact that the assumptions are too restrictive; this would, however, call into question the validity of the proof of Onsager reciprocal relations. Finally, it may also be proved that, as the time goes to infinity, the average value of the shift of the vector $\boldsymbol{\alpha}$ from its *equilibrium value* goes to zero:

$$\lim_{t\to\infty} \bar{\boldsymbol{\alpha}}^{\boldsymbol{\alpha_0}}(t) = \lim_{t\to\infty} e^{-Mt}\boldsymbol{\alpha_0} = 0 \tag{2.20}$$

Extension of Onsager relations for vectorial and tensorial phenomena

Onsager relations as derived in the statistical mechanics framework are valid for scalar macroscopic variables (more precisely, they are valid for the difference of the values of these variables from their equilibrium values) and for *fluxes* that are time derivatives of these variables. However, macroscopic phenomenological equations contain vectorial and tensorial variables, as well as fluxes which are not time derivatives of state variables. As a consequence of this, it is necessary to prove that Onsager reciprocal relations are valid for phenomenological equations too.

The rigorous proof is here omitted and only a general idea of the procedure is given, examining the case of an isotropic fluid characterized by heat conduction and mass diffusion phenomena [dGM84, p. 74]. Using the Gibbs equation, expressing the differential of the entropy as a function of the differentials of the extensive variables, and the mass and energy balance equations, the following expressions for the total derivative of the entropy production is obtained:

$$\frac{dS}{dt} = \rho \int_\Omega \left(\Delta \frac{1}{T}\frac{\partial u}{\partial t} - \sum_{k=1}^{n-1} \Delta \frac{\mu_k - \mu_n}{T}\frac{\partial c_k}{\partial t} \right) dV. \tag{2.21}$$

It must be noted that, in this Equation, and in the following **Eq.** 2.22, **Eq.** 2.23, **Eq.** 2.27 and **Eq.** 2.28, the symbol Δ is used to indicate a difference and not a Laplacian. It may be seen that the entropy production is still a bilinear form in the forces and fluxes; the first ones are the differences between the local value and a generic reference value at a reference point $\boldsymbol{r}_0$, whereas the second ones are the time derivatives of state variables. The sum has become an integral because it depends on the product of forces and fluxes in all points of the considered domain. Following this interpretation and the linear structure of phenomenological relationships, the fluxes may be written as a linear combination of the forces in all points of space, that is, as an integral:

$$\frac{\partial u}{\partial t} = \rho \int \left\{ K_{qq}(\boldsymbol{r}, \boldsymbol{r}')\Delta\frac{1}{T'} - \sum_{k=1}^{n-1} K_{qk}(\boldsymbol{r}, \boldsymbol{r}')\Delta\frac{\mu_k' - \mu_n'}{T'} \right\} d\boldsymbol{r}', \tag{2.22}$$

and

$$\frac{\partial c_i}{\partial t} = \rho \int \left\{ K_{iq}(\boldsymbol{r}, \boldsymbol{r}')\Delta\frac{1}{T'} - \sum_{k=1}^{n-1} K_{ik}(\boldsymbol{r}, \boldsymbol{r}')\Delta\frac{\mu_k' - \mu_n'}{T'} \right\} d\boldsymbol{r}' \quad (i = 1, 2, \ldots, n-1). \tag{2.23}$$

For these equations, Onsager reciprocal relations are assumed to hold, except for $\boldsymbol{r} = \boldsymbol{r_0}$ and $\boldsymbol{r}' = \boldsymbol{r_0}$:

$$K_{qq}(\boldsymbol{r}, \boldsymbol{r}') = K_{qq}(\boldsymbol{r}', \boldsymbol{r}), \tag{2.24}$$

$$K_{iq}(\boldsymbol{r}, \boldsymbol{r}') = K_{qi}(\boldsymbol{r}', \boldsymbol{r}) \qquad (i = 1, 2, \ldots, n-1), \tag{2.25}$$

$$K_{ik}(\boldsymbol{r}, \boldsymbol{r}') = K_{ki}(\boldsymbol{r}', \boldsymbol{r}) \qquad (i = 1, 2, \ldots, n-1). \tag{2.26}$$

On the other hand, by inserting the linear phenomenological equations into the energy and mass balance equations, the following expressions may be obtained:

$$\rho\frac{\partial u}{\partial t} = -\mathrm{div}\left(L_{qq}\, \boldsymbol{\nabla}\, \Delta\frac{1}{T} - \sum_{k=1}^{n-1} L_{qk}\, \boldsymbol{\nabla}\, \Delta\frac{\mu_k - \mu_n}{T} \right), \tag{2.27}$$

$$\rho\frac{\partial c_i}{\partial t} = -\mathrm{div}\left(L_{iq}\, \boldsymbol{\nabla}\, \Delta\frac{1}{T} - \sum_{k=1}^{n-1} L_{ik}\, \boldsymbol{\nabla}\, \Delta\frac{\mu_k - \mu_n}{T} \right) (i = 1, 2, \ldots, n-1). \tag{2.28}$$

The term on the right of these equations may be written as an integral by multiplying the whole expression for a three-dimensional delta function centred in $\boldsymbol{r}$.

Successively integrating by parts, this integro-differential equations takes an aspect similar to **Eq.** (2.22) and **Eq.** (2.23), allowing to identify the values of the coefficients K. Mathematical manipulation of the equalities that have been obtained lead to the proof that Onsager relations hold also for the phenomenological coefficients L, as it was originally stated in Chapter 6 of [dGM84].

The procedure used in the previous paragraph to *prove* that Onsager reciprocal relationships are valid also for vectorial and tensorial phenomena on a macroscopic point of view may not seem totally convincing. On one hand, in the specific case that has been examined, taking into consideration only the phenomena of heat diffusion and mass diffusion leads to a lack of generality; on the other hand, however, the proof may seem to be built to suit *ad hoc* the formulation of the reciprocal relationships that had been obtained from the microscopic treatise.

As a consequence of this, other scholars have tried to derive a proof of Onsager reciprocal relations in other ways. Without entering a subject that is maybe more epistemological than strictly physical, the fact is that Onsager relations are shown to be actually valid in a wide range of situations as it has been showed experimentally by de Groot and Mazur [dGM84], with non-equilibrium molecular dynamics (NEMD) by Hafskjold and Kjelstrup Ratkje [HK95] and, always through NEMD, also for surface interactions by Xu et al. [XKB+06]. It may also be argued that the reciprocity of the phenomenological coefficients may be *obvious* in the light of the original hypothesis made by Onsager, that is, microscopic reversibility.

Variational formulation

Onsager's theory and its reciprocal relations may be derived from a variational principle, as Onsager himself pointed out in his 1931 article [O31] and Gyarmati further explored [Gya70].

This is formulated as follows. If thermodynamic forces are prescribed, fluxes are the solution of the optimization problem

$$\begin{cases} \max[D(\boldsymbol{J})] \\ \text{s. t.} \\ D(\boldsymbol{J}) = \boldsymbol{X} \cdot \boldsymbol{J} \end{cases} \qquad \text{where} \quad \Phi(\boldsymbol{J}) = \frac{1}{2}\boldsymbol{J}^T \mathbb{R} \boldsymbol{J}, \tag{2.29}$$

that is, seeking the maximum of a dissipation function D subject to the constraint that $D(\boldsymbol{J}) = \boldsymbol{X} \cdot \boldsymbol{J}$. Indeed, this corresponds to having chosen a quadratic dissipation function. The physical significance and the reason for this choice will be clear after Chapter 3, and in particular after **Subsec.** 3.1.2.

Onsager Reciprocal Relations in the Steepest Entropy Ascent framework

In his latest work [Ber13], proposing a unified treatment for the dynamical evolution of thermodynamic systems, Beretta deduces the existence of a linear, symmetric relationship between the fluxes and an expression related to the forces, which is based on assumptions that seem more general than the ones assumed for the traditional proof. In particular, Onsager relations appear as a natural consequence of the use of a *metric tensor*, which is symmetric by definition. In a similar way, the GENERIC theory claims that Onsager's reciprocity emerges naturally from their dynamical theory. These aspects will be clarified in the following chapters when both SEA and GENERIC will be explained in detail.

2.1.5 Minimum Entropy Production for stationary states

The proof of Onsager reciprocal relations is related to an important principle of physics which was first stated by Prigogine [Pr47] for *discontinuous systems*, generalized by Glansdorff and Prigogine [GP54], and which has been obtained by Mazur [Ma52] for *continuous systems*: the ***principle of minimum entropy production***.

The following hypotheses are assumed for the proof:

- linear phenomenological equations linking fluxes and forces are assumed to hold:
$$J_i = \sum_k L_{ik} X_k; \tag{2.30}$$
- Onsager reciprocal relationships for the phenomenological coefficients are assumed to be valid:
$$L_{ik} = L_{ki}; \tag{2.31}$$
- Onsager's phenomenological coefficients are assumed to be constant;
- constraints on the border of the considered domain are time-independent;

- the system is at mechanical equilibrium, that is,

$$\frac{d\boldsymbol{v}}{dt} = 0. \tag{2.32}$$

If these hypotheses are assumed to be valid, it can be proved that imposing the stationarity of the global entropy production P, that is $\delta P = 0$, implies that the system is at a stationary state. This means that the partial derivatives of the thermodynamic variables of the system with respect to time are zero. Also the inverse proof is valid: if the system is assumed to be in a stationary state (partial time derivatives of the state variables are zero) and the previous hypotheses are assumed to hold, then the global entropy production is stationary.

Under the same hypotheses, it may also be proved that the partial derivative of the total entropy production with respect to time is non-positive:

$$\frac{\partial P}{\partial t} \leq 0. \tag{2.33}$$

where the total entropy production P is defined as

$$P = \int_\Omega \sigma \mathrm{dV}. \tag{2.34}$$

This means that the total entropy production diminishes over time, until it reaches a minimum in a stationary state. As a consequence of this, it may be stated that stationary states are stable states, characterized by a minimum of the entropy production. In light of these considerations, it may then be remarked that *thermodynamic equilibrium is a special case of a stationary state which is reached if the boundary conditions are compatible with the equilibrium conditions* [dGM84, p. 53].

As it has been previously mentioned, a more general statement of the previous theorem was proposed by Glansdorff and Prigogine [GP54], where no use of the phenomenological relationships is made. The total entropy production may be written as

$$P = \int_\Omega \sigma \mathrm{dV} = \int_\Omega \sum_i J_i X_i \mathrm{dV}. \tag{2.35}$$

The time derivative of P may thus be written as the sum of two terms:

$$\frac{\partial P}{\partial t} = \int_\Omega \sum_i J_i \frac{\partial X_i}{\partial t} \mathrm{dV} + \int_\Omega \sum_i \frac{\partial J_i}{\partial t} X_i \mathrm{dV}. \tag{2.36}$$

This expression shows that the entropy production is either due to the time variation of the forces (first term) or to the time variation of the fluxes (second term). It may be proved that only the first term is non-positive, while the second one does not have a definite sign. As a consequence of this, in the more general case, it may be stated that the stationary state is a stable state only if the entropy production term due to the time variation of the fluxes is negative.

It is easy to confuse the ***minimum entropy production principle*** with the ***maximum entropy production principle***, whose historical evolution will be thoroughly illustrated in the following chapter. The two principles are not mutually exclusive and refer to different features of the dynamical evolution of a thermodynamic system. First of all, it must be noted that the *maximum entropy production principle claims to have a broader generality.* Moreover, the maximum entropy production principle is a local principle that rules the evolution of a system at a fixed instant of time. In simple words, the idea behind it is that, among all possible paths to equilibrium, the system chooses the one that reaches it more quickly. In particular, among all possible infinitesimal advancements at a fixed instant of time, the system chooses the one that brings to a maximum production of entropy. At the successive instant of time, the system will choose again the infinitesimal path that maximizes the entropy production among all those that are possible at that instant of time. On the other hand, as it may be understood from the present paragraph, the *minimum entropy production principle* refers to a minimization of the entropy production *during the course of time.* It states, indeed, that the time derivative of the entropy production is non-positive and, thus, that the entropy production decreases as the system evolves towards equilibrium. That is, the entropy production at one instant of time is lower than the entropy production at a previous instant of time and higher than the one at a later instant. Moreover, the principle refers globally to the whole system (the integral of the local entropy production is said to be *minimized*) and is valid under the restrictive hypotheses that have been stated, while, on the other hand, the *maximum entropy production principle* is said to be valid much more generally. As some scholars claim to deduce the reciprocity relationships from the *maximum entropy production principle* and as the *minimum entropy production principle* is based also on this hypothesis, it may be said that the latter is a consequence of the former. The link between the two will be more clear at the end of the following chapter when the features of the *maximum entropy production* principle will be illustrated. However, we felt that this digression was necessary at this point of the treatise to avoid generating confusion in the reader's mind.

2.1.6 The general problem

Phenomenological equations are necessary to *close* the problem posed by the balance equations. If we consider a system with n components, the balance equations give $n+4$ differential equations: one scalar equation for each component's mass conservation, a three-dimensional vectorial equation for momentum balance and one scalar equation for the energy balance. However, the presence of the fluxes makes the problem underconstrained.

The insertion of the phenomenological equations, expressing the fluxes as functions of the thermodynamic forces (that is, as functions of the other thermodynamic variables), leads to the reduction of the variables of the problem. The variables are then $n+4$: the density ρ, the $n-1$ concentrations c_i, the three components of the velocity $\boldsymbol{v}$ and the temperature T. Specific internal energy u and chemical potentials μ_i can be evaluated form the other variables through the equations of state. As an example, the general problem for a one-component isotropic fluid with no external forces may be stated as follows [dGM84, p. 41]. The mass balance equation

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \boldsymbol{v}); \tag{2.37}$$

the momentum balance equation

$$\rho \frac{d\boldsymbol{v}}{dt} = -\nabla p + \eta \triangle \boldsymbol{v} + \left(\frac{1}{3}\eta + \eta_v \right) \nabla (\nabla \cdot \boldsymbol{v}); \tag{2.38}$$

finally, the energy balance equation

$$\rho \frac{du}{dt} = \lambda \triangle T - p \nabla \cdot \boldsymbol{v} + 2\eta (\nabla \boldsymbol{v})^s : (\nabla \boldsymbol{v})^s + \eta_v (\nabla \cdot \boldsymbol{v})^2. \tag{2.39}$$

Obviously, the coefficients of the phenomenological equations are assumed to be known and the value of the internal energy u may be calculated from the equation of state:

$$u = u(T, \rho). \tag{2.40}$$

For a one-component system, under the hypothesis of local equilibrium, all the intensive state variables may be expressed as functions of two generic variables of the system (in this case temperature and density). Simplification of the previous system of equations under particular hypotheses leads to the description of peculiar problems, such as heat conduction in a solid, isothermal hydrodynamics and others.

It may be then stated that *non-equilibrium thermodynamics has the purpose to study various irreversible processes as heat conduction, diffusion and viscous flow from a single point of view. It englobes a number of phenomenological theories such as the hydrodynamics of viscous fluids, the theory of diffusion and the theory of heat conduction.*[dGM84, p. 42]

The additional entropy balance equation is a scalar equation that introduces five additional variables with respect to the other balance equations: specific entropy s, the three-dimensional entropy flux $\boldsymbol{J}_s$ and the specific entropy production σ. With the so-called "heat-diffusion interaction hypothesis", the entropy flux $\boldsymbol{J}_s$ is assumed to depend on the heat flux $\boldsymbol{J}_q$ and on the n mass fluxes $\boldsymbol{J}_i$ [GFB94], thus reducing the number of new variables introduced by the entropy balance equation to two. If the hypothesis of local equilibrium is valid, however, specific entropy may be obtained through the equation of state, which is valid for each point in space. For a one-component system, specific entropy depends only on two generic thermodynamic variables. The entropy balance equation allows the calculation of the entropy production rate, which is the bilinear expression that is obtained combining the balance equations with the Gibbs equation.

The "heat-diffusion interaction hypothesis" is not accepted by everyone and some theories, such as "Rational Extended Thermodynamics" [MR98], consider the entropy flux as a *constitutive quantity*, that is, as an independent variable, rather than a quantity expressed as function of other fluxes.

2.1.7 Examples of application

Linear Irreversible Thermodynamics has had great success in the description of different types of phenomena, both discrete and continuous, as it has been illustrated, in addition to the previously cited book by de Groot and Mazur [dGM84], also in the books by Woods [Wo75] and by Kuiken [Ku94]. In order to give a feeling of the versatility and usefulness of Linear Irreversible Thermodynamics, we list some examples of applications, referring to the previously cited books for a deeper explanation of these phenomena:

- among the non-equilibrium phenomena that are more frequently used in the industrial sector, there are *thermoelectric effects*, including the *Peltier effect* and the *Seebeck effect*; the Peltier effect refers to the generation of a heat flow in the presence of an electric current, while, on the other hand, the Seebeck effect refers to the generation of a potential difference in the presence of a temperature gradient, phenomenon that is at the base of the functioning of a thermocouple (Chapter 13 of [dGM84]);

- *thermokinetic effects* that include the *thermomolecular pressure*, that is, the pressure difference rquired to suppress a heat flux in the presence of a temperature gradient, and *mechano-caloric effects* (heat flow caused by a pressure difference) (Chapter 15 of [dGM84]);

- *anisotropic heat conduction* in the presence of a magnetic field, including the Righi-Leduc effect (a temperature gradient giving rise to a heat flow in an orthogonal direction) (Chapter 11 of [dGM84]);

- two-fluid description of *superfluids*, including the influence of cross viscosities on wave propagation and a discussion of the role of mutual friction between the two fluids for rotating superfluids ([Wo75]).

2.2 Kinetic Theory

The goal of Kinetic Theory is to describe the behaviour and the macroscopic properties of a system made up by an enormous number of particles obeying the laws of Classical Mechanics. Each particle may be identified through six independent variables: the three-dimensional position vector $\boldsymbol{x_i}$ and the three-dimensional velocity vector $\boldsymbol{v_i}$. As a consequence of this, the whole system, made up of N particles, may be described by a total of $6N$ variables which are located in the so-called *phase space*. The time evolution of the $6N$-dimensional vector $\boldsymbol{z}$ is thus almost impossible to determine because positions and velocities of such a huge number of particles would be needed (for systems of *ordinary* dimensions, the number N is about the same order of magnitude as Avogadro's number). As a consequence of this, the aim of Statistical Mechanics is to determine the evolution equation for $P^{(N)}(\boldsymbol{z})$ in phase space, that is, the probability density of finding a particle in $\boldsymbol{z}$[2].

---

[2] Notice that the symbol $\boldsymbol{z}$ indicates both the random variable and the value of the random variable. A distinction has not been made in order not to burden the notation, but it should remembered in the successive passages.

2.2.1 Liouville Equation

With the use of the probability density, the system may be seen as a fluid with a density equals to $P^{(N)}(\boldsymbol{z})$. As the particles may escape the infinitesimal volume only because of a flux through its surface, the probability density obeys a a conservation equation that has the same form as a *continuity equation* [Ce88b]:

$$\frac{\partial P^{(N)}}{\partial t} + \nabla \cdot (P^{(N)} \boldsymbol{Z}) = 0, \tag{2.41}$$

with $\boldsymbol{Z}$ as the *velocity* vector $\boldsymbol{Z} = \dot{\boldsymbol{z}}$, analogously to what happens for a regular *continuity equation*. This equation is called **Liouville Equation**. Assuming that the forces $\boldsymbol{F} = \dot{\boldsymbol{v}} m$ do not depend on the velocities, Liouville equation may be re-written as:

$$\frac{\partial P^{(N)}}{\partial t} + \boldsymbol{Z} \cdot \frac{\partial P^{(N)}}{\partial \boldsymbol{z}} = 0 \tag{2.42}$$

which means that the material derivative of $P^{(N)}$ must be equal to zero.

2.2.2 Boltzmann Equation

The Liouville Equation is based on the hypothesis of the absence of collisions: indeed, collisions change the velocities of the colliding particles, making the vector $\boldsymbol{z}$ *disappear* in one point of phase space and making it *appear* in another point. In simpler words, in the presence of collisions, the continuity equation in the $6N$-dimensional space is not valid anymore because collisions cause a discontinuity in the velocity of the single particles. When collisions are considered, together with specific hypotheses, the Boltzmann Equation may be obtained. In fact, it has to be pointed out that the Boltzmann Equation is related to a different *level of description* with respect to the Liouville Equation, as it will be shown in the following pages. The hypotheses that have to be stated in order to obtain the Boltzmann Equation are the following:

- having a *diluted* gas, that is, a gas where the mean free path between collisions λ is much greater than the diameter a of the molecules;
- the forces among molecules are effective only in the short range;
- as a consequence of the previously stated hypotheses, collisions are *binary*, that is, the probability of having a collision among three or more particles is negligible;

- finally, collisions are *elastic*, conserving both momentum and kinetic energy (there is no excitation of internal degrees of freedom, which is reasonable at not-too-high temperatures).

In order to introduce the Boltzmann equation, the s-particle distribution function is introduced:

$$P^{(s)} = \int P^{(N)} \prod_{i=s+1}^{N} \mathrm{d}\boldsymbol{x}_i \mathrm{d}\boldsymbol{v}_i \tag{2.43}$$

with $s \in \{1, 2, \dots, N{-}1\}$. $P^{(s)}(\boldsymbol{x}_1, \boldsymbol{v}_1, \dots, \boldsymbol{x}_s, \boldsymbol{v}_s) \prod_{i=1}^{s} \mathrm{d}\boldsymbol{x}_i \mathrm{d}\boldsymbol{v}_i$ represents the probability density in $\boldsymbol{x}_1, \boldsymbol{v}_1, \dots, \boldsymbol{x}_s, \boldsymbol{v}_s$ of s randomly chosen molecules of finding s randomly chosen molecules. It may be proved that the following equation is obtained for the time evolution of the s-particle distribution function under the previously stated hypotheses for a system of N-particles made up by hard spheres of diameter σ in the absence of forces:

$$\frac{\partial P_N^{(s)}}{\partial t} + \sum_{i=1}^{s} \boldsymbol{v} \cdot \frac{\partial P_N^{(s)}}{\partial \boldsymbol{x}_i} =$$
$$= (N-s)\sigma^2 \sum_{i=1}^{s} \int \left[P_N^{(s+1)'} - P_N^{(s+1)} \right] |\boldsymbol{V}_i \cdot \boldsymbol{n}| \mathrm{d}^3\boldsymbol{n} \mathrm{d}^3\boldsymbol{u}, \tag{2.44}$$

with ($s \in [1, N] \cap \mathbb{N}$), where $P_N^{(s+1)'}$ is the $(s+1)$-particle distribution function before the collision, $P_N^{(s+1)}$ is the $(s+1)$-particle distribution function after the collision, $\boldsymbol{V_i}$ is the relative velocity between the two colliding particles whose absolute value is conserved during the collision as a consequence of the conservations of mass and momentum, and $\boldsymbol{n}$ is the unit vector directed along the line connecting the two centres. The $(s+1)$-particle distribution functions depend on the positions $\boldsymbol{x}_i$ of the particles, on time t and on their velocities: the velocities before the collision for $P_N^{(s+1)'}$ and the velocities after the collision for $P_N^{(s+1)}$. It is important to notice that the s particle distribution function depends on the $(s+1)$-particle distribution function; this would eventually lead to the original Liouville equation. In order to overcome this problem, Ludwig Boltzmann [Bo64; CC70] stated the famous *Stosszahlansatz* hypothesis [Ce88a], that is, the *hypothesis of molecular chaos*. This apparently *harmless* hypothesis states that the joint probability of finding a number of particles in a certain point of space, having a certain velocity, may be decomposed in the product of the single-particle probabilities, that is, the various stochastic events are independent. For a two-particle probability distribution function, we have

$$P^{(2)}(\boldsymbol{x_1}, \boldsymbol{v_1}, \boldsymbol{x_2}, \boldsymbol{v_2}, t) = P^{(1)}(\boldsymbol{x_1}, \boldsymbol{v_1}, t)P^{(1)}(\boldsymbol{x_2}, \boldsymbol{v_2}, t). \tag{2.45}$$

After introducing **Eq.** 2.45 (for a general s-particle distribution function) into **Eq.** 2.44, the Boltzmann equation is finally obtained. The Boltzmann equation is usually written with the mass distribution function f as the unknown variable. This function is related to the 1-particle distribution function through

$$f = NmP^{(1)} = MP^{(1)}, \tag{2.46}$$

where N is the number of particles of the system, m is the mass of the single particle and M is thus the total mass of the system.

The Boltzmann equation, in the presence of forces and in the limit for a very high number of particles, may thus be written as

$$\frac{\partial f}{\partial t} + \boldsymbol{v} \cdot \frac{\partial f}{\partial \boldsymbol{x}} + \frac{\boldsymbol{F}}{m} \cdot \frac{\partial f}{\partial \boldsymbol{v}} = \frac{\sigma^2}{m} \int [f' f'_u - f f_u] \, |\boldsymbol{V} \cdot \boldsymbol{n}| \mathrm{d}^3\boldsymbol{n} \mathrm{d}^3\boldsymbol{u} \tag{2.47}$$

where the dependencies of the mass distribution functions f, f', f_u and f'_u are the same of the one-particle distribution functions they are biunivocally related to (the subscript u indicates the dependence from the integration variable $\boldsymbol{u}$ and, as it has been shown before, the appendix indicates pre-collision distributions, that is, distributions depending on pre-collision velocities). In the derivation of the equation [Ce88b], it has also been supposed that the *interaction of a rigid sphere with a wall is independent of the evolution of the state of the other spheres and that no particles are captured by the solid walls* [Ce75]. The Boltzmann equation is thus a nonlinear integro-differential equation which is particularly difficult to solve even for the most simple cases. Because of this, various methods of approximate solution have been proposed and will be illustrated in the following paragraphs. From now on, the term on the right side of the equation will be called *collision term* or *collision integral* and indicated with $Q(f, f)$ to have a lighter notation.

2.2.3 Collision invariants

Given the *collision integral* illustrated in the preceding paragraph, a ***collision invariant*** is a function $\varphi(\boldsymbol{v})$, defined as follows:

$$\int Q(f, f)\varphi(\boldsymbol{v})\mathrm{d}\boldsymbol{v} = 0 \tag{2.48}$$

It may be proved ([Ce88b; Ce90]) that all functions satisfying the preceding equation also fulfill the equality

$$\varphi(\boldsymbol{v_1}) + \varphi(\boldsymbol{v_2}) = \varphi(\boldsymbol{v_1'}) + \varphi(\boldsymbol{v_2'}), \tag{2.49}$$

where the functions on the right-hand side depend on the velocities after the collision. The interpretation of this equation is quite clear: *collision invariants* are quantities that are conserved during an elastic collision. It may be proved that each *collision invariant* is a linear combination of the quantities that are conserved during an elastic collision, i.e. mass, momentum and kinetic energy [Ce88b; Ce90]:

$$\varphi(\boldsymbol{v}) = C_0 + C_1 v_x + C_2 v_y + C_3 v_z + C_4 \boldsymbol{v}^2 = a + \boldsymbol{b} \cdot \boldsymbol{v} + c\boldsymbol{v}^2 \tag{2.50}$$

where the C_i are the linear combination coefficients, $a = C_0$, $\boldsymbol{b} = \{C_1, C_2, C_3\}$ and $c = C_4$.

2.2.4 Equilibrium solutions and Maxwellian distribution

It is interesting to find the *equilibrium solution* of the Boltzmann equation. With the terms *equilibrium solution* and *equilibrium state* we refer to the conditions listed in the following passages. The reader must however be warned that the term has to be used carefully: for a deeper discussion we refer to [Ce88b]. The *equilibrium conditions* on the Boltzmann equation are the following:

- absence of external forces;
- the solution does not depend explicitly on time (the partial derivative with respect to time is equal to zero);
- the system is spatially homogeneous, that is, there are no density gradients, $\dfrac{\partial f}{\partial \boldsymbol{x}} = 0$.

As a consequence of these assumptions, the left-hand term of **Eq.** 2.47 becomes zero and the Boltzmann equation may be written as:

$$Q(f, f) = 0 \tag{2.51}$$

The necessary and sufficient condition to satisfy this equation is [Ce88b; Ce90]

$$f' f'_u = f f_u, \tag{2.52}$$

or, by taking the logarithm of both sides,

$$\ln f' + \ln f'_u = \ln f + \ln f_u. \tag{2.53}$$

Thus, we may see that $\ln f$ is a *collision invariant*, according to the definition in **Eq.** 2.49. As a consequence of this, it may be expressed as a linear combination of the basis for collision invariants:

$$\ln f^{eq} = a + \boldsymbol{b} \cdot \boldsymbol{v} + c\boldsymbol{v}^2. \tag{2.54}$$

Consequently, the *equilibrium solution* for the Boltzmann equation is

$$f^{eq}(\boldsymbol{v}) = e^{a+\boldsymbol{b}\cdot\boldsymbol{v}+c|\boldsymbol{v}|^2} \tag{2.55}$$

or, introducing a constant vector $\boldsymbol{w}$ and two constants α and A ((a, $\boldsymbol{b}$, c) $\longrightarrow$ (α, $\boldsymbol{w}$, A)),

$$f^{eq}(\boldsymbol{v}) = Ae^{-\alpha(\boldsymbol{v}-\boldsymbol{w})^2}, \tag{2.56}$$

which is a Maxwellian distribution for a gas characterized by linear uniform motion with velocity $\boldsymbol{w}$.

Macroscopic quantities are computed as suitable moments of the 1-particle distribution function and macroscopic balance equations are derived by proper integrations. For the sake of brevity, we will omit this derivation and only a simple example of the relationship between macroscopic quantities and microscopic distribution functions will be shown. Given the one-particle distribution function P such that $P(\boldsymbol{x}, \boldsymbol{v}, t)\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}$ is the probability of finding one particle in the infinitesimal neighbourhood of $\boldsymbol{x}$ and $\boldsymbol{v}$, the spatial probability density $P_{\boldsymbol{x}}$, that is, the probability of finding a particle around the spatial coordinate $\boldsymbol{x}$, independently of the velocity, may be computed by integration:

$$P_{\boldsymbol{x}} = \int P \mathrm{d}^3\boldsymbol{v}. \tag{2.57}$$

The infinitesimal mass $\mathrm{d}M(\boldsymbol{x}, t)$ is equal to

$$\mathrm{d}M = \rho \mathrm{d}^3\boldsymbol{x} = NmP_{\boldsymbol{x}}\mathrm{d}^3\boldsymbol{x} = \mathrm{d}^3\boldsymbol{x} \int f \mathrm{d}^3\boldsymbol{v}, \tag{2.58}$$

where **Eq.** 2.46 has been used. Thus, we have that:

$$\rho(\boldsymbol{x}, t) = \int f \mathrm{d}^3\boldsymbol{v}. \tag{2.59}$$

Considering this definition, the Maxwellian distribution expressed through the macroscopic quantities may be written as:

$$f^{eq} = \frac{\rho}{(2\pi\theta)^{3/2}} \cdot \exp\left[-\frac{(\boldsymbol{v}-\boldsymbol{w})^2}{2\theta}\right] \tag{2.60}$$

where $\theta = \frac{k_B T}{m}$, with k_B as Boltzmann's constant, T as the temperature of the gas (related to the average value of the Kinetic Energy of the particles) and $\boldsymbol{w}$ as the average macroscopic velocity of the medium.

By defining other macroscopic variables such as the stress tensor $\boldsymbol{\sigma}$, the internal energy u and the heat flux q, the macroscopic balance equations illustrated in **Eq.** 2.37, **Eq.** 2.38 and **Eq.** 2.39 are then obtained. The idea is quite simple: the Boltzmann equation has to be multiplied by the five collision invariants φ_i and then integrated over all velocities. Because of the definition of collision invariant, the term on the right-hand side of the equation disappears and only the terms on the left remain. These terms are then re-written according to the definition of the macroscopic variables, thus yielding the balance equations.

According to what has been shown, the Boltzmann equation admits a particular form of the mass distribution function f as the equilibrium distribution, even if nothing has been said on *how* this equilibrium distribution is approached. Boltzmann has also showed that the average value of $\ln f$ has a particular property:

$$\langle \ln f(\boldsymbol{v}, t) \rangle = \frac{1}{\rho} \int f(\boldsymbol{v}, t) \ln f(\boldsymbol{v}, t) \mathrm{d}\boldsymbol{v} = \frac{H(t)}{\rho}, \tag{2.61}$$

where $H(t)$ is Boltzmann's H-function. Boltzmann proved that [Ce75]:

$$\frac{\mathrm{d}H}{\mathrm{d}t} \leq 0, \tag{2.62}$$

where the equality sign only holds at equilibrium for a Maxwellian distribution. Generalizing with the use of the collision integral, we have:

$$\int Q(f,f) \ln f \mathrm{d}^3\boldsymbol{v} \leq 0 \tag{2.63}$$

where the equality holds if f is a Maxwellian.

2.2.5 Model equations and approximate solution

The main problem in solving the Boltzmann equation is related to the complex form of the collision integral. In order to simplify it, alternative collision operators, named *collision models* have been proposed. A collision model $J(f)$ must satisfy two fundamental properties in order to reproduce a behaviour similar to the one of the collision integral $Q(f,f)$:

- the collision model must conserve the *collision invariants* φ_α,

  $$\int \varphi_\alpha J(f) \mathrm{d}^3\boldsymbol{v} = 0 \qquad (\alpha \in [0,4] \cap \mathbb{N}_0); \tag{2.64}$$

- the collision term $J(f)$ must satisfy

  $$\int \ln f J(f) \mathrm{d}\boldsymbol{v} \leq 0, \tag{2.65}$$

  with the equality holding only for a Maxwellian.

Welander and BGK

The simplest and most widely used *collision model* has been proposed in 1954 by Bhatnagar, Gross, and Krook [BGK54] and it is named after the initials of the last names of the authors, that is, the BGK model. It is also called the *relaxation time model* and has been proposed independently by Welander [We54]. The idea underlying this model is that the effects caused by the collisions are proportional to the deviation from the *local Maxwellian*. As a consequence of this, the collision term takes the form

$$J(f) = \frac{f^{eq} - f}{\tau}, \tag{2.66}$$

where τ is the *relaxation time* and may be seen as constant in the simplest model, while it could be considered as dependent from other variables in more articulated models. It may be proved that this model satisfies the H-theorem that was illustrated before, while one of the main drawbacks is the fact that it is not able to predict the correct Prandtl number, which results always equal to one. It is also straightforward to prove that it satisfies the condition of preserving the collision invariants; indeed, we have that

$$\begin{cases} J(f) = \dfrac{[f^{eq}(\boldsymbol{v}) - f(\boldsymbol{v})]}{\tau} \\ \displaystyle\int \varphi_\alpha f^{eq}(\boldsymbol{v}) \mathrm{d}^3\boldsymbol{v} = \int \varphi_\alpha f(\boldsymbol{v}) \mathrm{d}^3\boldsymbol{v} \qquad (\alpha \in [0,4] \cap \mathbb{N}_0) \end{cases} \tag{2.67}$$

The last equation is satisfied because the *equilibrium Maxwellian* is, by definition, the one that has the same value of the collision invariants of the mass distribution function f, i.e. the same value of ρ, $\boldsymbol{v}$ and T. It must be noted that J is a non-linear function of f, since the parameters of the Maxwellian ρ, $\boldsymbol{v}$ and T are, in turn, function of f according to their definitions.

In order to achieve a more precise value of the Prandtl number, other models have been proposed, such as the Ellipsoidal-Statistical-BGK (ES-BGK) [CT66; ALPP00] and the velocity-dependent collision frequency model [Kro59; St97; MS04], $\nu(\boldsymbol{v})$-BGK, which will be used in Chapter 5. The collision term for this model may be written as:

$$Q_{\nu(\boldsymbol{v})-BGK}(f) = \nu(\boldsymbol{v})(f^P - f) \tag{2.68}$$

where $\nu(\boldsymbol{v})$ is the velocity-dependent collision frequency and f^P is a pseudo-Maxwellian whose meaning will be illustrated in Chapter 5. If a power law dependence of the type $\nu(\boldsymbol{v}) = |\boldsymbol{v} - \boldsymbol{w}|^\alpha$ is chosen, it may be proved, as illustrated in [St97], that the ratio between thermal conductivity and dynamic viscosity is a precise expression depending on the value of α. Thus, by adjusting the value of the exponent, the suitable value of the Prandtl number may be obtained.

Chapman-Enskog and Grad methods

Another way of facing the problem of the complicated form of the collision integral is the one adopted by the so-called *approximate methods of solution*. It is not a goal of the present work to illustrate the analytical development of these methods, which is extremely onerous from a mathematical point of view. For a precise exposition of the subject, we recall [Ce75]. We just aim to give an idea of their construction.

Regarding the *Chapman-Enskog expansion*, the starting point is the nondimensionalization of the Boltzmann equation, which may be written as [Li03]

$$\left(\frac{\partial}{\partial t}+\hat{D}\right)f=\frac{1}{Kn}\hat{I}f \tag{2.69}$$

where $\hat{D}$ is the differential operator containing partial derivatives with respect to space coordinates and velocities, $\hat{I}$ is the collision integral operator and Kn is a nondimensional parameter that emerges from the nondimensionalization of the equation: the Knudsen number. As it has already been stated in the previous paragraphs, the Knudsen number represents the ratio between the mean free path of the particle and a typical length scale of the system [Ce75]. The Chapman-Enskog expansion is based on the hypothesis that ϵ is small, that is, the system is collision-dominated (the mean free path of the particle is significantly smaller than the typical length of the system). Based on these hypothesis, the function f, the collision operator and the time derivative are expanded in series of Kn [Li03]. By stopping the expansion at a certain order, different systems of equations are obtained: each successive iterate of the Chapman-Enskog expansion gives a more detailed set of hydrodynamic equations. For example, stopping to the first iterate bring to the *Euler equations*, while the second iterate gives the *Navier-Stokes equations* and the third iterate gives the so-called *Burnett equations* [Li03].

The so-called *Grad's moment method* is instead based on the idea of obtaining macroscopic transport equations from different *moments* of the velocity of the Boltzmann equation: the Boltzmann equation is multiplied by different powers of the velocity and then integrated over all velocity space. This method has, obviously the problem of *closure*: each moment equation is characterized by the presence of a higher moment. As a consequence of this, in order to *close* the problem, a higher moment must be expressed as a function of *lower moments*. The method proposed by Grad is based on the assumption that the fourth-order moments are related to the second and zero order moments. In this way, the so-called *twenty-moments approximation* is obtained. With further simplifications and assumptions, approximations with a lower number of moments may be obtained, corresponding to the well known *Navier-Stokes* or *Euler* equations [Li03].

2.2.6 Examples of application

The Boltzmann equation has been used and is currently used to model many different physical phenomena. The following is just a list of the main applications of this *tool* in the most diverse fields of physics.

Rarefied gas dynamics

As it has also been stated in the introductory Chapter, the kinetic approach is used to model physical situations characterized by a Knudsen number $Kn = \frac{\lambda}{L}$ significantly different from zero, when the Navier-Stokes equations fail [Had06]. Traditionally, most of the interest towards this kind of situations was related to the modelling of external, high-speed flows associated with flight in the upper atmosphere. Recently, however, the interest in microscale and nanoscale internal flows has given a strong input to the study of other physical situations when Navier-Stokes description is expected to fail and kinetic effects are important. Because of this historical reason, flows characterized by the failure of the Navier-Stokes description are referred as *rarefied gas dynamics* [Bi94; Ko69] or *noncontinuum* flows, which may lead to the idea that this kind of flows are associated also to a breakdown of the continuum assumption. It must however be highlighted that, for a large class of these problems, even though the Navier-Stokes equations fail, the hydrodynamic fields and the conservation laws remain well defined, arising from moments of the Boltzmann equation [VK65]. It must then be noted that, for a large class of scholars, failure of the Navier-Stokes equations is different from the failure of the *continuum* approach [Had06]. Another important application of the Boltzmann equation is the one related to **Lattice Boltzmann Methods** (LBM). These are methods of Computational Fluid Dynamics (CFD) simulation where, instead of the traditional conservation equations (i.e. mass, momentum and energy), a discretized Boltzmann equation is solved for fictive particles undergoing collisions and streaming processes over a discrete mesh. The description in terms of macroscopic variables arises form the resolution of this discretized system of partial differential equations and these methods have been shown to be more useful than regular CFD methods in the modelling of complex geometries and situations with high Knudsen number [Su13]. In **Fig.** 2.1 the possible different approaches to Fluid Dynamics simulation are shown.

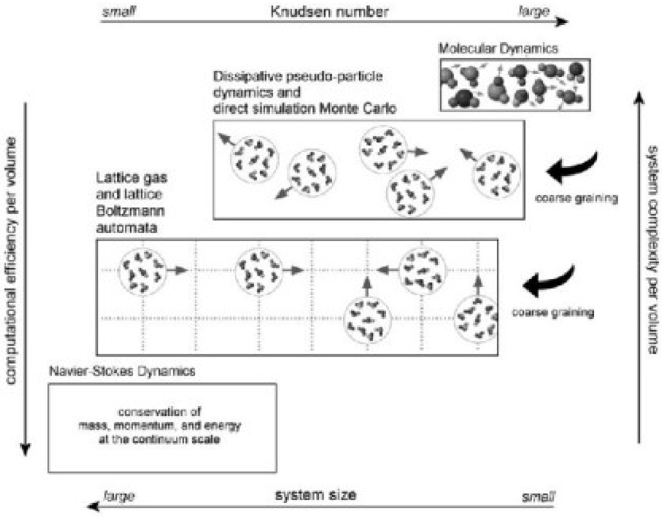

Fig. 2.1: Different Approaches to Fluid Dynamics simulation, classified according to relevant parameters, such as system size and complexity of modelling per unit volume.

Electron and phonon transport in solids

Another important application of the Boltzmann equation is the study of electron and phonon transport in solids, with particular regard to the determination of transport coefficients. The key idea is that electrons in a solid are assumed to obey the probability distribution function $f = f(\boldsymbol{r}, \boldsymbol{p}, t)$ in phase-space. If collisions are not considered, the governing equation is the Liouville equation; on the other hand, if the collisions are considered, the collision term is usually modelled with a *mean relaxation time* $\tau(\boldsymbol{p})$ term:

$$Q(f, f) = -\frac{f(\boldsymbol{r}, \boldsymbol{p}, t) - f_0(\boldsymbol{r}, \boldsymbol{p}, t)}{\tau(\boldsymbol{p})}. \tag{2.70}$$

Electrons are then modelled as semiclassical Bloch wave packets moving in a partially occupied band, the equilibrium distribution f_0 is assumed to be the Fermi-Dirac distribution and the probability distribution function f is approximated through a first order expansion in $\boldsymbol{r}$ and $\boldsymbol{p}$. After further developments, the expression for the electric conductivity tensor and the thermal conductivity due to electrons may be obtained. On the other hand, also phonon packets in an insulator may be modelled through the Boltzmann equation in order to compute the expression for the thermal conductivity of the material. Analytical passages are similar to those that have been described for electron transport, even though it must be noted that, as phonons are bosons and not fermions, they obey the Bose-Einstein distribution [Z72].

Neutron transport

A further application of the Boltzmann Equation is the modelling of neutron behaviour in a reactor (in this case, the equation is usually referred to as the Boltzmann Transport Equation, or BTE). Neutron transport may be modelled either deterministically, through the BTE, or stochastically, through Montecarlo methods. With the first method, the system is described through a differential equation related to the underlying physical process: the equation is then discretized in order to be solved numerically. On the other hand, with the use of Montecarlo methods, the physical process is simulated directly and only the probability density functions (pdfs) related to its different phases are needed: Montecarlo simulation then proceeds by random sampling from the pdfs. The Boltzmann Transport Equation for neutrons is obtained as a balance between the neutrons that are generated and the neutrons that are *consumed*. It is a complicated integro-differential equation in seven dimensions: three in space $\boldsymbol{r}$, two in direction $\boldsymbol{\Omega}$ and one each in energy E and time t. It may be written as

$$\begin{aligned}\left[\frac{1}{v}\frac{\partial}{\partial t} + \boldsymbol{\Omega}\cdot\nabla + \Sigma(\boldsymbol{r},E,t)\right]\phi(\boldsymbol{r},\boldsymbol{\Omega},E,t) &= \\ = \int_0^\infty \mathrm{d}E' \int_{4\pi} \mathrm{d}\boldsymbol{\Omega}'\Sigma_s(\boldsymbol{r},\boldsymbol{\Omega}\cdot\boldsymbol{\Omega},E\to E)\phi(\boldsymbol{r},\boldsymbol{\Omega}',E',t) &+ \\ + \frac{\chi(E)}{4\pi}\int_0^\infty \mathrm{d}E' \int_{4\pi} \mathrm{d}\boldsymbol{\Omega}'\nu(E')\Sigma_f(\boldsymbol{r},E',t)\phi(\boldsymbol{r},\boldsymbol{\Omega}',E',t) + Q(\boldsymbol{r},\boldsymbol{\Omega},E,t), \quad (2.71)\end{aligned}$$

where v is the absolute value of the velocity of the neutron, biunivocally associated to its energy $E = 1/2mv^2$, Σ is the macroscopic total cross section, Σ_s is the macroscopic scattering cross section, Σ_f is the macroscopic fission cross section, ϕ is the neutron flux, $\nu(E)$ is the average number of neutrons produced per fission and $\chi(E)$ represents the probability distribution function related to the neutrons produced isotropically in a fission. The left-hand side of the equation represents the neutrons lost either because of leakage or because of collisions that result in scattering or absorption. The right-hand side of the equation represents the source terms: the neutrons scattered in that direction and energy, from all others directions and energies, the fission neutrons produced in that direction and energy and the source term Q. The equation is usually discretized in space, directions and time and neutrons are *clustered* into energy groups [La66].

Kinetic description of plasmas

Plasmas are also objects of study of Non-Equilibrium Thermodynamics. Whenever the continuum hypothesis is valid for them, Classical Non-Equilibrium Thermodynamics (as outlined in **Sec.** 2.1) can be applied, and one deals with multiple-fluid or single-fluid (Magneto-HydroDynamic) models. There are certain situations in which fluid models of the Navier-Stokes type are not applicable and one has to switch to kinetic models. Of course, it is again a matter of time scales: if one wants to study the response of a plasma to perturbations comparable to its natural resonance frequency, fluid models are not appropriate neither from the conceptual nor from the mathematical standpoint. The most important example is represented by the study of longitudinal modes, which are also called *electrostatic modes* since, for their description, the magnetic field may be neglected and the plasma model is studied in conjunction with Gauss Law

$$\boldsymbol{\nabla}\cdot\boldsymbol{E}=\frac{\rho}{\varepsilon_0};$$

this, indeed, is responsible of the longitudinal component of the wave.

Usually, the time scales of the phenomena under consideration are longer than the typical relaxation time of the system (thermalisation of the plasma population), so that one can neglect collisions and focus onto *Vlasov Equation*, another name to say a *collisionless Boltzmann Equation*. After this equation has been coupled with Gauss Law and the electrostatic potential has been introduced, the resulting system of equation is called *Vlasov-Poisson system*. For instance, here is the system for an electronic population:

$$\begin{cases}\boldsymbol{E}(\boldsymbol{x},t)=-\boldsymbol{\nabla}\,\phi(\boldsymbol{x},t)\\ \boldsymbol{\nabla}\cdot\boldsymbol{E}(\boldsymbol{x},t)=\dfrac{e}{\varepsilon_0}\displaystyle\int f(\boldsymbol{x},\boldsymbol{v},t)\mathrm{d}^3\boldsymbol{v}\\ \dfrac{\partial f(\boldsymbol{x},\boldsymbol{v},t)}{\partial t}+\boldsymbol{v}\cdot\dfrac{\partial f(\boldsymbol{x},\boldsymbol{v},t)}{\partial \boldsymbol{x}}-\dfrac{e}{m_e}\cdot\dfrac{\partial f(\boldsymbol{x},\boldsymbol{v},t)}{\partial \boldsymbol{v}}=0\end{cases}$$

Analytically solving the linearised version of this system was a great challenge which required complex analysis and the theory of Schwartz' distributions. Once this has been done, one disposes of the dispersion relation related to electrostatic modes, and observes certain peculiar phenomena, such as the famous one of Landau damping, which could not be uncovered with a fluid model.

References

[ALPP00] P. Andries, P. Le Tallec, J.-P. Perlat, and B. Perthame. "The Gaussian-BGK model of the Boltzmann equation with small Prandtl number". In: *European Journal of Mechanics: B - Fluids* 19 (2000), pp. 813–830 (cit. on p. 52).

[Ber13] G. P. Beretta. "Steepest-Entropy-Ascent and Maximal-Entropy-Production Dynamical Models of Irreversible Relaxation to Stable Equilibrium from Any Non-Equilibrium State. Unified Treatment for Six Non-Equilibrium Frameworks". In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013.* Ed. by M. Pilotelli and G. P. Beretta. Brescia: Cartolibreria Snoopy, 2013, pp. 100–109 (cit. on p. 39).

[BGK54] P. L. Bhatnagar, E. P. Gross, and M. Krook. "A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems". In: *Physical Review* 94.3 (1954), pp. 511–525 (cit. on p. 51).

[Bi94] G. A. Bird. *Molecular Gas Dynamics and the Direct Simulation of Gas Flows.* Oxford: Clarendon, 1994 (cit. on p. 54).

[Bo64] L. Boltzmann. *Lectures on Gas Theory.* Berkeley: University of California Press, 1964 (cit. on p. 46).

[CC70] S. Chapman and T. G. Cowling. *The Mathematical Theory of Nonuniform Gases.* Cambridge: Cambridge University Press, 1970 (cit. on p. 46).

[Ce75] C. Cercignani. *Theory and Application of the Boltzmann Equation.* Edinburgh and London: Scottish Academic Press, 1975 (cit. on pp. 47, 51, 53).

[Ce88a] C. Cercignani. *Ludwig Boltzmann, The Man Who Trusted Atoms.* Oxford: Oxford University Press, 1988 (cit. on p. 46).

[Ce88b] C. Cercignani. *The Boltzmann Equation and Its Applications.* New York: Springer-Verlag, 1988 (cit. on pp. 45, 47–49).

[Ce90] C. Cercignani. *Mathematical Methods in Kinetic Theory.* New York: Plenum Press, 1990 (cit. on pp. 48–49).

[CT66] C. Cercignani and G. Tironi. “Some Applications of a Linearized Kinetic Model with Correct Prandtl Number”. In: *Il Nuovo Cimento B* 43 (1966), pp. 64–78 (cit. on p. 52).

[Cu08] P. Curie. *Oeuvres de Pierre Curie*. Ed. by S. F. de Physique. Paris: Gauthier-Villars, 1908 (cit. on p. 32).

[dGM84] S. R. de Groot and P. Mazur. *Non-equilibrium Thermodynamics*. 2nd ed. (1st ed., North-Holland, 1962). Mineola (New York): Dover Publications, 1984 (cit. on pp. 29, 32, 36, 38, 40, 42–44).

[GFB94] E. P. Gyftopoulos, M. I. Flik, and G. P. Beretta. “What is Diffusion?” In: *Journal of Energy Resources Technology* 116 (1994), pp. 136–139 (cit. on p. 43).

[GP54] P. Glansdorff and I. Prigogine. “Sur les proprietes differentielles de la production d’entropie”. In: *Physica* 20 (1954), pp. 773–780 (cit. on pp. 39–40).

[Gya70] I. Gyarmati. *Non-equilibrium thermodynamics: Field theory and variational principles*. Engineering science library. Springer, 1970 (cit. on p. 38).

[Had06] N. G. Hadjiconstantinou. “The limits of Navier-Stokes theory and kinetic extensions for describing small-scale gaseous hydrodynamics”. In: *Physics of Fluids* 18 (2006) (cit. on p. 54).

[HK95] B. Hafskjold and S. Kjelstrup Ratkje. “Criteria for Local Equilibrium in a System with Transport of Heat and Mass”. In: *Journal of Statistical Physics* 78 (1995), pp. 463–494 (cit. on p. 38).

[JCL10] D. Jou, J. Casas-Vázquez, and G. Lebon. *Extended Irreversible Thermodynamics*. 4th ed. Springer, 2010 (cit. on p. 31).

[Ko69] M. N. Kogan. *Rarefied Gas Dynamics*. New York: Plenum, 1969 (cit. on p. 54).

[Kro59] M. Krook. “Continuum Equations in the Dynamics of Rarefied Gases”. In: *Journal of Fluid Mechanics* 6 (1959), pp. 523–541 (cit. on p. 52).

[Ku94] G. D. Kuiken. *Thermodynamics of Irreversible Processes*. Chichester: Wiley, 1994 (cit. on p. 43).

[La66] J. R. Lamarsh. *Introduction to Nuclear Reactor Theory*. Boston: Addison-Wesley, 1966 (cit. on p. 57).

[Li03] R. L. Liboff. *Kinetic Theory: Classical, Quantum and Relativistic Descriptions*. New York: Springer, 2003 (cit. on p. 53).

[Ma52] P. Mazur. "Sur les etats a production d'entropie minimum dans les systemes continus". In: *Bulletin de l'Academie Royale de Belgique* 38 (1952) (cit. on p. 39).

[MR98] I. Muller and T. Ruggeri. *Rational Extended Thermodynamics*. 2nd ed. (1st ed., Springer-Verlag, 1993). New York: Springer Verlag, 1998 (cit. on p. 43).

[MS04] L. Mieussens and H. Struchtrup. "Numerical comparison of Bhatnagar-Gross-Krook models with proper Prandtl number". In: *Physics of Fluids* 16 (2004), pp. 2797–2813 (cit. on p. 52).

[O31] L. Onsager. "Reciprocal relations in irreversible processes. I." In: *Physical Review* 37.4 (1931), pp. 405–426.
L. Onsager. "Reciprocal relations in irreversible processes. II". In: *Physical Review* 38.12 (1931), pp. 2265–2279. Cit. on pp. 33, 35, 38.

[Pr47] I. Prigogine. *Étude thermodynamique des phénomènes irréversibles*. Liège: Desoer, 1947 (cit. on p. 39).

[St97] H. Struchtrup. "The BGK-model with velocity-dependent collision frequency". In: *Continuum Mechanics and Thermodynamics* 9.1 (1997), pp. 23–31 (cit. on p. 52).

[Su13] S. Succi. *The Lattice Boltzmann equation for Fluid Dynamics and Beyond*. Oxford: Oxford Science Publications, 2013 (cit. on p. 54).

[VK65] W. G. Vincenti and C. H. Kruger. *Introduction to Physical Gas Dynamics*. Malabar, FL: Krieger, 1965 (cit. on p. 54).

[We54] P. Welander. "On the temperature jump in a rarefied gas". In: *Arkiv Fysik* 7 (1954), pp. 507–553 (cit. on p. 51).

[Wo75] L. C. Woods. *The Thermodynamics of Fluid Systems*. Oxford: Clarendon Press, 1975 (cit. on pp. 43–44).

[XKB+06] J. Xu et al. “Verification of Onsager’s Reciprocal Relations for Evaporation and Condensation Using Non-Equilibrium Molecular Dynamics”. In: *Journal of Colloid and Interface Science* 299 (2006), pp. 452–463 (cit. on p. 38).

[Z72] J. M. Ziman. *Principles of the Theory of Solids*. 2nd ed. (1st ed., Cambridge University Press, 1964). Cambridge: Cambridge University Press, 1972 (cit. on p. 56).

3

Seeking a dynamical Law

In their view of Thermodynamics, which we like to adopt in this thesis, Gyftopoulos and Beretta take the position that the two Laws of Thermodynamics are nothing but two general features of the Dynamics of any (thermodynamic) system from any state to another. In other words, if an equation were discovered which describes the motion of the system in state space (without discussing if and to what extent could the equation be used in practical terms to calculate motions of complex systems), the two Laws could be straightforwardly derived from it as theorems.

Thus, as it has been pointed out by many authors such those in [CLaw; MS06], Thermodynamics lacks a Dynamical Law which, superimposed to the chosen Kinematics of the system, completes the general description of all its possible trajectories.

The **Maximum Entropy Production Principle** (MEPP) is a valuable candidate in that it gives a possible and consistent choice of evolution of a system and, perhaps, it has been the most studied in the last decades. As Martyushev and Seleznev wrote in their large review on this principle [MS06], the MEPP was proposed independently by several scientists in different contexts where it was used to solve specific problems. Its most general gist may be stated with Martyushev and Seleznev:

> *"By this principle, a nonequilibrium system develops so as to maximize its entropy production under present constraint." [MS06, p. 3]*

On the other hand, it has been largely criticized in its general validity by many authors, both on theoretical and experimental grounds. In this thesis, we do not aim to prove the generality of the principle, but only that it is useful to describe the dynamics of certain systems.

3.1 Ziegler's extremum principle

As far as we now, although the MEPP was proposed differently by some authors, the first clear statement and use in Nonequilibrium Thermodynamics was by Hans Ziegler in 1957 [Z57]. Changing the notation of the original articles, this reads as follows:

> *"If the force vector $\boldsymbol{X}$ is prescribed, the actual flux vector $\boldsymbol{J}$ maximizes the dissipation rate σ subject to the side condition $D(\boldsymbol{J}) = \boldsymbol{J} \cdot \boldsymbol{X}$" [Z68, p. 416]*

The original goal of the MEPP is to generalize Onsager's theory with a more general principle applying to nonlinear phenomena too.

> *"Once the connection between fluxes and forces is known, the rate of dissipation can be expressed, by means of the dissipation function in terms of the fluxes (and possibly the instantaneous values of the state variable) alone. The reverse is not true: the dissipation function does not completely determine the relationship between fluxes and forces unless an additional statement is provided by Onsager's theory which thus yields the forces in terms of the dissipation function. [Z68, p. 415]"*

The "statement" that substitutes the other one in the quotation is chosen to be his *orthogonality principle*. Let's summarize the mathematical problem.

3.1.1 Mathematical formulation

Following Ziegler, we formulate the problem in terms of prescribed forces and fluxes as unknowns. As in Classical Nonequilibrium Thermodynamics, it is also supposed that fluxes depend on forces and state variables, and knowing the relationship between fluxes and forces closes the description of the dynamics. For the benefit of the reader, we omit the dependence on state variables, which plays no role in the following treatment. In particular, the dissipation function D is that function of fluxes which, if calculated in correspondence of a certain set of fluxes, gives the entropy production σ:

$$\sigma = D(\boldsymbol{J}) \tag{3.1}$$

As described by Ziegler, the mathematical form of classical linear force-flux relations in terms of the dissipation function may be written as follows. First of all, fluxes may be expressed as linear combination of forces:

$$\boldsymbol{J}(\boldsymbol{X}) = \mathbb{L}\boldsymbol{X}, \tag{3.2}$$

with the *resistivity matrix* $\mathbb{R}$ defined as

$$\mathbb{R} \overset{\text{def}}{=} \mathbb{L}^{-1}. \tag{3.3}$$

The dissipation function, given by the product of forces and fluxes, may then be written as

$$\boldsymbol{J}(\boldsymbol{X}) = \mathbb{L}\boldsymbol{X}, \tag{3.4}$$

and its derivative with respect to the fluxes is

$$\frac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} = (\mathbb{R} + \mathbb{R}^T)\boldsymbol{J}, \tag{3.5}$$

and, considering Onsager symmetry relations, we obtain that

$$\mathbb{R}^T = \mathbb{R} \tag{3.6}$$

$$\frac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} = 2\mathbb{R}\boldsymbol{J}(\boldsymbol{X}) = 2\boldsymbol{X} \tag{3.7}$$

where **Eq.** 3.5 and **Eq.** 3.2 have been used in the last passages.

The classical linear force-flux relations may be then be rewritten in the equivalent form

$$\begin{cases} \lambda \dfrac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} = \boldsymbol{X} \\ \lambda = \dfrac{1}{2} \end{cases}, \tag{3.8}$$

which leads to the wanted relation by inverting the function on the left hand side to get $\boldsymbol{J}$ as a function of $\boldsymbol{X}$. Moreover, it suggests a generalization by letting λ

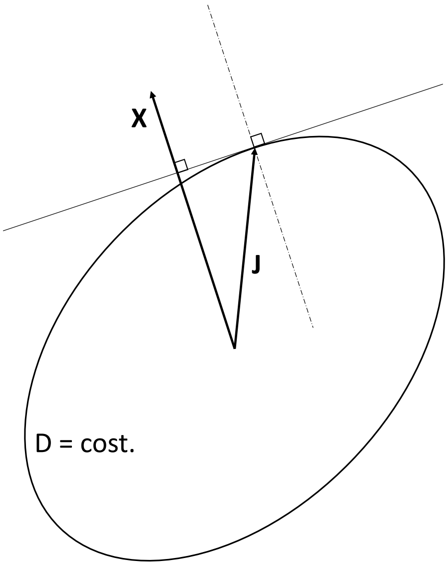

Fig. 3.1: Vector representation of fluxes and forces.

be a function of $\boldsymbol{J}$. In particular, the final result obtained by Ziegler is

$$\frac{D(\boldsymbol{J})}{\dfrac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} \cdot \boldsymbol{J}} \frac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} = \boldsymbol{X} \tag{3.9}$$

and was also called *orthogonality principle* because this is equivalent to state that, if the thermodynamics forces are prescribed, the flux vector must be such to point the contour line of $D(\boldsymbol{J})$ where the force is orthogonal to the latter (see **Fig.** 3.1). Indeed, we may see that the first term on the left-hand side of the equation is a scalar; thus, the previous expression means that the gradient of the dissipation function with respect to the fluxes must be parallel to the forces. As the gradient is locally orthogonal to the curve at constant D, the tangent at any point on a curve at constant D, which is orthogonal to the local gradient, must also be orthogonal to the forces. Euler Theorem for homogeneous functions clearly implies that **Eq.** (3.9) reduces to

$$\frac{1}{k} \frac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} = \boldsymbol{X} \tag{3.10}$$

for a homogeneous function of degree k, and – for $k = 2$ (the case of linear phenomenological relations) – to (3.7). Therefore, any linear phenomenological law (which makes the dissipation function be a bilinear form of the fluxes) implies a MEPP: however, the set of possible mathematical functions which satisfy (3.9) is larger than the set of bilinear forms, so that Ziegler's principle may constitute a proper generalization of the linear theory.

Eq. (3.9) can be derived from the variational principle stated above, which takes the expression

$$\begin{cases} \max D(\boldsymbol{J}) \\ \text{s. t.} \\ D(\boldsymbol{J}) = \boldsymbol{X} \cdot \boldsymbol{J} \end{cases} \tag{3.11}$$

The physical significance of the constraint is given by the balance equations. Indeed, as explained in **Subsec.** 2.1.2, the form of the entropy production is due to the balances of energy, momentum, etc., so that – focusing on the entropy production term only – the constraint appears to be their full representative. The constrained optimization is performed by means of the usual Lagrange multipliers. We define the Lagrangian

$$L(\boldsymbol{J}, \lambda) = D(\boldsymbol{J}) - \lambda[D(\boldsymbol{J}) - \boldsymbol{X} \cdot \boldsymbol{J}] \tag{3.12}$$

and solve for it the unconstrained problem by putting its partial derivatives to zero:

$$\begin{cases} (1-\lambda)\dfrac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} + \lambda \boldsymbol{X} = \boldsymbol{0} \\ D(\boldsymbol{J}) - \boldsymbol{X} \cdot \boldsymbol{J} = 0 \end{cases} \tag{3.13}$$

where the first equation is the derivative with respect to $\boldsymbol{J}$, while the second equation of the system is the derivative with respect to the Lagrange multiplier.

From the first equation,

$$\frac{\lambda - 1}{\lambda} \frac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} = \boldsymbol{X} \tag{3.14}$$

Substituting back into the constraint,

$$D(\boldsymbol{J}) - \frac{\lambda - 1}{\lambda} \frac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} \cdot \boldsymbol{J} = 0 \qquad \Longrightarrow \qquad \frac{\lambda - 1}{\lambda} = \frac{D(\boldsymbol{J})}{\dfrac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} \cdot \boldsymbol{J}}, \tag{3.15}$$

and then using **Eq.** 3.14, we get the result

$$\frac{D(\boldsymbol{J})}{\dfrac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} \cdot \boldsymbol{J}} \frac{\partial D(\boldsymbol{J})}{\partial \boldsymbol{J}} = \boldsymbol{X}. \tag{3.16}$$

We may visualize the constrained optimization performed in **Eq.** (3.11) by means

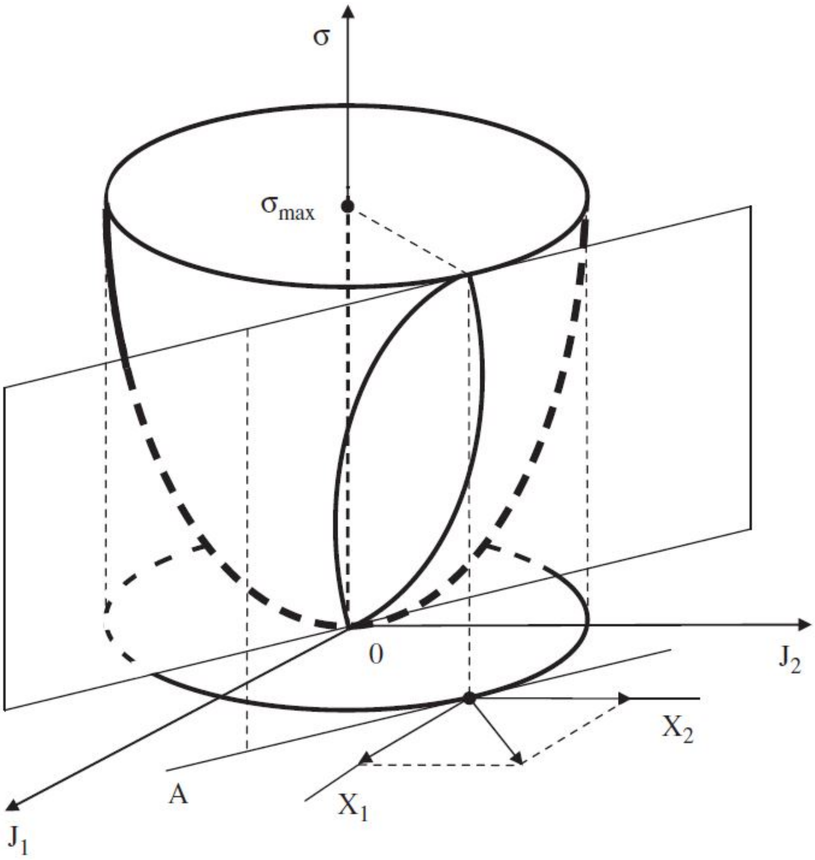

Fig. 3.2: Simplest (quadratic) dissipative surface $D(\boldsymbol{J})$ during a linear irreversible process. Here J_1 and J_2 are two thermodynamic fluxes and 0 denotes the point corresponding to the equilibrium state. [MS06]

of **Fig.** 3.2. The graph of the dissipation function is represented as a paraboloid, and the constraint is the plane $X_1 J_1 + X_2 J_2 = 0$ (where X_1, X_2 are constants). If the dissipation function has suitable convexity features, the problem has a unique solution, which is also proven to be a maximum. In this regards, Ziegler imposes the condition that the component of the thermodynamic force vector along the flux vector increases for increasing fluxes, that is $\boldsymbol{X} \cdot \boldsymbol{J}/|\boldsymbol{J}|$ grows if $|\boldsymbol{J}|$ grows, in order "*to exclude instabilities (in the form of self-excited oscillations) [...]. It then follows from* (3.1) *that $D(\boldsymbol{J})$ does not only increase on any ray emanating from the origin O in $\boldsymbol{J}$-space, but increases faster than $|\boldsymbol{J}|$. Thus, the D-surfaces are ordered, each one of them containing those with smaller values of D. Moreover, they become more dense with increasing D, in the sense that the distances between the points of intersection of an equidistant family of D-surfaces with a given ray from O decrease monotonically.*" [Z68, p. 419] In other words, the dissipation function is convex with respect to fluxes. This *stability* condition assures that, to each value of the force vector, a unique value of the flux vector corresponds. We thus have the constitutive law $\boldsymbol{J} = \boldsymbol{j}(\boldsymbol{X})$ or $\boldsymbol{X} = \boldsymbol{x}(\boldsymbol{J})$ (as the relationship between forces and fluxes is bijective, it is also invertible). If we take the dissipation function be a bilinear form of the fluxes, we have already noticed that the form of **Eq.** (3.9) is preserved, so that linear phenomenological laws imply a MEPP.

> *"It appears that, physically as well as mathematically, the principle of maximum dissipation rate (or maximum rate of entropy production) is the simplest and most satisfactory of the principles enumerated [in the article]. Subject to the stability condition, it supplies the generalization we have been looking for, and it even ensures the convexity of the D-surfaces." [Z68, p. 419]*

3.1.2 Ziegler's MEPP implies Onsager relations

A very important characteristic of the MEPP is that it entails Onsager reciprocity in the near-equilibrium domain. This is significant because Onsager relations are a widely accepted feature of non-equilibrium dynamics, which – as outlined in **Subsec.** 2.1.4 – has been proven valid through different theoretical explanations and experiments. Below, we shall give the proof of the fact that Onsager's reciprocity may be deduced if the Maximum Entropy Production Principle and linear relationships between forces and fluxes are assumed as hypotheses.

By expanding the flux vector in a Maclaurin series to first order:

$$\boldsymbol{J} = \boldsymbol{j}(\boldsymbol{X}) = \frac{\partial \boldsymbol{j}(\boldsymbol{0})}{\partial \boldsymbol{X}} \boldsymbol{X} + o(\boldsymbol{X}) = \mathbb{L}\boldsymbol{X} + o(\boldsymbol{X}) \qquad \text{for} \quad |\boldsymbol{X}| \to 0, \tag{3.17}$$

we see that the phenomenological matrix $\mathbb{L}$ is the coefficient of the first-order term. Assuming the function $\boldsymbol{j}$ to be invertible in a neighbourhood of the equilibrium point, and considering that $\boldsymbol{j}(\boldsymbol{0}) = \boldsymbol{0}$ by choice, by the inverse function theorem $\boldsymbol{x}(\boldsymbol{J}) \overset{\text{def}}{=} \boldsymbol{j}^{-1}(\boldsymbol{J})$ satisfies

$$\frac{\partial \boldsymbol{x}(\boldsymbol{0})}{\partial \boldsymbol{J}} = \left[\frac{\partial \boldsymbol{j}(\boldsymbol{0})}{\partial \boldsymbol{X}}\right]^{-1} \qquad \Longrightarrow \qquad \mathbb{R} = \mathbb{L}^{-1}.$$

Therefore, proving Onsager reciprocity is equivalent to proving the symmetry of this matrix. Moreover, using the first order approximation and the conservation equations (which have as a consequence that the dissipation function is the product of forces and fluxes),the entropy production takes the form

$$\sigma = D(\boldsymbol{J}) = \boldsymbol{X}^T(\boldsymbol{J})\boldsymbol{J} \simeq \boldsymbol{J}^T \left[\frac{\partial \boldsymbol{x}(\boldsymbol{0})}{\partial \boldsymbol{J}}\right]^T \boldsymbol{J}; \tag{3.18}$$

that is, the dissipation function is quadratic in the vicinity of equilibrium. By using these simple facts, it is easy to arrive at the desired result.

We shall use the following notation:

$$\partial_k \overset{\text{def}}{=} \frac{\partial}{\partial J_k}.$$

Assuming as hypothesis the functional form (3.9) of the constitutive relations between forces and fluxes, it is sufficient to prove that

$$\partial_j x_k = \partial_k x_j,$$

which is Onsager's reciprocity.

Proof. Using Euler Theorem for homogeneous functions (the dissipation function is an homogeneous function of the second order, using a linear approximation of the force-flux relations), we have that:

$$J_\ell \partial_\ell D = 2D.$$

we obtain

$$\begin{aligned}
\partial_j x_k &\overset{(3.9)}{=} \partial_j \left(\frac{D}{J_\ell \partial_\ell D} \partial_k D\right) \overset{(3.18)}{=} \frac{1}{2}\partial_j \partial_k D;\\
\partial_k x_j &\overset{(3.9)}{=} \partial_k \left(\frac{D}{J_\ell \partial_\ell D} \partial_j D\right) \overset{(3.18)}{=} \frac{1}{2}\partial_k \partial_j D.
\end{aligned}$$

The result follows from Schwarz theorem. Thus, we have proved that, if the generalized Ziegler relationship **Eq.** 3.9 and linearity of the force-flux relations are assumed, Onsager reciprocity may be derived. ☐

3.2 Steepest Entropy Ascent

The Steepest Entropy Ascent (SEA) principle to model the dynamics of a thermodynamic system was originally proposed in Intrinsic Quantum Thermodynamics by Beretta [Ber81] and it was recently adapted to meso- and macroscopic thermodynamic frameworks [Ber13]. Since the Fisher metric that was adopted in the quantum theory is not general enough to reproduce the dynamics of the other thermodynamic theories, an extension of the previous formulation was needed.

The idea behind the SEA construction is to *geometrize* the thermodynamic system and choose the direction of evolution of the state in a proper space. The formulation of the principle may be expressed as follows:

> *"The time evolution and transport equations advance the local state representative in the direction of maximal entropy production per unit of distance traveled in state space compatible with the conservation constraints." [Ber13, p. 101].*

In this thesis, we shall give a more mathematical and abstract formulation of the SEA principle with respect to the one that was expressed originally, in order to reach the following, equally important, objectives:

- making SEA more comparable to other abstract theories, such as GENERIC;
- clearly recognizing certain mathematical concepts.

For now, we do not have the tools to deal with the rigorous mathematical formulation. Hence, we postpone the full discussion to Chapter 4. Here we only anticipate the final results.

The time evolution of the state is represented by a curve on a manifold, and this is an integral curve of a vector field (the velocity vector is parallel to the vector field at each point of the curve). The vector field is composed of two distinguished parts: the first one is a *reversible* contribution, whose form is known from each physical theory; the second one is a *dissipative* contribution, which is modelled with a Maximum Entropy Production principle (MEPP). After concentrating the attention on the latter part, one would want the state to go in the direction of the gradient of the entropy in state space.

But an "obstacle" is imposed by the conservation laws: if the states is pushed in the direction of the entropy gradient, the conserved quantities may really not be conserved. Therefore, the dissipative vector field must point in the direction of maximal directional derivative of the entropy compatibly with the orthogonality to the gradients of all conserved quantities.

3.2.1 SEA main features

The principal advantages of the SEA formalism are:

- Direct implementation of the Second Law of Thermodynamics into the dynamic law itself.
- Implication of Onsager reciprocal relations in the near-equilibrium domain.

3.3 Edelen's theory

During the seventies, Dominic G. B. Edelen proposed a generalization of Onsager's linear theory based on a few thermodynamically consistent assumptions. Here we present the most mature formulation in [BEK79]. Since a wider view of the subject will be available towards the end of the thesis, we postpone the discussion and comments to Chapter 4.

3.3.1 Thermodynamic requirements for fluxes

According to the exposition in [BEK79], the basic thermodynamic requirements for a nonlinear system of constitutive relations

$$\boldsymbol{J} = \boldsymbol{j}(\boldsymbol{X}, \boldsymbol{\omega}) \qquad (\boldsymbol{X} \in \mathbb{R}^n) \tag{3.19}$$

are the following:

- **forces vanishes in the equilibrium state**;
- **fluxes can be expressed as functions of forces and state variables**:

$$\boldsymbol{J} = \boldsymbol{j}(\boldsymbol{X}, \boldsymbol{\omega}) \qquad \boldsymbol{j} \in C^{\infty}(\mathbb{R}^n) \quad \forall \boldsymbol{\omega}; \tag{3.20}$$

- **the dissipation inequality**:

$$\sigma = \boldsymbol{X} \cdot \boldsymbol{J} = \boldsymbol{X} \cdot \boldsymbol{j}(\boldsymbol{X}, \boldsymbol{\omega}) = D(\boldsymbol{X}, \boldsymbol{\omega}) \geq 0; \tag{3.21}$$

- the linear approximation should yield **Onsager's reciprocal relations**:

$$\begin{aligned} & \boldsymbol{j}(\boldsymbol{X}, \boldsymbol{\omega}) = \mathbb{L}(\boldsymbol{\omega})\boldsymbol{X} + o(\boldsymbol{X}) \quad \text{for} \quad |\boldsymbol{X}| \to 0 \\ \Longrightarrow \qquad & \mathbb{L}(\boldsymbol{\omega})^T = \mathbb{L}(\boldsymbol{\omega}). \end{aligned} \tag{3.22}$$

3.3.2 Onsager's linear theory

In the linear theory, we have:

$$\sigma = D(\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{X}\cdot\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{X}^T\mathbb{L}(\boldsymbol{\omega})\boldsymbol{X} \tag{3.23}$$

$$\boldsymbol{\nabla}\, D(\boldsymbol{X},\boldsymbol{\omega}) = [\mathbb{L}(\boldsymbol{\omega}) + \mathbb{L}(\boldsymbol{\omega})^T]\boldsymbol{X} = 2\mathbb{L}(\boldsymbol{\omega})\boldsymbol{X} \tag{3.24}$$

The differential form $\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega})\cdot \mathrm{d}\boldsymbol{X}$ – in this case – is exact, since:

$$\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega})\cdot \mathrm{d}\boldsymbol{X} = \mathbb{L}(\boldsymbol{\omega})\boldsymbol{X}\cdot \mathrm{d}\boldsymbol{X} = \mathrm{d}\left[\frac{D(\boldsymbol{X},\boldsymbol{\omega})}{2}\right] = \mathrm{d}\Phi(\boldsymbol{X},\boldsymbol{\omega}) \tag{3.25}$$

3.3.3 Generalizations via integrating factor

Some attempts to generalize Onsager's linear theory are based on admitting that the differential form has an integrating factor $u(\boldsymbol{X},\boldsymbol{\omega})$. That is, we can write

$$\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega}) = u(\boldsymbol{X},\boldsymbol{\omega})\,\boldsymbol{\nabla}\,\Phi(\boldsymbol{X},\boldsymbol{\omega}), \tag{3.26}$$

so that the following differential form is *exact*:

$$\frac{\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega})}{u(\boldsymbol{X},\boldsymbol{\omega})}\cdot \mathrm{d}\boldsymbol{X} = \mathrm{d}\Phi(\boldsymbol{X},\boldsymbol{\omega}). \tag{3.27}$$

and $\Phi(\boldsymbol{X},\boldsymbol{\omega})$ is called *dissipation function*. The notion of *integrating factor* is defined based on **Eq.** 3.27.

This and the linear theory are based on the additional requirement that, in order to render the differential form *closed* (which is a necessary condition for the *exactness*), these *modified* fluxes $\dfrac{\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega})}{u(\boldsymbol{X},\boldsymbol{\omega})}$ must be *Onsager fluxes*, i.e. fluxes which satisfy

$$\left\{\frac{\partial}{\partial \boldsymbol{X}}\left[\frac{\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega})}{u(\boldsymbol{X},\boldsymbol{\omega})}\right]\right\}^T = \frac{\partial}{\partial \boldsymbol{X}}\left[\frac{\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega})}{u(\boldsymbol{X},\boldsymbol{\omega})}\right]. \tag{3.28}$$

and this follows from Schwarz theorem if we assume the function Φ is C^1 in the forces.

3.3.4 Generalization via Helmholtz-Hodge decomposition

The most general method to decompose a differential form on a compact, oriented, Riemannian manifold M and more explicit than the previous attempt with an integrating factor is the so called *Helmholtz-Hodge decomposition* (see Chapter B). Since $\boldsymbol{X} \in \mathbb{R}^n$ with the standard Euclidean metric, the same decomposition may be applied directly to the vector fields[1], so that:

$$\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{j}_{\text{cons}}(\boldsymbol{X},\boldsymbol{\omega}) + \boldsymbol{j}_{\text{noncons}}(\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{\nabla}\,\Phi(\boldsymbol{X},\boldsymbol{\omega}) + \boldsymbol{y}(\boldsymbol{X},\boldsymbol{\omega}), \tag{3.29}$$

where $\boldsymbol{j}_{\text{cons}}$ is a 'conservative' vector field ($\boldsymbol{j}_{\text{cons}} \in \mathfrak{X}_{\text{cons}}(\mathbb{R}^n)$), namely a vector field that can be written as the gradient of a smooth function $\Phi \in C^\infty(\mathbb{R}^n)$; and $\boldsymbol{y}$ belongs to the orthogonal complement of $\mathfrak{X}_{\text{cons}}(\mathbb{R}^n)$ ($[\mathfrak{X}_{\text{cons}}(\mathbb{R}^n)]^\perp = \mathfrak{X}(\mathbb{R}^n)\backslash\mathfrak{X}_{\text{cons}}(\mathbb{R}^n)$).

As explained by Bataille, Edelen, and Kestin (we do not report the arguments, about which we refer to [BEK79]), the dissipation inequality imposes a condition on $\boldsymbol{y}$ (which we rename $\boldsymbol{U}$), so that the final result for a thermodynamically consistent decomposition is:

$$\begin{cases} \boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{\nabla}\,\Phi(\boldsymbol{X},\boldsymbol{\omega}) + \boldsymbol{U}(\boldsymbol{X},\boldsymbol{\omega}) \\ \boldsymbol{X}\cdot\boldsymbol{U}(\boldsymbol{X},\boldsymbol{\omega}) = 0, \end{cases} \tag{3.30}$$

where $\Phi(\boldsymbol{X},\boldsymbol{\omega})$ is a smooth scalar-valued function such that

$$\Phi(\boldsymbol{X},\boldsymbol{\omega}) = \int_0^1 D(\lambda\boldsymbol{X},\boldsymbol{\omega})\frac{\mathrm{d}\lambda}{\lambda}, \qquad D(\lambda\boldsymbol{X},\boldsymbol{\omega}) = O(\lambda^2). \tag{3.31}$$

Therefore, the requirement about *Onsager fluxes* may be relaxed in the following way. Define

$$\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{j}_D(\boldsymbol{X},\boldsymbol{\omega}) + \boldsymbol{j}_N(\boldsymbol{X},\boldsymbol{\omega}),$$

where

$$\boldsymbol{j}_D(\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{\nabla}\,\Phi(\boldsymbol{X},\boldsymbol{\omega})$$

is the dissipative part, and

$$\boldsymbol{j}_N(\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{U}(\boldsymbol{X},\boldsymbol{\omega}), \qquad \boldsymbol{X}\cdot\boldsymbol{U}(\boldsymbol{X},\boldsymbol{\omega}) = 0,$$

the nondissipative part of $\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega})$. Moreover, $\boldsymbol{\nabla}\,\Phi(\lambda\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{O}(\lambda)$ and $\boldsymbol{U}(\lambda\boldsymbol{X},\boldsymbol{\omega}) = \boldsymbol{O}(\lambda^2)$ (the nondissipative fluxes, near equilibrium, go to zero faster than the dissipative ones).

---

[1] This is possible, as shown in Chapter B, through the isomorphism $g^\sharp$.

The exact differential form is instead:

$$[\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega}) - \boldsymbol{j}_N(\boldsymbol{X},\boldsymbol{\omega})] \cdot \mathrm{d}\boldsymbol{X} = \mathrm{d}\Phi(\boldsymbol{X},\boldsymbol{\omega}). \tag{3.32}$$

Hence, the *Onsager* requirement is equivalent to the closedness of this differential form:

$$\left\{\frac{\partial}{\partial \boldsymbol{X}}\left[\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega}) - \boldsymbol{j}_N(\boldsymbol{X},\boldsymbol{\omega})\right]\right\}^T = \frac{\partial}{\partial \boldsymbol{X}}\left[\boldsymbol{j}(\boldsymbol{X},\boldsymbol{\omega}) - \boldsymbol{j}_N(\boldsymbol{X},\boldsymbol{\omega})\right]. \tag{3.33}$$

These symmetry relations can be written in the equivalent form

$$\left[\frac{\partial \boldsymbol{j}_D(\boldsymbol{X},\boldsymbol{\omega})}{\partial \boldsymbol{X}}\right]^T = \frac{\partial \boldsymbol{j}_D(\boldsymbol{X},\boldsymbol{\omega})}{\partial \boldsymbol{X}} \tag{3.34}$$

since the dissipative part of the flux, $\boldsymbol{j}_D$, is given by $\boldsymbol{j} - \boldsymbol{j}_N$.

In the light of this discussion, we see that in Edelen's formalism there is place for a non-dissipative (reversible) contribution. This opens up to the problem of the one- vs. two-generator formalisms fronted by GENERIC: we shall take it up again in Chapter 4, where we will also resume the theory above from the standpoint of GENERIC.

References

[BEK79] J. Bataille, D. G. B. Edelen, and J. Kestin. “On the structuring of thermodynamic fluxes: a direct implementation of the dissipation inequality”. In: *International Journal of Engineering Science* 17.5 (1979), pp. 563–572 (cit. on pp. 72, 74).

[Ber13] G. P. Beretta. “Steepest-Entropy-Ascent and Maximal-Entropy-Production Dynamical Models of Irreversible Relaxation to Stable Equilibrium from Any Non-Equilibrium State. Unified Treatment for Six Non-Equilibrium Frameworks”. In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013*. Ed. by M. Pilotelli and G. P. Beretta. Brescia: Cartolibreria Snoopy, 2013, pp. 100–109 (cit. on p. 71).

[Ber81] G. P. Beretta. “On the general equation of motion of quantum thermodynamics and the distinction between quantal and nonquantal uncertainties”. thesis. 1981 (cit. on p. 71).

[CLaw] A. Bejan. “Constructal-theory network of conducting paths for cooling a heat generating volume”. In: *International Journal of Heat and Mass Transfer* 40.4 (1997), pp. 799–816.
A. Bejan and S. Lorente. “The constructal law and the thermodynamics of flow systems with configuration”. In: *International Journal of Heat and Mass Transfer* 47.14-16 (2004), pp. 3203–3214 .
A. Bejan and S. Lorente. “Constructal law of design and evolution: Physics, biology, technology, and society”. In: *Journal of Applied Physics* 113.15 (2013). Cit. on p. 63.

[GB05] E. P. Gyftopoulos and G. P. Beretta. *Thermodynamics: Foundations and Applications*. 2nd ed. (1st ed., Macmillan, 1991). Mineola (New York): Dover Publications, 2005 (cit. on p. 63).

[MS06] L. M. Martyushev and V. D. Seleznev. “Maximum entropy production principle in physics, chemistry and biology”. In: *Physics Reports* 426.1 (2006), pp. 1–45 (cit. on pp. 63–64, 68).

[Z57] H. Ziegler. “Thermodynamik und rheologische Probleme”. In: *Ingenieur-Archiv* 25.1 (1957), pp. 58–70 (cit. on pp. 64, 66).

[Z68] H. Ziegler. “A Possible Generalization of Onsager’s Theory”. In: *Irreversible Aspects of Continuum Mechanics and Transfer of Physical Characteristics in Moving Fluids*. Ed. by Parkus, H. and Sedov, L. I. IUTAM Symposia. Springer Vienna, 1968, pp. 411–424 (cit. on pp. 64–66, 68–69).

"[...] an 'ideal physical theory' [...is ...] a theory in which there are well defined mathematical constructs and well defined rules for translating physical reality into these constructs; having done so the mathematics then grinds out whatever answers it can and these are then translated back into physical statements. The point here is that while 'physical intuition' is a useful guide for formulating the mathematical structure and may even be a source of inspiration for constructing mathematical proofs, it should not be necessary to rely on it once the initial 'translation' into mathematical language has been given."

Elliot H. Lieb and Jakob Yngvason *in*
The Physics and Mathematics of the Second Law of Thermodynamics

4

Geometric Thermodynamics

The quotation opening this chapter describes the most ambitious goal of Physics, and it is very similar to the sixth of the famous 'Mathematische Probleme', whose translation in [Hi02] reads:

> *"The investigations on the foundations of geometry suggest the problem:* To treat in the same manner, by means of axioms, those physical sciences in which mathematics plays an important part *[...]."*

In this chapter, we will make use of very important and known concepts in Differential Geometry (we have written Chapter A for the reader not used to the subject). In the spirit of this chapter, we add the beautiful introduction by Mrugała to his article [Mr00]:

> *"The basic problem of any physical theory is to find the proper set M of all plausible states, i.e. the so-called state space of the system. The other necessary ingredient is the structure of this space. The structure is usually defined by means of a tensor, vector or covector field, or by connection. The group preserving the geometrical structure of M is considered as the symmetry group of the theory. From this point of view any physical theory can treated as a branch of geometry in*

> *the broadest meaning of this word. This approach is well known for instance in classical mechanics, special and general relativity, electrodynamics, gauge fields, quantum mechanics, and so on.*
>
> *Contrary to the above-mentioned theories the situation is not so clear in thermodynamics. First of all there are two phenomenological and statistical approaches to the thermal phenomena. For the second, the situation is relatively simple only for homogeneous equilibrium systems. It complicates remarkably for general nonequilibrium systems where one has to deal with a big number of macroscopic variables of various types."*

At this point, we are thus faced to the following situation:

- As testified by Einstein at **Page** 5, Thermodynamics is the most 'consolidated' physical theory.
- On the other hand, according to Mrugała, Thermodynamics has not been clearly 'geometrized' yet, at variance with other theories.

Following the line of thought of Grmela and Öttinger, we recognize the following steps on the route towards this 'geometrization':

- Classical Mechanics can be formulated in an abstract (general) setting, in the context of Geometric Mechanics ([A89; MR03]): the natural arenas are symplectic manifolds, and their generalization, i.e., Poisson manifolds.
- Also Thermodynamics may be made a physical theory as in the quotation above. The most famous examples are, in Equilibrium Thermodynamics, the work by Carathéodory [C09], the book by Hermann [He73], and – for example – the references in [Q07]; in Nonequilibrium Thermodynamics, the important geometric structure of *metriplectic manifolds* (see some history and references in [Mor09] and in [F05]), on which we shall focus our attention herein.
- In its most renowned presentation, metriplectic dynamics has been called *General Equation for the Non-Equilibrium Reversible-Irreversible Coupling* (GENERIC) [GÖ97], which represents also a generalization in the context of *contact manifolds*.

- A simpler and less structured approach (*Steepest-Entropy-Ascent* dynamics) was proposed by Beretta in a quantum landscape [B81], and recently adapted for meso- and macroscopic systems in [B13].

In general, a reversible evolution is modelled with an antisymmetric tensor, while the irreversible one with a symmetric object. When the latter is a tensor, clearly it may be represented by a kind of metric tensor. This is why below we shall associate dissipation with a metric tensor.

As for the interpretations applied to thermodynamics, we have already touched the subject in Chapter 1. Here we briefly mention the two main lines of thought in the narrower context of its geometric formulation:

- According to spirit of the Keenan school of Thermodynamics, we take the position that Thermodynamics is valid at every scale, entropy is an intrinsic property of matter: this is brought to the point that even microscopically there is dissipation (an example is represented by Intrinsic Quantum Thermodynamics). A particular dynamical equation is typical of the Kinematics pertaining to a certain level of description.

- Multiscale dynamics: Grmela sees Thermodynamics as "*a meta-theory addressing relations among dynamical theories formulated on different levels of description*" [Gr13]; this relations are expressed in the framework of contact structures. At every scale there is a GENERIC, and in passing from a more detailed level to a less detailed one through the method of Coarse Graining [Ö05] one sees dissipation, even if microscopic dynamics were reversible.

As a final comment before entering the beautiful geometrical exposition of Classical Mechanics, we anticipate and stress that the most consolidated dynamical models are designed for **isolated systems**, which we shall consider in this work. Therefore, the parameters characterizing the dynamics do not explicitly depend on time: both thermodynamic potentials and vector fields are fixed, and the state representative moves on the integral lines of them (time is the parameter of these curves). For driven systems, some solutions have been proposed, as we shall discuss later.

4.1 Mechanics and Geometry

In nineteenth and twentieth centuries, and much research is still alive, Classical Mechanics has been formulated in the mathematical language of Differential Geometry, so that an entire new field was born under the name of Geometric Mechanics. Following this route, some attempts have been made to translate also Thermodynamics (in particular, Non-Equilibrium Thermodynamics) in this setting.

In this section, we shall give the main notions of Geometric Mechanics following one the main reference books in the field: *Introduction to Mechanics and Symmetry* by Marsden and Ratiu.

4.1.1 Symplectic manifolds

Let M be a Banach manifold and

$$\begin{array}{rcl} \Omega \; : \; M & \rightarrow & \Lambda^2 T^* M \\ z & \mapsto & \Omega_z \end{array}$$

a smooth differential two-form, namely $\Omega \in \Omega^2(M)$.[1]

Ω is **(weakly) nondegenerate** if, $\forall z \in M$,

$$\Omega_z(v_1, v_2) = 0 \quad \forall v_2 \in T_z M \qquad \Longrightarrow \qquad v_1 = 0; \tag{4.1}$$

that is, if v_1 is 'Ω_z-orthogonal' to every vector in $T_z M$, it is the zero vector.

Ω_z induces the associated vector bundle homomorphism

$$\begin{array}{rcl} \Omega_z^\flat \; : \; T_z M & \rightarrow & T_z^* M \\ v & \mapsto & \Omega_z^\flat(v) \end{array}$$

defined by $\Omega_z^\flat(v_1)(v_2) = \Omega_z(v_1, v_2)$. Therefore, weakly nondegeneracy translates into "$\Omega_z^\flat(v) = 0$ implies $v = 0$", which – for the linearity of the map $\Omega_z^\flat$ – is equivalent to injectivity of $\Omega_z^\flat$.

In case of surjectivity in addition to injectivity (so that $\Omega_z^\flat$ is a vector space isomorphism), Ω is called **strongly nondegenerate**. This guarantees the continuity of the inverse function as well, thanks to a theorem of Functional Analysis (in particular, from a corollary of the *open mapping theorem*). In finite dimen-

---

[1] We apologize for using the same symbol for the differential form and the global section $\Omega^2(M)$, but this is the usual notation.

sions, since a linear map between finite-dimensional spaces of the same dimension is injective if and only if it is surjective, weak and strong degeneracies coincide.

Definition 4.1. A **symplectic manifold** is a pair (M, Ω), where M is a manifold and Ω is a closed (weakly) nondegenerate two-form on M (called **symplectic form**). The manifold is **strongly symplectic** is Ω is strongly nondegenerate.

One of the most important results in symplectic geometry is Darboux Theorem, which essentially says that every symplectic manifold is locally 'flat'; namely, that is possible to find coordinates in which the components of the symplectic form are constants.[2]

Theorem 4.2 (**Darboux Theorem**). *If (M, Ω) is a strong symplectic manifold, in a neighbourhood of each $z \in M$ there is a local coordinate chart in which Ω is constant.*

Corollary 4.3. *If (M, Ω) is a finite-dimensional symplectic manifold, M is even-dimensional ($\mathrm{dim} M = 2n$), and in a neighbourhood of each $z \in M$ there are local coordinates $(q^1, \dots q^n, p^1, \dots p^n)$ in which*

$$\Omega = \sum_{k=1}^{n} \mathrm{d}q^i \wedge \mathrm{d}p^i .$$

*Such coordinates are called **canonical coordinates** (or **Darboux coordinates**, or **symplectic coordinates**).*

In Chapter A we have not talked about integration and orientation on manifolds. However, for completeness, and for the benefit of those who are acquainted to the subject, we report the following

Theorem 4.4. *Let (M, Ω) be a $2n$-dimensional symplectic manifold. Then, M is oriented by the **Liouville volume form***

$$\Lambda = \frac{(-1)^{n/2}}{n!} \underbrace{\Omega \wedge \Omega \wedge \cdots \wedge \Omega}_{n \text{ times}} = \frac{(-1)^{n/2}}{n!} \Omega^n \tag{4.2}$$

which – thanks to the non-degeneracy of the symplectic form – is nowhere vanishing. In canonical coordinates, it takes the expression

$$\Lambda = \mathrm{d}q^1 \wedge \mathrm{d}q^2 \wedge \cdots \wedge \mathrm{d}q^n \wedge \mathrm{d}q^1 \wedge \mathrm{d}q^2 \wedge \cdots \wedge \mathrm{d}p^n . \tag{4.3}$$

---

[2] This does not happen in Riemannian geometry: a manifold with non-zero curvature is not locally isometric to Euclidean space; namely, it is not flat.

A further important point is the abstract definition of *canonical transformations*, which are smooth maps preserving the symplectic structure.

Definition 4.5. Let (M_1, Ω_1) and (M_2, Ω_2) be symplectic manifolds. A diffeomorphism $f \colon M_1 \to M_2$ satisfying $f^*\Omega_2 = \Omega_1$ is called **symplectomorphism** or **canonical transformation**.

Hamiltonian systems

Definition 4.6. Let (M, Ω) be a symplectic manifold. A vector field X^H is **Hamiltonian** if there is a function $H \colon M \to \mathbb{R}$ such that

$$X^H \lrcorner \Omega = \Omega^\flat(X^H) = \mathrm{d}H \tag{4.4}$$

In a (weak) symplectic infinite-dimensional manifold, the injectivity of $\Omega^\flat$ (4.4) does not guarantee that the inverse function

$$\Omega^\sharp = \left(\Omega^\flat\right)^{-1} : T^*M \to TM$$

be defined on all the cotangent bundle. Therefore, for a given smooth function H, we cannot infer *a priori* whether a Hamiltonian vector field exists or not, but we can only affirm its uniqueness. Generally, one assumes its existence abstractly and then verify it in examples.

The triple (M, Ω, H) is a **Hamiltonian system**.

Let's focus on the finite-dimensional case. In Darboux coordinates, X^H can be computed explicitly as follows. After writing

$$X^H = \sum_{k=1}^{n} \left(a^k \frac{\partial}{\partial q^k} + b^k \frac{\partial}{\partial p^k} \right)$$

for some coefficient (a^k, b^k) to be determined, we compute

$$X^H \lrcorner \Omega = \sum_{k=1}^{n} \left(a^k \frac{\partial}{\partial q^k} + b^k \frac{\partial}{\partial p^k} \right) \lrcorner \sum_{k=1}^{n} \mathrm{d}q^k \wedge \mathrm{d}p^k = \sum_{k=1}^{n} (a^k \mathrm{d}p^k - b^k \mathrm{d}q^k).$$

On the other hand,

$$\mathrm{d}H = \sum_{k=1}^{n} \left(\frac{\partial H}{\partial q^k} \mathrm{d}q^k + \frac{\partial H}{\partial p^k} \mathrm{d}p^k \right),$$

so that we find

$$X^H = \sum_{k=1}^{n} \left(\frac{\partial H}{\partial p^k} \frac{\partial}{\partial q^k} - \frac{\partial H}{\partial q^k} \frac{\partial}{\partial p^k} \right). \tag{4.5}$$

The integral curves of this vector field are **Hamilton's Equations**

$$\begin{cases} \dot{q}^i(t) = a^i(q(t), p(t)) = \dfrac{\partial H}{\partial p^k}(q(t), p(t)) \\ \dot{p}^i(t) = b^i(q(t), p(t)) = -\dfrac{\partial H}{\partial q^k}(q(t), p(t)) \end{cases} . \tag{4.6}$$

Theorem 4.7. *If X^H is a Hamiltonian vector field on the symplectic manifold (M, Ω) and θ is the flow of X^H, θ_t is a symplectomorphism, namely $\theta_t^* \Omega = \Omega \ \forall t$.*

Since, for $t = 0$, $\theta_0(p) = \theta(p, 0)$ is the identity, this is equivalent to show that

$$\left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t_0} (\theta_t^* \Omega) = \theta_{t_0}^* \left(\pounds_{X^H} \Omega \right) = 0 \qquad \forall \ t_0 \tag{4.7}$$

or, equivalently, that

$$\pounds_{X^H} \Omega = 0 \tag{4.8}$$

Proof. The last expression is proven, by Cartan's magic formula (A.35)

$$\pounds_{X^H} \Omega = \mathrm{d}(X^H \lrcorner \Omega) + X^H \lrcorner \mathrm{d}\Omega = 0,$$

thanks to the facts that X^H is Hamiltonian and Ω is closed. □

We note that the closure of the symplectic form implies Hamiltonian flows are symplectomorphisms.

Corollary 4.8 (Liouville Theorem)**.** *A corollary of **Th.** 4.7 in finite dimensions is **Liouville Theorem**: Hamiltonian flows preserve the Liouville volume form.*

Poisson Brackets on Symplectic Manifolds

Another well-known concept from Classical Mechanics is represented by *Poisson brackets*. On symplectic manifolds, they are defined after the symplectic form.

Definition 4.9. Let (M, Ω) be a strong symplectic manifold. We define the **Poisson bracket** of two functions $F, G \colon M \to \mathbb{R}$ by

$$\{F, G\} \overset{\text{def}}{=} \Omega(X^F, X^G) = \mathrm{d}F(X^G) = X^G F = \pounds_{X^G} F \tag{4.9}$$

Poisson brackets measure the rate of change of F along the Hamiltonian flow of G; or, equivalently, thanks to the skew-symmetry of the symplectic form, it is the opposite of the rate of change of G along the Hamiltonian flow of F.

While bilinearity and skew-symmetry are obvious from the definition, Jacobi identity

$$\{F,\{G,h\}\}+\{G,\{h,F\}\}+\{h,\{F,G\}\}=0$$

can be proven, for every triplet of smooth functions F, G, h. This is a consequence of the fact that Hamiltonian flows are symplectomorphisms, or – in other words – that the symplectic form is closed (and the converse is also true: the validity of Jacobi identity implies Hamiltonian flows are symplectomorphisms). We thus have the following

Theorem 4.10. *The space of the smooth functions (modulo constant[3]) on a symplectic manifold is a* Lie algebra *under the Poisson bracket.*

Moreover, we have the following

Theorem 4.11. *In view of the identity*

$$\left[X^F,X^G\right]=-X_{\{F,G\}}, \tag{4.10}$$

the set of Hamiltonian vector fields $\mathfrak{H}(M)$ is a Lie subalgebra of the set of vector fields $\mathfrak{X}(M)$.

By means of Poisson bracket, we may also write the equations of motion of Mechanics in a concise form.

Theorem 4.12. *If θ is the flow of the Hamiltonian vector field X^H and F is a smooth function, we have*

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(F\circ\theta^{(z)}\right)(t)=\mathrm{d}F(\dot\theta^{(z)}(t))=X^H_{\theta^{(z)}(t)}F=&\\=\{F,H\}(\theta^{(z)}(t))=\left(\{F,H\}\circ\theta^{(z)}\right)(t),&\end{aligned} \tag{4.11}$$

or, equivalently,

$$\frac{\partial}{\partial t}\left(F\circ\theta\right)(z,t)=\left(\{F,H\}\circ\theta\right)(z,t). \tag{4.12}$$

---

[3] Namely, smooth functions defined up to a constant: $C^\infty(M)/\mathbb{R}$. Obviously, a constant doesn't affect the definition of Hamiltonian vector field 4.4.

One often writes **Eq.** (4.11) in the compact form[4]

$$\dot{F} = \{F, H\} \tag{4.13}$$

and calls it **equation of motion in Poisson bracket form**. Clearly, if one chooses $F = z_i$, where z_i is the ith component function in a given coordinate chart, one obtains the equations of motion for coordinates, i.e., Hamilton's Equations.

Example 4.13 (*Symplectic view of the Schrödinger Equation*)**.** Let $\mathcal{H}$ be a complex Hilbert space equipped with the inner product $\langle , \rangle$ and the symplectic form

$$\Omega(\psi_1, \psi_2) = 2\hbar \Im\langle \psi_1, \psi_2 \rangle.$$

Given a (self-adjoint) Hamiltonian operator $\hat{H}$ on $\mathcal{H}$, we write its expectation value $H \colon \mathcal{H} \to \mathbb{R}$ as

$$H(\psi) \overset{\text{def}}{=} \langle \psi, \hat{H}\psi \rangle,$$

which is also our Hamiltonian function.

With straightforward computations, one verifies that the vector field defined by

$$X^H_\psi = -\frac{i}{\hbar}\hat{H}\psi,$$

is indeed the Hamiltonian vector field associated to H:

$$\begin{aligned} \mathrm{d}H_\psi(\phi) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} \langle \psi + t\phi, \hat{H}(\psi + t\phi)\rangle = \langle \phi, \hat{H}\psi\rangle + \langle \psi, \hat{H}\phi\rangle = 2\Re\langle \phi, \hat{H}\psi\rangle = \\ = 2\Im\langle -i\hat{H}\psi, \phi\rangle = \frac{1}{\hbar}\Omega(-i\hat{H}\psi, \phi) = (X^H_\psi \lrcorner\, \Omega)(\phi) \end{aligned}$$

The integral curves of that vector field are the solutions of the *Schrödinger Equation*

$$\dot{\psi}(t) = -\frac{i}{\hbar}\hat{H}\psi(t).$$

By means of a similar calculation, one may verify that

$$\{F, G\}(\psi) = \Omega_\psi(X^F_\psi, X^G_\psi) = \frac{1}{i\hbar}\langle \psi, [\hat{F}, \hat{G}]\psi\rangle,$$

where $[,]$ is the Lie bracket (quantum commutator) of the vector fields (operators) $\hat{F}$ and $\hat{G}$ associated to their expectation values F and G.

---

[4] This is only a symbolic notation: with $\dot{F}$ we denote neither the total nor the partial derivative of F with respect to time, but precisely what is written above.

Cotangent Bundle[5]

The most important symplectic manifolds are total spaces of cotangent bundles, which have a natural symplectic structure and are also called **phase spaces**.

Definition 4.14. Let (q, p) be a point in T^*M with $q \in M$, $p \in T_q^*M$ and $\pi \colon T^*M \to M$ is the projection. We define the (**tautological**) one-form $\tau \in \Omega^1(T^*M)$ by

$$\tau_{(q,p)} = \mathrm{d}\pi^*_{(q,p)} p \tag{4.14}$$

so that, if $v \in T_{(q,p)}\left(T^*M\right)$,

$$\tau_{(q,p)}(v) = p\left(\mathrm{d}\pi_{(q,p)}(v)\right). \tag{4.15}$$

Theorem 4.15. *The tautological one-form τ is smooth and $\Omega = -\mathrm{d}\tau$ is a symplectic form on the total space of T^*M.*

We note that Ω is automatically closed, since it is exact by definition.

4.1.2 Poisson manifolds

Poisson manifolds represent a generalization of symplectic manifolds, in the sense they allow for a degeneracy we shall study below. One usually starts from Poisson brackets instead of a skew-symmetric tensor, and later defines the latter in term of the former. The first concept is the one of *Poisson algebra.*

Definition 4.16 (Poisson algebra). A **Poisson bracket** on a manifold M is a bilinear operation $\{,\}$ on the ring of smooth functions $C^\infty(M)$ such that:

$(C^\infty(M), \{,\})$ is a Lie algebra;

$\{,\}$ is a derivation in each factor, that is

$$\{FG, K\} = \{F, K\}G + F\{G, K\} \qquad \forall\ F, G, K \in C^\infty(M). \tag{4.16}$$

The analogous property

$$\{F, GK\} = \{F, G\}K + G\{F, K\}$$

follows from **Eq.** (4.16) and the skew-symmetry of the Poisson brackets.

---

[5] This paragraph may be omitted on a first reading: the contents will not be used, but we report them here for completeness.

A pair of a manifold M and a Poisson bracket on it is called **Poisson manifold**.

Then, since – fixing K – the map $G \mapsto \{G, K\}$ is a derivation, we have the following

Theorem 4.17. *Let M be a Poisson manifold. For each smooth function H on M there is a unique vector field X^H on M such that*

$$X^H G = \{G, H\} \qquad \forall\ G \in C^\infty(M) \tag{4.17}$$

and we call X^H the ***Hamiltonian vector field*** *of H.*

Theorem 4.18. *If θ is the flow of the Hamiltonian vector field X^H and F is a smooth function, we have*

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(F \circ \theta^{(z)}\right)(t) = \left(\{F, H\} \circ \theta^{(z)}\right)(t), \tag{4.18}$$

or, equivalently,

$$\frac{\partial}{\partial t}\left(F \circ \theta\right)(z, t) = \left(\{F, H\} \circ \theta\right)(z, t). \tag{4.19}$$

One often writes **Eq.** (4.18) in the compact form

$$\dot{F} = \{F, H\}. \tag{4.20}$$

So far, Poisson manifolds are very similar to symplectic manifolds. The main feature distinguishing them each other is degeneracy. Indeed, smooth functions C satisfying

$$\{C, F\} = 0 \qquad \forall F \in C^\infty(M) \tag{4.21}$$

are called **Casimir functions** of the Poisson bracket. The set of Casimir functions forms the *centre* of the Poisson algebra.[6] This means that a Casimir function is constant along the flow of every Hamiltonian vector field or, equivalently, $X^C = 0$, that is, C generates trivial dynamics.

On a symplectic manifold M, $X^C = 0$ implies $\mathrm{d}C = \Omega^\flat(X^C) = 0$ since $\Omega^\flat$ is injective (for the nondegeneracy of the symplectic form): therefore, any Casimir function is constant on connected components of M. This is an important point for the construction of a thermodynamic theory: if entropy were chosen among the

---

[6] The **centre** of a group (or algebra) is the set of elements which commute with every element of the group (or algebra).

Casimir functions of the Poisson bracket on a symplectic manifold (thus, it would certainly be conserved by reversible dynamics), it could not vary, and we would not be describing but constant-entropy states (for example, the zero-entropy states of Mechanics). In other words, there would be no room for dissipation. Instead, the degeneracy of a Poisson structure allows a Casimir function to vary on the manifold, but not along Hamiltonian vector fields (reversible dynamics). Hence, we see (and anticipate) that an irreversible contribution may be generated by a 'dissipative' vector field, which increases entropy along its integral curves, and – summed to the Hamiltonian one – yields the total dynamics of the system.

In this context, canonical transformation take the name of *Poisson transformations* and are defined analogously:

Definition 4.19. Let $(M_1, \{,\}_1)$ and $(M_2, \{,\}_2)$ be Poisson manifolds. A diffeomorphism $f\colon M_1 \to M_2$ satisfying

$$f^*\{F, G\}_2 = \{f^*F, f^*G\}_1 \qquad \forall\ F, G \in C^\infty(M_2) \tag{4.22}$$

is called **Poisson** or **canonical transformation**.

Even in this case, Hamiltonian flows are canonical transformation. In the context of symplectic manifolds, this property was assured by the closedness of the symplectic form, or – equivalently – by Jacobi identity. In the present context, we do not dispose of a closed form, but Jacobi identity still holds and implies the following

Theorem 4.20. *If X^H is a Hamiltonian vector field on the Poisson manifold $(M, \{,\})$ and θ is the flow of X^H, θ_t is a Poisson transformation, namely*

$$\theta_t^*\{F, G\} = \{\theta_t^*F, \theta_t^*G\} \quad \forall\ F, G \in C^\infty(M_2).$$

Poisson tensor

The derivation property of the Poisson bracket implies the existence of a contravariant antisymmetric two-tensor field $P \in \Gamma(T^2TM)$ such that

$$P_z(\mathrm{d}F_z, \mathrm{d}G_z) = \{F, G\}(z).$$

P is called **cosymplectic** or **Poisson tensor**.

Let $P^\sharp\colon T^*M \to TM$ the vector bundle homomorphism associated to P. Through this map, we have an even more direct rule to calculate Hamiltonian vector fields.

$$\{F,H\} = \begin{cases} X^H F = \mathrm{d}F(X^H) \\ P(\mathrm{d}F,\mathrm{d}H) = \mathrm{d}F\left(P^\sharp(\mathrm{d}H)\right) \end{cases} \qquad \Longrightarrow \qquad X^H = P^\sharp(\mathrm{d}H) \quad (4.23)$$

Lie-Poisson structures

In addition to the case of symplectic manifolds, there is another situation where Poisson brackets arise naturally.

Theorem 4.21. *Consider a Lie algebra $\mathfrak{g}$ with bracket $[,]$ and a (weakly) nondegenerate pairing between this and another Banach space $\mathfrak{h}$:*

$$\langle,\rangle\colon \mathfrak{g}\times\mathfrak{h}\to\mathbb{R}.$$

Then, $\mathfrak{h}$ is a Poisson manifold with bracket

$$\{F,G\}_h = \left\langle h, \left[\frac{\delta F}{\delta g},\frac{\delta G}{\delta f}\right]\right\rangle \qquad (h\in\mathfrak{h},\ F,G\in C^\infty(\mathfrak{h}))\,, \qquad (4.24)$$

which is often called ***Lie-Poisson bracket****. The functional derivatives are defined in the usual way:*

$$\mathrm{d}F_{h_1}(h_2) = \left\langle h_2, \left.\frac{\delta F}{\delta h}\right|_{h_1}\right\rangle \qquad \left(h_1,h_2\in\mathfrak{h},\ \frac{\delta F}{\delta f}\in\mathfrak{g}\right).$$

Symplectic stratification

The first relation from Poisson and symplectic manifold arises when the Poisson tensor is strongly nondegenerate.

Theorem 4.22. *Let $(M,\{,\})$ be a Poisson manifold. If the Poisson tensor is strongly nondegenerate, that is, it yields an isomorphism $P^\sharp\colon \mathrm{d}F\mapsto X^F$ of T^*M with TM, then (M,Ω) is a symplectic manifold with $\Omega(X^F,X^G)=\{F,G\}$ for any locally defined vector fields X^F and X^G. The closure of the symplectic form follows from Jacobi's identity.*

In the general finite-dimensional case, a Poisson manifold is the disjoint union of symplectic manifolds, in the sense we shall now discuss briefly. The subset

$P^\sharp(T^*M)$ of TM is called the **characteristic field** or **distribution** of the Poisson structure.[7] Since the rank of the Poisson operator $P^\sharp$ is not constant in general, this is not a regular distribution, but a **singular** (or **generalized**) **distribution**. The point is: when is a singular distribution integrable (in the sense of **Sec.** A.5)? While the answer for the constant-rank case is given by **Th.** A.7 (Frobenius Theorem), for singular distributions the problem is solved by the following

Theorem 4.23 (**Stefan-Sussmann Theorem**). *A smooth (singular) distribution D on M is integrable if and only if it is spanned by a set of smooth local vector fields which are* infinitesimal automorphisms *of D, that is, for each point $z \in M$, horizontal vector field $X\colon U \to D$ ($U \subseteq M$) (we say that $X \in \mathfrak{X}^D_{loc}(M)$) with flow ϕ^X_t, it happens that*

$$\mathrm{d}(\phi^X_t)_z(D_z) = D_{(\phi^X_t)_z}. \tag{4.25}$$

Another way to say it is that the distribution is ϕ^X_t-invariant $\forall X \in \mathfrak{X}^D_{loc}(M)$, or that it is invariant under the flow of every $X \in \mathfrak{X}^D_{loc}(M)$.

Theorem 4.24 (**Symplectic Stratification Theorem**). *In our case, $D = P^\sharp(T^*M)$ and $\mathfrak{X}^D_{loc}(M)$ is the set of local Hamiltonian vector fields. Thanks to Jacobi identity, flows of local Hamiltonian vector fields are Poisson maps (***), so that **Eq.** (4.25) is automatically satisfied, and $P^\sharp(T^*M)$ is integrable: its maximal integral submanifolds are called **symplectic leaves** because the induced Poisson structure on them is symplectic.*

Another way to characterise a symplectic leaf is the following

Definition 4.25. Let M be a Poisson manifold. We say that $z_1, z_2 \in P$ are on the same symplectic leaf of M if there is a piecewise smooth curve in M joining z_1 and z_2, each segment of which is a trajectory of a local Hamiltonian vector field.

Thanks to this alternative definition of foliation, we have a better feeling on why the maximal integral submanifolds (i.e., leaves) of $P^\sharp(T^*M)$ are symplectic: indeed, Hamiltonian flows preserve the Poisson tensor, thus – in particular – its rank, which is constant on the leaves. Hence, the Poisson tensor, when restricted on submanifolds of dimension equal to its rank, 'becomes' a symplectic tensor.[8]

---

[7] For a brief discussion of distributions and foliations, see **Sec.** A.5.

[8] See, e.g., [OR03] for a rigorous discussion.

4.2 Steepest Entropy Ascent

The main ideas of the SEA principle were exposed in **Sec.** 3.2 and here are resumed and developed in mathematical language.

For each thermodynamic theory, we are given:

- a (possibly infinite-dimensional) smooth real Banach manifold M (whose points are named γ);
- a set of functions (or operators) $C^i\colon M \to \mathbb{R}$, which represent the conserved properties of the process;
- another function $S\colon M \to \mathbb{R}$, the entropy of the system;
- a (strongly nondegenerate) Riemannian metric tensor field $g \in \mathfrak{T}^2(M)$, which – at every point of the manifold – takes two vectors as an input:

$$g_p\colon (u, v) \mapsto g_p(u, v)$$

  and gives a real number (which is the inner product of u and v).

In particular – as explained in **Sec.** A.4.4 –, the property of strong non-degeneracy implies that the vector bundle map (at every point p) $g_p^\flat\colon T_pM \to T_p^*M$

$$[g_p^\flat(u)](v) = g_p(u, v), \tag{4.26}$$

which brings a vector into a covector. Therefore, the inverse map $g_p^\sharp\colon T_p^*M \to T_pM$ is defined too, and we can define the *gradient* of a smooth function by

$$\mathrm{d}F_p(v) = g_p((\operatorname{grad} F)_p, v);$$

or, explicitly,

$$(\operatorname{grad} F)_p = g_p^\sharp \mathrm{d}F_p. \tag{4.27}$$

Furthermore, we suppose to know the reversible dynamics of any process, through a vector field X^H, which always takes known forms, as we will see in examples. Instead, we focus on the irreversible part of the dynamics, represented by the *dissipative vector field* Y^S. A thermodynamic process $\alpha\colon I \to M\,(I \subseteq \mathbb{R})$ is an integral curve of the sum of the reversible and dissipative vector fields:

$$\dot{\alpha}(t) = X^H_{\alpha(t)} + Y^S_{\alpha(t)} \tag{4.28}$$

Now comes the choice of the dynamics.

- We require the constancy of the properties C^i during a process: the conservation laws (also for the entropy) along the reversible vector field are taken for granted, while the conservations along the dissipative one must be assured:

  $$\frac{\mathrm{d}(C^i \circ \alpha)(t)}{\mathrm{d}t} = \mathrm{d}C^i(\dot{\alpha}(t)) = \mathrm{d}C^i\left(Y^S_{\alpha(t)}\right) = 0, \tag{4.29}$$

  where we have used the chain rule in the first step.

- We build the dissipative vector field

  $$\begin{aligned} Y^S &= \frac{1}{\tau}\left(\operatorname{grad} S - \sum\nolimits_j \beta^j \operatorname{grad} C^j\right) = \frac{1}{\tau} g^\sharp\left(\mathrm{d}S - \sum\nolimits_j \beta^j \mathrm{d}C^j\right) = \\ &= -\frac{1}{\tau} g^\sharp \mathrm{d}\Phi \end{aligned} \tag{4.30}$$

  where τ is a dimensionality constant, Φ is the potential [GÖ97]

  $$\Phi(\gamma, a, \boldsymbol{\beta}) = -S(\gamma) + aE(\gamma) + \sum\nolimits_j \beta^j \mathrm{d}C^j(\gamma), \tag{4.31}$$

  and the β^js are Lagrange multipliers related to the conservation constraints. Their values are computed by inserting **Eq.** (4.30) into **Eq.** (4.29):

  $$\frac{1}{\tau}\left[\mathrm{d}C^i(\operatorname{grad} S) - \mathrm{d}C^i\left(\sum\nolimits_j \beta^j \operatorname{grad} C^j\right)\right] = 0 \tag{4.32}$$

  $$g^\flat \operatorname{grad} C^i\,(\operatorname{grad} S) - \sum\nolimits_j \beta^j g^\flat \operatorname{grad} C^i\left(\operatorname{grad} C^j\right) = 0$$

  $$\sum_j g(\operatorname{grad} C^i, \operatorname{grad} C^j)\beta^j = g(\operatorname{grad} C^i, \operatorname{grad} S). \tag{4.33}$$

  and solving this linear system of equations.

This construction is justified in [B08; B13] and can be recognized in **Fig.** 4.1.

We see that the rate of entropy production in a process is

$$\begin{aligned} \frac{\mathrm{d}(S \circ \alpha)(t)}{\mathrm{d}t} &= \mathrm{d}S(\dot{\alpha}(t)) = \frac{1}{\tau} g\left(\operatorname{grad} S, \operatorname{grad} S - \sum\nolimits_j \beta^j \operatorname{grad} C^j\right) = \\ &= \frac{1}{\tau} g\left(\operatorname{grad} S - \sum\nolimits_i \beta^i \operatorname{grad} C^i, \operatorname{grad} S - \sum\nolimits_j \beta^j \operatorname{grad} C^j\right) \geq 0, \end{aligned} \tag{4.34}$$

where the last equality follows from **Eq.** (4.33), and the inequality from the non-negativity of the metric tensor. The Second Law of Thermodynamics is thus

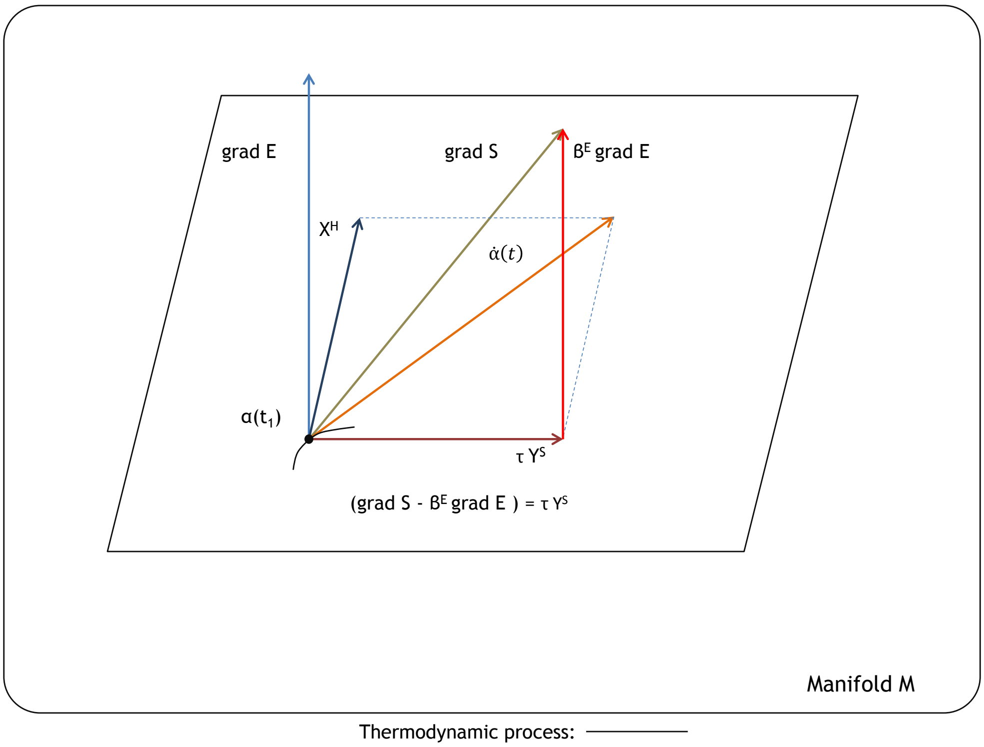

Fig. 4.1: SEA construction with energy as the only conserved quantity.

automatically satisfied.

The reader might have wondered about the meaning of the thermodynamic potential introduced in **Eq.** (4.31). This is the function indicating the equilibrium states. Indeed, optimising it equals maximising entropy subjected to the conservation constraints. At equilibrium, it assumes the value of a Legendre transform of entropy, depending on which variables are conserved. This is why it has sometimes been called *non-equilibrium Massieu function.*

4.3 GENERIC

In this section we present the simplest form of GENERIC (see [GÖ97]), the one that best resembles the SEA equation. Call M the manifold of all possible states γ ($\gamma \in M$) and build the following structure.

- There exist two potentials $H\colon M \to \mathbb{R}$ and $S\colon M \to \mathbb{R}$ representing energy and entropy respectively.

- M is a (possibly infinite-dimensional) Banach *(co)metriplectic* manifold, that is a manifold carrying two compatible structures:

- A Poisson structure describing the reversible part of the dynamics, which consists of an antisymmetric contravariant 2-tensor field $P\colon M \to T^2 T^{**}M$, which yields the Poisson operator $P^\sharp\colon T^*M \to T^{**}M$ satisfying $P^\sharp(T^*M) \subseteq T^{**}M \subseteq TM$[9]. Since P is assumed to be possibly degenerate, $P^\sharp$ is in general non-invertible (it is not a vector-space isomorphism, but only an homomorphism). To this tensor the *Poisson bracket*
$$\{F, G\}_z = P_z(\mathrm{d}F_z, \mathrm{d}G_z) \qquad \forall\ F, G \in C^\infty(M)$$
is associated on the set of smooth functions $C^\infty(M)$. This bracket must also satisfy Jacobi identity
$$\{F, \{G, K\}\} + \{G, \{K, F\}\} + \{K, \{F, G\}\} = 0 \qquad \forall\ F, G, K \in C^\infty(M),$$
which represents a further constraint on the Poisson tensor.
- A degenerate co-Riemannian structure (i.e., we have a degenerate *co-metric* instead of a nondegenerate metric) describing the irreversible dynamics.[10] This is suggested by various arguments that we don't enter in this thesis, and consists of a symmetric and non-negative definite contravariant 2-tensor field $h\colon M \to T^2 T^{**}M$, which yields the vector bundle map $h^\sharp\colon T^*M \to TM$ satisfying $h^\sharp(T^*M) \subseteq T^{**}M \subseteq TM$, in general non-invertible and sometimes called *friction operator*. Analogously, we have the *dissipative bracket*
$$[F, G]_z = h_z(\mathrm{d}F_z, \mathrm{d}G_z) \qquad \forall\ F, G \in C^\infty(M).$$

The time evolution of the state is thus represented by the curve $\alpha\colon I \to M$ ($I \subseteq \mathbb{R}$), which is the integral curve of the sum of two vector fields X^H and Y^S such that
$$\dot{\alpha}(t) = X^H_{\alpha(t)} + Y^S_{\alpha(t)}. \tag{4.35}$$

The two vector fields are:

---

[9] This condition is needed to guarantee that $P^\sharp(\mathrm{d}H)$ is a vector field, and is automatically satisfied whenever the manifold is modelled on a reflexive Banach space or, as a particular case, on a Hilbert space. See [OR03].

[10] Here is the point where a generalization has already been proposed, especially in the context of *contact manifolds*: see [He73; Gr08].

- a Hamiltonian vector field X^H, computed through

$$X^H = P^\sharp(\mathrm{d}H); \tag{4.36}$$

- a dissipative vector field Y^S, computed through

$$Y^S = h^\sharp(\mathrm{d}S). \tag{4.37}$$

Moreover, we impose the following complementary requirements of degeneracy:

- the entropy function S is chosen among the distinguished functions (Casimir functions) of the Poisson structure, that is

$$\{F, S\} = P(\mathrm{d}F, \mathrm{d}S) = \mathrm{d}F\big(P^\sharp(\mathrm{d}S)\big) = 0 \qquad \forall F \in C^\infty(M), \tag{4.38}$$

$$\text{or} \quad P^\sharp(\mathrm{d}S) = 0; \tag{4.39}$$

- analogously, the Hamiltonian function H is chosen among the distinguished functions of the dissipative structure, that is

$$[F, H] = h(\mathrm{d}F, \mathrm{d}H) = \mathrm{d}F\big(h^\sharp(\mathrm{d}H)\big) = 0 \qquad \forall F \in C^\infty(M), \tag{4.40}$$

$$\text{or} \quad h^\sharp(\mathrm{d}H) = 0; \tag{4.41}$$

- lastly, we have to assure that other conserved properties of the system are kept constants by the dynamics or, in other words, that they are distinguished functions of both brackets.

If α is a thermodynamic process and F is a smooth function, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(F \circ \alpha\right)(t) &= \dot{\alpha}(t)F = \\ &= X^H_{\alpha(t)}F + Y^S_{\alpha(t)}F = \\ &= \{F, H\}(\alpha(t)) + [F, S](\alpha(t)) = \\ &= \left(\{F, H\} \circ \alpha\right)(t) + \left([F, S] \circ \alpha\right)(t); \end{aligned} \tag{4.42}$$

or, more synthetically,

$$\dot{F} = \{F, H\} + [F, S]. \tag{4.43}$$

From the degeneracy conditions, one easily sees that, for $F = H$,

$$\dot{H} = 0, \tag{4.44}$$

which reflects the conservation of energy for an isolated system; and, for $F = S$,

$$\dot{S} = [S, S] \geq 0, \tag{4.45}$$

in accordance with the principle of entropy non-decrease.

We note in passing that the expression

$$h^\sharp(\mathrm{d}S)$$

is similar in form to **Eq.** (4.27). The difference is the degeneracy of the pairing, which hinders us from disposing of an expression like **Eq.** (4.27); in other words, there is no one-to-one correspondence between covectors and vectors. Howsoever, by means of a further condition we may associate to it the meaning of *gradient*, in the sense we will discuss later.

4.4 Discussion

Thanks to having rewritten the SEA method with a more rigorous differential geometry language, we made it more comparable to the GENERIC model.

First of all, since the reversible part of the dynamics in the SEA model is not rationalized as in GENERIC, but only described case by case, we see that the Poisson structure may be fully imported from the second to the first dynamics without changes. Hence, we shall focus on the dissipative part, analyse similarities and differences between the two models and highlight the aspects not completely clear that deserve to be further developed.

4.4.1 Purposes

In their original article [GÖ97], Grmela and Öttinger declared the two main purposes of GENERIC:

1. to reproduce known equations of motion of known physical theories by casting them in a single abstract form;

2. to suggest new equations for new thermodynamic theories dealing with complex systems.

The goal of the SEA method [B13] applied to meso- and macroscopic systems was similar:

1. to show that a broad selection of known theoretical frameworks for the description of non-equilibrium thermodynamics at various levels of description can all be unified when viewed as implementations of the SEA principle;
2. to provide rigorous mathematical formalization of the so-called Maximum Entropy Principle (MEP) Principle, as an attempt to clarifying its meaning, scope and domain of validity;
3. to propose a formalization of known theories which reduces to the linear theories in the proximity of equilibrium, entailing Onsager reciprocity. Hence, showing that such theories are indeed SEA with respect to any metric, which at equilibrium reduces to a generalized Onsager conductivity matrix.

4.4.2 Geometric structure

By construction,

$$Y^S F = \mathrm{d}F(Y^S) = \mathrm{d}F\left(h^\sharp(\mathrm{d}S)\right) = h(\mathrm{d}F, \mathrm{d}S) = [F, S] = 0;$$

for every distinguished function F of the dissipative bracket. That is, the degeneracy condition (4.40) imposed on the geometrical structure of the manifold implies that every distinguished function of the dissipative bracket cannot vary along the dissipative vector field: the entropy ‘gradient’ $h^\sharp(\mathrm{d}S)$ is automatically parallel to the level sets of all the distinguished functions (thus, for example, to the level set of energy). The information of constancy of the conserved functions is contained already in the co-metric tensor h. Conversely, the SEA model is built on a manifold with less structure and, consequently, there cannot be preliminary assumptions on the functions related to the conserved properties, i.e., that they are distinguished functions of some structure. As a consequence, in the SEA method the conservations must be imposed later.

The GENERIC dissipative structure is a ‘weak version’ of the SEA one: by looking at **Fig.** 4.1 and **Fig.** 4.4, we see that in the GENERIC picture there cannot be any dissipative vector outside the metric leaf for the degeneracy of the dissipative bracket. In the SEA picture the situation is different, but can easily be transformed into the GENERIC one by simply allowing for a degeneracy of the metric tensor.

The geometric structure of the dissipative dynamics of GENERIC is similar to a sub-Riemannian structure. The difference lies in the nature of the distribution $h^\sharp(T^*M)$: whereas, in the sub-Riemannian case, this is a regular distribution (i.e., of constant dimension), in the case of GENERIC this may not happen (since the rank of $h^\sharp$ may vary from point to point), and one is faced to a *generalized* (or *singular*) distribution, whose dimension is not constant.[11] As outlined in **Par.** 4.1.2, a generalized distribution is integrable if and only if it is generated by a family of smooth vector fields, and is invariant with respect to their flows. In Classical Mechanics, the condition that assures this integrability is Jacobi identity, since it forces Hamiltonian flows to be canonical transformation (Poisson maps), that is, to preserve the Poisson structure. Indeed, as in **Th.** 4.7, since – for $t = 0$, the flow of X^H is the identity – we have to prove that

$$\pounds_{X^H} P = 0. \tag{4.46}$$

Proof. We make use of the identities **Eqs.** (A.26) and (A.36) and Jacobi's (JI),

$$\begin{aligned}
&\pounds_{X^H}(P(\mathrm{d}F, \mathrm{d}G)) \overset{(A.26)}{=} \\
&= (\pounds_{X^H} P)(\mathrm{d}F, \mathrm{d}G) + P(\pounds_{X^H} \mathrm{d}F, \mathrm{d}G) + P(\mathrm{d}F, \pounds_{X^H} \mathrm{d}G) \overset{(A.36)}{=} \\
&= (\pounds_{X^H} P)(\mathrm{d}F, \mathrm{d}G) + P(\mathrm{d}(\pounds_{X^H} F), \mathrm{d}G) + P(\mathrm{d}F, \mathrm{d}(\pounds_{X^H} G)) = \\
&= (\pounds_{X^H} P)(\mathrm{d}F, \mathrm{d}G) + P(\mathrm{d}\{F, H\}, \mathrm{d}G) + P(\mathrm{d}F, \mathrm{d}\{G, H\}) = \\
&= (\pounds_{X^H} P)(\mathrm{d}F, \mathrm{d}G) + \{\{F, H\}, G\} + \{F, \{G, H\}\} \overset{(JI)}{=} \\
&= (\pounds_{X^H} P)(\mathrm{d}F, \mathrm{d}G) + \{\{F, G\}, H\} = \\
&= (\pounds_{X^H} P)(\mathrm{d}F, \mathrm{d}G) + \pounds_{X^H}(P(\mathrm{d}F, \mathrm{d}G)).
\end{aligned}$$

□

What about the dissipative part of the dynamics? Going through the same

---

[11] Hereafter, we shall consider finite-dimensional manifolds, because the theorems we will mention are valid for this case.

steps, we notice that, since

$$
\begin{aligned}
&\pounds_{Y^S}(h(\mathrm{d}F,\mathrm{d}G)) \overset{(A.26)}{=}\\
&= (\pounds_{Y^S} h)(\mathrm{d}F,\mathrm{d}G) + h(\pounds_{Y^S}\mathrm{d}F,\mathrm{d}G) + h(\mathrm{d}F,\pounds_{Y^S}\mathrm{d}G) \overset{(A.36)}{=}\\
&= (\pounds_{Y^S} h)(\mathrm{d}F,\mathrm{d}G) + h(\mathrm{d}(\pounds_{Y^S} F),\mathrm{d}G) + h(\mathrm{d}F,\mathrm{d}(\pounds_{Y^S} G)) =\\
&= (\pounds_{Y^S} h)(\mathrm{d}F,\mathrm{d}G) + h(\mathrm{d}[F,S],\mathrm{d}G) + h(\mathrm{d}F,\mathrm{d}[G,S]) =\\
&= (\pounds_{Y^S} h)(\mathrm{d}F,\mathrm{d}G) + [[F,S],G] + [F,[G,S]]\,,
\end{aligned}
$$

flows of the dissipative vector field do not preserve the cometric tensor. For this reason, the distribution $h^\sharp(T^*M)$ is not integrable, and we do not even have the notion of distance, as we would have in the sub-Riemannian (constant rank) case (*Carnot-Carathéodory distance*: see, e.g., [Mon06]).

Instead, if we endowed the dissipative structure with the *Leibniz identity*

$$[[F,G]\,,K] = [F,[G,K]] + [[F,K]\,,G]\,, \tag{4.47}$$

which is a generalization of Jacobi identity for non skew-symmetric brackets, we see that dissipative flows preserve the cometric tensor, thus guaranteeing the integrability of the generalized distribution $h^\sharp(T^*M)$ to *metric leaves*. On metric leaves, one has a (non-degenerate) metric, can calculate distances with it, and define gradients by **Eq.** (4.27). In this case, one may also interpret GENERIC dynamics as a SEA dynamics on metric leaves.

At the present time, we do not know the implications of this possible assumption from the physical standpoint, nor whether it is satisfied in known examples of dynamics (such as the Boltzmann Equation) because of its complexity of verification. We only recognize its mathematical coherence and notice that it leads to a more symmetric treatment of the two structures, maybe paving the way to an unification into a single one.

It is interesting to understand the relationship between metric leaves, where metriplectic dynamics takes place, and symplectic leaves, where Hamiltonian dynamics takes place. The first ones are surfaces with constant energy, while the second ones are surfaces with constant entropy (because Hamiltonian dynamics is reversible). As a consequence of this, the intersection of symplectic leaves on a metric leaf produces isentropic lines, and the reversible vector X^H is always contained in a symplectic leaf. This is illustrated in **Fig.** 4.4, where the constancy of

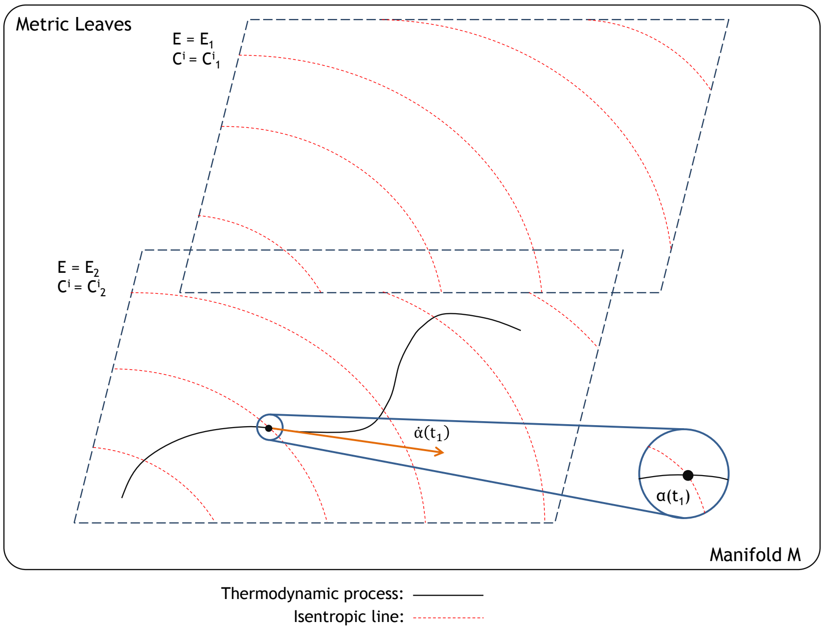

Fig. 4.2: Metric leaves in a manifold: GENERIC dynamics takes place on a single metric leaf.

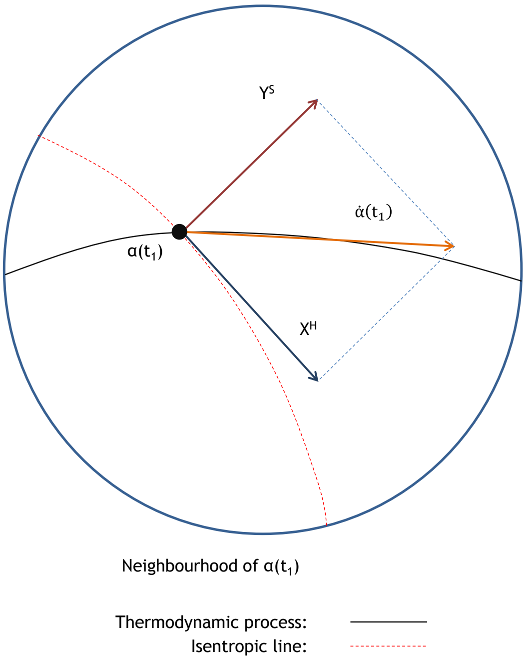

Fig. 4.3: GENERIC dynamics on a metric leaf, with the velocity vector decomposed in a reversible or Hamiltonian part and a irreversible or dissipative one.

all the other conserved quantities is assumed.

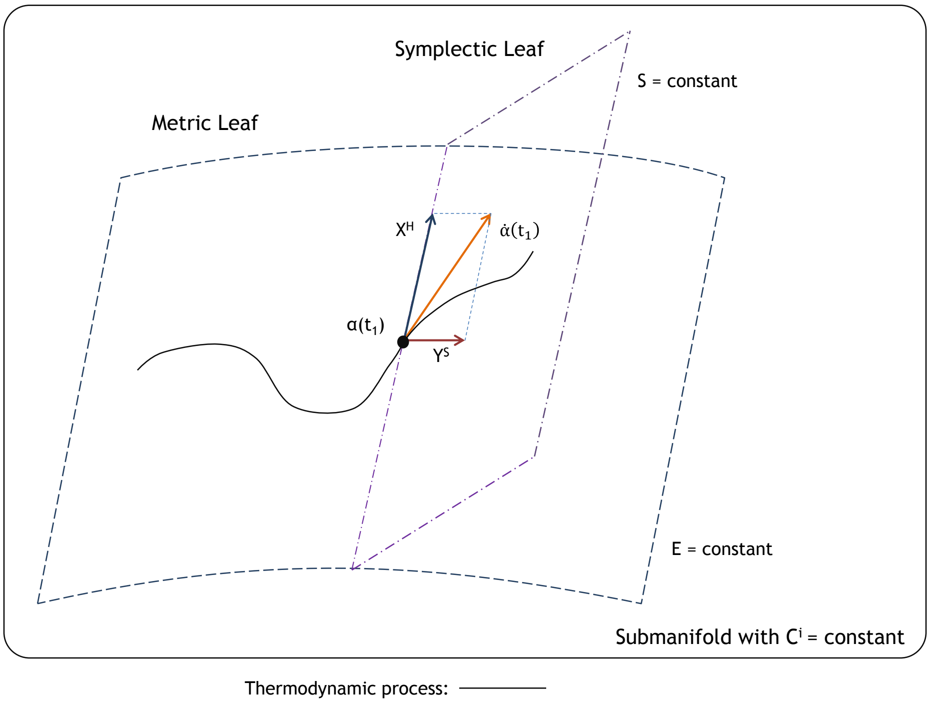

Fig. 4.4: Intersection of a metric leaf with a symplectic leaf and correlation with the evolution of a thermodynamic process.

4.4.3 MEPP

Let's resume our last statement and suppose, for a moment, that there is only dissipative dynamics, confining ourselves on a metric leaf, where the degenerate contravariant tensor h is restricted into the non-degenerate one h_L. In this way, we can build the corresponding covariant metric tensor g_L, which acts on vectors as $g_L(u, v) = h_L(h^\flat_L(u), h^\flat_L(v))$[12], take a unit vector v ($g_L(v, v) = 1$) and wonder in which direction the directional derivative of entropy is greater. By definition (4.27) of gradient of a smooth function (given the non-degenerate bilinear form g_L) and the Cauchy-Schwarz inequality,

$$\begin{aligned} |\mathrm{d}S(v)|^2 = |g_L(\operatorname{grad}_L S, v)|^2 \leq \\ \leq g_L(\operatorname{grad}_L S, \operatorname{grad}_L S)\ g_L(v, v) = \|\operatorname{grad}_L S\|^2_L \,. \end{aligned} \tag{4.48}$$

---

[12] In finite dimension, it has matrix $[g_{L,ij}] = [h^{ij}_L]^{-1}$. For a more rigorous treatment of this procedure for the case of symplectic leaves, see [OR03].

That is, the absolute value of the directional derivative is always smaller than the norm of the gradient vector, and reaches its maximum value when $v = \dfrac{\operatorname{grad}_L S}{\|\operatorname{grad}_L S\|_L}$. The restriction of the total entropy 'gradient' $h^\sharp(\mathrm{d}S)$ to the metric leaf is indeed $\operatorname{grad}_L S$[13], since the evolution of the dissipative part takes place uniquely on the metric leaf by definition: the entropy gradient has zero components outside of it.[14]

Therefore, any nonequilibrium dynamics that can be written in GENERIC form and satisfying Leibniz identity (4.47) is automatically SEA.

4.4.4 Relaxation time

Once the state has been chosen, namely the manifold where a thermodynamic process occurs, in the SEA concept we may identify the following two degrees of freedom:

- distance between states (or, more generally, the metric);
- trajectory of the state representative (or, locally, direction of evolution on the manifold).

It is clear that the rate of evolution is regulated by the metric, since the velocity of a curve is the scalar

$$\|\dot{\alpha}(t)\| = \sqrt{g(\dot{\alpha}(t), \dot{\alpha}(t))},$$

which can be scaled by a constant in the metric tensor.

GENERIC and SEA show that many thermodynamic theories may be given a metric with respect to which the dynamics is fully reproduced. If, indeed, one selects the GENERIC or SEA assumption, the problem falls entirely onto the choice of the metric.

A question might arise at this point: is there a metric natural to the manifold of each nonequilibrium framework? That is, can one forget about the "standard" equations of motion in each framework (GENERIC always finds a metric when it starts from the equations of motion) and find another form for them through a new metric not imposed by the "standard" equations themselves? Perhaps, this is a point where the emerging field of Information Geometry could give a hint.

In the previous formulation of the SEA model, in the IQT framework, a metric (the Fisher-Rao metric) was chosen *ab initio*: this was inspired by the fact that

---

[13] This vector is sometimes called *horizontal gradient*, because it belongs to the distribution $\mathrm{Im}(h^\sharp)$.

[14] It has no *vertical* components.

the state is, essentially, a probability measure. On the other hand, because of the novelty of the theory, there was no equation having to be reproduced by the model, as in the other meso- and macroscopic frameworks. Due to this fact, the relaxation time τ was allowed to be a functional of the state, in order to keep the degree of freedom *rate of evolution* in the dynamics.

But now that the metric is not chosen *a priori*, the SEA relaxation time may be merged in the metric tensor (indeed, the operator $\hat{L} \overset{\text{def}}{=} g^\sharp/\tau$ introduced in [B13] does so), since it simply scales the intensity of the dissipative vector field in the direction of the entropy gradient. That's why, in this work, we assumed it to be only a dimensionality constant.

4.4.5 Reversible-irreversible coupling

One the most interesting successes of metriplectic dynamics is the description of the mutual relation between the reversible and irreversible parts of the dynamics. Indeed, let's do a comparison among the dynamic theories we have met so far:

Theory	Equation	Additional constraints
Onsager's dynamics	$\dot{\boldsymbol{\alpha}} = \boldsymbol{L}\dfrac{\partial S}{\partial \boldsymbol{\alpha}}$	$\mathrm{d}H = 0, \quad \mathrm{d}C^i = 0$
Edelen's theory	$\begin{aligned} \boldsymbol{j} &= \boldsymbol{j}_N + \boldsymbol{j}_D \\ &= \boldsymbol{U} + \boldsymbol{\nabla}\Phi \end{aligned}$	$\boldsymbol{X}\cdot\boldsymbol{U} = 0$
GENERIC	$\begin{aligned} \dot{\boldsymbol{\alpha}} &= \dot{\boldsymbol{\alpha}}_{\text{rev}} + \dot{\boldsymbol{\alpha}}_{\text{irr}} \\ &= P^\sharp \mathrm{d}H + h^\sharp \mathrm{d}S \end{aligned}$	$P^\sharp \mathrm{d}C^i = 0, \quad h^\sharp \mathrm{d}C^i = 0$ $P^\sharp \mathrm{d}S = 0, \quad h^\sharp \mathrm{d}H = 0$
Steepest-Entropy-Ascent	$\begin{aligned} \dot{\boldsymbol{\alpha}} &= \dot{\boldsymbol{\alpha}}_{\text{rev}} + \dot{\boldsymbol{\alpha}}_{\text{irr}} \\ &= \dot{\boldsymbol{\alpha}}_{\text{rev}} - g^\sharp \mathrm{d}\Phi \end{aligned}$	

Recalling the definition (4.31) of the thermodynamics potential and exploiting the degeneracy conditions, GENERIC may be also rewritten as

$$\dot{\boldsymbol{\alpha}} = \dot{\boldsymbol{\alpha}}_{\text{rev}} + \dot{\boldsymbol{\alpha}}_{\text{irr}} = \frac{1}{a}P^\sharp \mathrm{d}\Phi - h^\sharp \mathrm{d}\Phi \tag{4.49}$$

This was called by Grmela a *nonlinear Onsager-Casimir equation.*

In its reasoning, which we outlined in **Subsec.** 2.1.4, Onsager interprets $\dot{\boldsymbol{\alpha}}$ as a flux. Indeed, this represents the rate of evolution of the state as pulled out from equilibrium. If we interpret the operator $\boldsymbol{L}$ (or $h^\sharp$ and $g^\sharp$) as a kind of 'spring constant', $\boldsymbol{X} \overset{\text{def}}{=} \mathrm{d}\Phi$ takes the role of a thermodynamic force, i.e., an indicator of how much a system is far from equilibrium (think of the fact that the derivative of the potential vanishes at equilibrium!). In the case of Onsager, the force $\dfrac{\partial S}{\partial \boldsymbol{\alpha}}$ vanishes at equilibrium since the constraints of energy and number of particle conservation are automatically satisfied (he considers a micro-canonical ensemble).

In light of this, we are open to two different interpretations of the work of Edelen:

- Edelen's total flux may be interpreted as $\dot{\boldsymbol{\alpha}}$, so that the two components (dissipative and non-dissipative) may be naturally associated to $\dot{\boldsymbol{\alpha}}_{\text{rev}}$ and $\dot{\boldsymbol{\alpha}}_{\text{irr}}$ in the following way (in this form of GENERIC, we take the quadratic potential that leads to linear phenomenological relations):

  $$\boldsymbol{j}_D = \boldsymbol{L}\boldsymbol{X} \quad \longleftrightarrow \quad \dot{\boldsymbol{\alpha}}_{\text{irr}} = -h^\sharp \mathrm{d}\Phi = h^\sharp \mathrm{d}S \tag{4.50}$$

  $$\boldsymbol{j}_N = \boldsymbol{U} \quad \longleftrightarrow \quad \dot{\boldsymbol{\alpha}}_{\text{rev}} = \frac{1}{a} P^\sharp \mathrm{d}\Phi = P^\sharp \mathrm{d}H \tag{4.51}$$

  Edelen's condition $\boldsymbol{X} \cdot \boldsymbol{U} = 0$ becomes

  $$\mathrm{d}\Phi(P^\sharp \mathrm{d}H) = \{-S + aH, H\} = 0, \tag{4.52}$$

  which is automatically verified for the degeneracy condition (4.39) and the skew-symmetry of the Poisson bracket.

  Hence, if Edelen suggestion to consider non-dissipative phenomena is interpreted as above, Onsager dynamics appears to contain only the irreversible component, and could be extended as follows:

  $$\dot{\boldsymbol{\alpha}} = \boldsymbol{P}\frac{\partial H}{\partial \boldsymbol{\alpha}} + \boldsymbol{L}\frac{\partial S}{\partial \boldsymbol{\alpha}}, \tag{4.53}$$

  which is basically the GENERIC.

- The second interpretation is more natural to Edelen's presentation itself, and it is advocated by Goddard in [Go11], or Ostoja-Starzewski and Zubelewicz in [OZ11]. Consider a kind of 'non-dissipative irreversible' contribution due to some symmetric part of the friction operator:

  $$\dot{\boldsymbol{\alpha}}_{\text{irr}} = -h^\sharp_{\text{sym}} \mathrm{d}\Phi - h^\sharp_{\text{skew}} \mathrm{d}\Phi. \tag{4.54}$$

> The skew-symmetric part obviously does not contribute to the entropy production. Taking a look to **Eq.** (4.49), we don't see why such a 'non-dissipative irreversible' contribution should not directly incorporated in the *reversible* dynamics. This is why we prefer the first interpretation.
>
> An alternative explanation by Grmela and Öttinger in [GÖ97] is that they "*consider isolated systems (no external magnetic fields) and [they] assume that a sufficiently detailed level of description has been chosen*".

4.4.6 Thermodynamic forces and fluxes

As we see from the previous considerations, the quantity that is 'maximized' is the total entropy production, and not a local one. A local MEPP was originally proposed by Ziegler in [Zi57] and brought to the fore by Martyushev and Seleznev in [MS06] as a possible 'first principle' entailing the Second Law, linear phenomenological relations near equilibrium, and Onsager reciprocity; in particular it constitutes an extension of Linear Irreversible Thermodynamics (LIT).

In order to see the consistency of the SEA and GENERIC approaches to LIT, we have to identify thermodynamic forces and fluxes. Let's concentrate ourselves onto GENERIC. Öttinger in [Ö05], citing the article [E98] by Edwards, decomposes the friction operator into

$$h^\sharp = B \circ D \circ B^T, \tag{4.55}$$

where $B\colon Z \to TM$ is a "*mechanical component*", $B^T\colon T^*M \to Z^*$ is its transpose map, $D\colon Z^* \to Z$ contains all the dynamic material information and may be identified with the phenomenological matrix of classical nonequilibrium thermodynamics ($L = D$), and Z is a vector space. Then, he defines the thermodynamic forces

$$X \overset{\text{def}}{=} -B^T \mathrm{d}S \in Z^*, \tag{4.56}$$

so that the classical form of linear irreversible thermodynamics arises naturally:

$$\begin{aligned}\dot{S} = [S,S] = \mathrm{d}S(h^\sharp \mathrm{d}S) = \mathrm{d}S(B \circ L \circ B^T \mathrm{d}S) = \\ = \mathrm{d}S \circ B \circ L \circ B^T(\mathrm{d}S) = (B^T \mathrm{d}S)(LB^T \mathrm{d}S) = X(LX)\end{aligned} \tag{4.57}$$

and the (Onsager) symmetry and non-negative definiteness of L are deduced from the analogous properties of $h^\sharp$ (cf. **Eq.** (4.55)). Defining also the thermodynamic fluxes

$$J = LX, \tag{4.58}$$

he realizes that the structure of this GENERIC version carries the one of Linear Irreversible Thermodynamics. The transformation B^T basically moves some differential operator from M to B, thus to X, allowing us to recognize the forces of Classical Non-Equilibrium Thermodynamics.

In CNET, however, the thermodynamic force takes the role of how far a system is from equilibrium. The correct quantity to be optimized to find equilibrium states is not the unconstrained entropy, but the potential (4.31)

$$\Phi(x, a, \boldsymbol{\beta}) = -S(x) + aE(x) + \sum\nolimits_i \beta^i C^i(x), \tag{4.59}$$

whose differential $\mathrm{d}\Phi$ vanishes at equilibrium states.

In [Ö05] no explanation on how to make the decomposition (4.55) is given, but only a few examples are presented, and the only imposed constraint is

$$B^T \mathrm{d}E = 0. \tag{4.60}$$

If the following further constraints are added:

$$B^T \mathrm{d}C^i = 0 \tag{4.61}$$

for each distinguished function of the dissipative bracket, we see that

$$-B^T \mathrm{d}\Phi = B^T \mathrm{d}S, \tag{4.62}$$

and we are allowed to treat $-B^T \mathrm{d}S$ as a thermodynamic force, since

$$J = -LB^T \mathrm{d}\Phi = LB^T \mathrm{d}S. \tag{4.63}$$

We note that, in this reasoning, no reference is made to physical space $\mathbb{R}^3$, no local force $X = X(\boldsymbol{r})\,(\boldsymbol{r} \in \mathbb{R}^3)$ emerges, but a force is a covector field on a manifold, hence function of the state. For example, if we choose a set of fields as state (as in hydrodynamic: see **Sec.** 5.1)

$$x = \tilde{x}(\boldsymbol{r}) = (x_1(\boldsymbol{r}), x_2(\boldsymbol{r}), \ldots, x_n(\boldsymbol{r})) \qquad \left(\boldsymbol{r} \in \mathbb{R}^3\right),$$

a thermodynamic force is function of the entire fields, and is a field itself. **Eq.** (4.57) becomes

$$\frac{\mathrm{d}S}{\mathrm{d}t} = [S,S] = \left\langle \frac{\delta S}{\delta x}, h^{\sharp}\frac{\delta S}{\delta x} \right\rangle = \left\langle C_M^T \frac{\delta S}{\delta x}, L C_M^T \frac{\delta S}{\delta x} \right\rangle = \langle X, LX \rangle ; \tag{4.64}$$

and, if one doesn't make the integration implied by the scalar product in **Eq.** (4.64), but performs only the scalar product on the discrete indexes

$$X^T L X, \tag{4.65}$$

so that

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \int X^T(\boldsymbol{r}) L(\boldsymbol{r}) X(\boldsymbol{r}) \mathrm{d}^3\boldsymbol{r}, \tag{4.66}$$

he may be led to recognize the *local* entropy production $\sigma(\boldsymbol{r})$ that appears in

$$\frac{\mathrm{d}S}{\mathrm{d}t} = \int \sigma(\boldsymbol{r}) \mathrm{d}^3\boldsymbol{r}. \tag{4.67}$$

by

$$\sigma(\boldsymbol{r}) = X^T(\boldsymbol{r}) L(\boldsymbol{r}) X(\boldsymbol{r}) \tag{4.68}$$

and infer the non-negativity of $\sigma(\boldsymbol{r})$ by the one of L. However, since the domain of integration is all the region of $\mathbb{R}^3$ occupied by the system (it is not arbitrary, because it is closed by choice!), the above identification is not fully justified: thus, the local entropy production is not assured to be non-negative. Moreover, it is not clear why one should distinguish between discrete and continuous indexes in this setting of global reasoning (the state is the collection of fields, there is not a state at each point of $\mathbb{R}^3$).

As a final question, we ask ourselves whether the transformation rule (4.56) is a change of variables from $x \in M$ (the coordinates of the manifold where we started the construction) to some $y \in N$, which – in general – is not known, since the properties of the transformation are not known as well. If we were able to invert B, we could invert the following

$$\dot{\alpha}(t) = \left[P\mathrm{d}H + (B \circ L \circ B^T)\mathrm{d}S\right]_{\alpha(t)} = \left[P\mathrm{d}H + BJ\right]_{\alpha(t)}, \tag{4.69}$$

to get

$$[\dot{y}(t) =] \quad \dot{\beta}(t) \equiv B^{-1}_{\alpha(t)}\dot{\alpha}(t) = \left[B^{-1}P\mathrm{d}H + J\right]_{\alpha(t)} = U_{\alpha(t)} + J_{\alpha(t)}, \tag{4.70}$$

and go back to y.[15] For instance, in the case of hydrodynamics, B is not even invertible: hence, the representation in terms of forces and fluxes of LIT is not fully equivalent to the GENERIC evolution equation. Furthermore, we doubt its usefulness: in the spirit of SEA and GENERIC, the thermodynamic force may be taken as $\mathrm{d}\Phi$ (or $\mathrm{d}S$ for GENERIC), and the corresponding flux as $h^\sharp \mathrm{d}\Phi$ ($= h^\sharp \mathrm{d}S$ for GENERIC), without need for making a transformation of variables.

As we have just seen, the GENERIC approach in the present form (namely, for closed system only) doesn't claim to deal with a maximization of the local entropy, nor assure its non-negativity. In fact, the really fundamental quantity is the total entropy S, which is assured not to decrease for closed system. From the point of view of GENERIC, the most promising mathematical tool for modelling open systems resides in the Dirac structures, which represent an extension of Poisson structures and a geometrical framework to treat constraints, i.e., driven systems (see [JÖ04; Ö05]).

In the SEA model, an attempt to treat open systems was made in [B08; B13], and a simple phenomenological model for quantum systems were developed in [B09]). An interesting parallel could be played out between this attempt and the more mathematical framework of Dirac structures.

Another interesting way of making thermodynamic forces and fluxes emerge from the geometric structure is in [Gr08], in the context of contact manifolds.

4.5 (Free) thoughts on brackets

This section may be omitted since it does not contain information needed to understand what will be exposed in the subsequent chapters. Moreover, some notions have not been introduced in the remaining part of the thesis. We presently do not know the physical implications of the following mathematical statements, but we only want – sometimes – to stress the mathematical consistence of our reasoning, and – other times – only to suggest possible future developments, even if we have not the time and the necessary mathematical tools to tackle the topics.

---

[15] More rigorously, as in **Sec.** A.5, we could wonder if $B^{-1}(TM)$ is the tangent bundle of some manifold N.

4.5.1 Leibniz-Leibniz bracket

Under some conditions, we may think of the Poisson and cometric tensors as a unique tensor in the following way.

Given the entropy S, the energy H, and all the conserved properties C^j, we introduce, as before, the potential

$$\Phi = -S + aH + \sum_j \beta^j C^j. \tag{4.71}$$

Then, define the tensor field

$$T = P - h \qquad \big(: M \to T^2 T^{**} M\big)\,, \tag{4.72}$$

yielding the vector bundle homomorphism $T^\sharp \colon T^*M \to T^{**}M$ that satisfy $P^\sharp(T^*M) \subseteq T^{**}M \subseteq TM$, and associate to this the *Leibniz-Leibniz bracket*[16]

$$\begin{aligned}\Big(F, G\Big)_z = T_z(\mathrm{d}F_z, \mathrm{d}G_z) = P_z(\mathrm{d}F_z, \mathrm{d}G_z) + h_z(\mathrm{d}F_z, \mathrm{d}G_z) = \\ = \{F, G\}_z + [F, G]_z \qquad \forall\ F, G \in C^\infty(M). \quad (4.73)\end{aligned}$$

on the set of smooth functions $C^\infty(M)$. This bracket must also satisfy Leibniz identity

$$\Big(\Big(F, G\Big), K\Big) = \Big(F, \Big(G, K\Big)\Big) + \Big(\Big(F, K\Big), G\Big) \qquad \forall\ F, G, K \in C^\infty(M), \tag{4.74}$$

which represents a further constraint on the tensor T. As discussed in **Sec.** 4.4.2, adding the requirement of the Leibniz identity on the symmetric bracket makes dissipative processes preserve the cometric tensor. Imposing the Leibniz identity on the bracket (4.73) is a further generalization. This does not imply Jacobi and Leibniz identities on the Poisson and dissipative brackets. Not even the converse is true: Jacobi and Leibniz identity do not entail **Eq.** (4.74), as we can see by

---

[16] We decided to call this bracket in this way because it satisfies both Leibniz identity for derivations and Leibniz identity (4.47).

expanding **Eq.** (4.74) in terms of Poisson and dissipative brackets:

$$\begin{aligned}
&\{\{F,G\},K\}+[\{F,G\},K]+\{[F,G],K\}+[[F,G],K]=\\
&\qquad=\{F,\{G,K\}\}+[F,\{G,K\}]+\{F,[G,K]\}+[F,[G,K]]+\\
&\qquad+\{\{F,K\},G\}+[\{F,K\},G]+\{[F,K],G\}+[[F,K],G]\,.
\end{aligned}\tag{4.75}$$

However, the conditions are compatible: satisfaction of Leibniz identity by Poisson and dissipative brackets may be seen as a 'first step' towards **Eq.** (4.75); the other 'steps' are the mutual Leibniz identities

$$[\{F,G\},K]=[F,\{G,H\}]+[\{F,K\},G]$$

and

$$\{[F,G],K\}=\{F,[G,H]\}+\{[F,K],G\}\,.$$

Recalling the arguments of **Subsec.** 4.4.2, all these relations would guarantee that Hamiltonian flows preserve the dissipative structure, that dissipative flows preserve the Poisson structure, hence that thermodynamic processes preserve the whole tensor field defined in **Eq.** (4.72).

The time evolution of the state is thus represented by the curve $\alpha\colon I\to M\;(I\subseteq\mathbb{R})$, which is the integral curve of the (structure-preserving) vector field Z^{Φ} such that

$$\dot{\alpha}(t)=Z^{\Phi}_{\alpha(t)},\tag{4.76}$$

and computed through

$$Z^{\Phi}=T^{\sharp}(\mathrm{d}\Phi).\tag{4.77}$$

4.5.2 (Complex) SEA-Nambu dynamics

Steepest-Entropy-Ascent dynamics is rather similar to Nambu dynamics in that, unlike symplectic and Poisson ones, it adopts as generators not only the Hamiltonian, but also all the other conserved functions, which are thus brought to the same level of energy. In addition, however, also entropy is present to generate the irreversible part of the time evolution, so that, in metriplectic systems, the latter is associated to a symmetric tensor or bracket. On the other hand, as one may learn from Kähler geometry, a complex structure J can provide the link between a symmetric and a skew-symmetric tensor in the following way:

$$g(v, w) = \omega(v, Jw),$$

where (as in Kähler manifolds) a symplectic and Riemannian tensor have been taken as examples.[17] Therefore, we have the feeling (but, so far, we have not developed the necessary mathematical instruments to tackle the problem) that energy and entropy could be unified in a single entity such as $E + iS$, and in such a way that the full skew-symmetry of the Nambu bracket united to the 'complexification' of entropy might assure both constancy of the conserved properties and non-decrease of entropy along a thermodynamic process.

4.5.3 Complex GENERIC

The same idea might be applied to the GENERIC formalism. The difference is degeneracy, in such a way that the only generators are energy and entropy, and all the conserved functions are distinguished functions of the bracket(s). It would be interesting to look whether the complex structure capable of capturing the features of GENERIC is already present in the literature, or it have to be designed for this occasion.

---

[17] Think of the fact that i times a hermitian matrix gives an anti-hermitian one.

References

[A89] V. I. Arnol'd. *Mathematical Methods of Classical Mechanics.* 2nd ed. (1st ed., Nauka, 1974). Springer-Verlag, 1989 (cit. on p. 80).

[B08] G. P. Beretta. "Modeling non-equilibrium dynamics of a discrete probability distribution: General rate equation for maximal entropy generation in a maximum-entropy landscape with time-dependent constraints". In: *Entropy* 10.3 (2008), pp. 160–182 (cit. on pp. 94, 110).

[B09] G. P. Beretta. "Nonlinear quantum evolution equations to model irreversible adiabatic relaxation with maximal entropy production and other nonunitary processes". In: *Reports on Mathematical Physics* 64.1-2 (2009), pp. 139–168 (cit. on p. 110).

[B13] G. P. Beretta. "Steepest-Entropy-Ascent and Maximal-Entropy-Production Dynamical Models of Irreversible Relaxation to Stable Equilibrium from Any Non-Equilibrium State. Unified Treatment for Six Non-Equilibrium Frameworks". In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013.* Ed. by M. Pilotelli and G. P. Beretta. Brescia: Cartolibreria Snoopy, 2013, pp. 100–109 (cit. on pp. 81, 94, 99, 105, 110).

[B81] G. P. Beretta. "On the general equation of motion of quantum thermodynamics and the distinction between quantal and nonquantal uncertainties". thesis. 1981 (cit. on p. 81).

[C09] C. Carathéodory. "Untersuchungen über die Grundlagen der Thermodynamik". In: *Mathematische Annalen* 67.3 (1909), pp. 355–386 (cit. on p. 80).

[E98] B. J. Edwards. "An analysis of single and double generator thermodynamic formalisms for the macroscopic description of complex fluids". In: *Journal of Non-Equilibrium Thermodynamics* 23.4 (1998), pp. 301–333 (cit. on p. 107).

[F05] D. Fish. "Metriplectic Systems". thesis. Portland State University, 2005 (cit. on p. 80).

[Go11] J. D. Goddard. "On the thermoelectricity of W. Thomson: Towards a theory of thermoelastic conductors". In: *Journal of Elasticity* 104.1-2 (2011), pp. 267–280 (cit. on p. 106).

[GÖ97] M. Grmela and H. C. Öttinger. "Dynamics and thermodynamics of complex fluids. I. Development of a general formalism". In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 56.6 (1997), pp. 6620–6632 (cit. on pp. 80, 94–95, 98, 107).

[Gr08] M. Grmela. "Contact geometry of nonequilibrium thermodynamics". In: *Meeting the Entropy Challenge: An International Thermodynamics Symposium in Honor and Memory of Professor Joseph H. Keenan*. Ed. by G. P. Beretta, A. F. Ghoniem, and G. N. Hatsopoulos. Vol. 1033. 2008, pp. 235–240 (cit. on pp. 96, 110).

[Gr13] M. Grmela. "Extensions of nondissipative continuum mechanics toward complex fluids and complex solids". In: *Continuum Mechanics and Thermodynamics* 25.1 (2013), pp. 55–75 (cit. on p. 81).

[Gr86] M. Grmela. "Bracket formulation of diffusion-convection equations". In: *Physica D: Nonlinear Phenomena* 21.2-3 (1986), pp. 179–212 (cit. on p. 105).

[He73] R. Hermann. *Geometry, Physics and Systems*. 2nd ed. New York: Marcel Dekker, 1973 (cit. on pp. 80, 96).

[Hi02] D. Hilbert. "Mathematical Problems". In: *Bulletin of the New York Mathematical Society* 8.10 (1902), pp. 437–479 (cit. on p. 79).

[JÖ04] R. J. J. Jongschaap and H. C. Öttinger. "The mathematical representation of driven thermodynamic systems". In: *Journal of Non-Newtonian Fluid Mechanics* 120.1-3 (2004), pp. 3–9 (cit. on p. 110).

[LY99] E. H. Lieb and J. Yngvason. "The Physics and Mathematics of the Second Law of Thermodynamics". In: *Physics Report* 310.1 (1999), pp. 1–96.

[Mon06] R. Montgomery. *A Tour of Subriemannian Geometries, Their Geodesics and Applications*. Mathematical Surveys and Monographs. American Mathematical Society, 2006 (cit. on p. 101).

[Mor09] P. J. Morrison. “Thoughts on brackets and dissipation: Old and new”. In: *Journal of Physics: Conference Series* 169 (2009) (cit. on p. 80).

[Mr00] R. Mrugała. “On contact and metric structures on thermodynamic spaces”. In: *RIMS Kokyuroku* 1142 (2000), pp. 167–181 (cit. on pp. 79–80).

[MR03] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry. A Basic Exposition of Classical Mechanical Systems.* 2nd ed. (1st ed., 1994). New York: Springer-Verlag, 2003 (cit. on pp. 80, 82).

[MS06] L. M. Martyushev and V. D. Seleznev. “Maximum entropy production principle in physics, chemistry and biology”. In: *Physics Reports* 426.1 (2006), pp. 1–45 (cit. on p. 107).

[Ö05] H. C. Öttinger. *Beyond Equilibrium Thermodynamics.* 1st ed. Hoboken (New Jersey): Wiley, 2005 (cit. on pp. 81, 107–108, 110).

[OR03] A. Odzijewicz and T. S. Ratiu. “Banach Lie-Poisson Spaces and Reduction”. In: *Communications in Mathematical Physics* 243.1 (2003), pp. 1–54 (cit. on pp. 92, 96, 103).

[OZ11] M. Ostoja-Starzewski and A. Zubelewicz. “Powerless fluxes and forces, and change of scale in irreversible thermodynamics”. In: *Journal of Physics A: Mathematical and Theoretical* 44.33 (2011) (cit. on p. 106).

[Q07] H. Quevedo. “Geometrothermodynamics”. In: *Journal of Mathematical Physics* 48.1 (2007) (cit. on p. 80).

[Zi57] H. Ziegler. “Thermodynamik und rheologische Probleme”. In: *Ingenieur-Archiv* 25.1 (1957), pp. 58–70 (cit. on p. 107).

5

Applications

5.1 Metriplectic formulation of Hydrodynamics

The simplest framework for which we may see metriplectic dynamics realise itself is the system of equations of **Subsec.** 2.1.6. This formulation was given by Morrison in [Mo84] and Öttinger in [Ö05] and we report it here in a schematic way, in order to highlight how the various building blocks take place in the geometrical formulation. We do not discuss the derivation of the starting equations, since the latter is on the same streamline of **Subsec.** 2.1.2, where we outlined the necessary steps without directly showing them.

Firstly, the two most natural choices of state are:

- an energy representation:
$$x = (\rho, \boldsymbol{M}, \mathtt{u})\,;$$
- an entropy representation:
$$x' = (\rho, \boldsymbol{M}, \mathtt{s})\,,$$
which we shall adopt below.

ρ is the mass density, $\boldsymbol{M}$ the bulk volume-specific linear momentum of the fluid, $\mathtt{u}$ the volume-specific internal energy, and $\mathtt{s}$ the volume-specific entropy. Another common option consists in choosing the velocity $\boldsymbol{v}$ instead of the linear momentum.

Then, one makes the assumptions of local thermal equilibrium, linear relations between the viscous stress tensor[1] and the bulk velocity

$$\begin{cases} \tau^{ik} = \Lambda^{ikmn} \partial_m \left(\dfrac{M_n}{\rho} \right) \\ \Lambda^{ikmn} \overset{\text{def}}{=} -\eta \left(\delta^{ni}\delta^{mk} + \delta^{nk}\delta^{mi} - \dfrac{2}{3}\delta^{ik}\delta^{mn} \right) - \zeta\delta^{ik}\delta^{mn} \end{cases} , \tag{5.1}$$

and between heat flux and temperature gradient (Fourier Law).

The whole system of equation is the following:

$$\begin{cases} \dfrac{\partial\rho}{\partial t} = -\operatorname{div}(\boldsymbol{M}) \\ \dfrac{\partial\boldsymbol{M}}{\partial t} = -\operatorname{div}\left(\dfrac{\boldsymbol{M}\otimes\boldsymbol{M}}{\rho} + \tau \right) - \operatorname{grad}(p) + \rho\operatorname{grad}(\phi_{\text{grav}}) \\ \dfrac{\partial\mathtt{s}}{\partial t} = -\operatorname{div}\left(\dfrac{\mathtt{s}\boldsymbol{M}}{\rho} \right) + \dfrac{\kappa}{T}\partial^2 T - \dfrac{\boldsymbol{\tau}}{T} : \operatorname{grad}\left(\dfrac{\boldsymbol{M}}{\rho} \right) \end{cases} , \tag{5.2}$$

where ϕ_{grav} is the gravitational energy per unit mass, and ∂^2 is the Laplacian. Temperature and pressure arise for the local equilibrium assumption and are not new variables, since they are functions of mass and entropy density:

$$\begin{cases} T = \dfrac{\partial\mathtt{u}(\rho,\mathtt{s})}{\partial\mathtt{s}} \\ p = \mathtt{s}T(\rho,\mathtt{s}) + \rho\mu(\rho,\mathtt{s}) - \mathtt{u}(\rho,\mathtt{s}) = \mathtt{s}\dfrac{\partial\mathtt{u}(\rho,\mathtt{s})}{\partial\mathtt{s}} + \rho\dfrac{\partial\mathtt{u}(\rho,\mathtt{s})}{\partial\rho} - \mathtt{u}(\rho,\mathtt{s}) \end{cases} . \tag{5.3}$$

The other parameters are $\begin{array}{rl} \eta : & \text{shear viscosity} \\ \zeta : & \text{bulk viscosity} \\ \kappa : & \text{thermal conductivity} \end{array}$.

Having identified the system of interest, we are ready to give its metriplectic formulation. First of all, some mathematical points should be clarified. Indeed, the whole problem of the description of a fluid could be situated in the more general framework of its abstract Hamiltonian description. Here we do not want to give such deeply mathematical (beautiful) insights (we shall give some basic ideas before introducing the discussion about the Boltzmann Equation), but we

---

[1] As usual, the stress tensor is decomposed as: $\sigma^{ik} = \tau^{ik} + p\delta^{ik}$.

limit to notice that, henceforth, thanks to the L^2 inner product

$$\int x \cdot y \mathrm{d}V, \tag{5.4}$$

covectors and vectors (which, on a vector space, are identified with the vector space itself) will be identified. For example, the differential $\mathrm{d}A_y$ of a function A

$$\mathrm{d}A_y(z) = \left\langle \left.\frac{\delta A}{\delta x}\right|_y, z \right\rangle = \int \left.\frac{\delta A}{\delta x}\right|_y \cdot z \mathrm{d}V$$

is identified with the functional derivative $\left.\frac{\delta A}{\delta x}\right|_y$. Therefore, **Eq.** (4.35) takes the form

$$\frac{\partial x}{\partial t} = P_x^\sharp \frac{\delta H}{\delta x} + h_x^\sharp \frac{\delta S}{\delta x}. \tag{5.5}$$

The first ingredient is the generator of reversible dynamics, namely, the total energy

$$H = \int \left[\frac{\boldsymbol{M}^2}{2\rho} + \rho \phi_{\mathrm{grav}} + \mathtt{u}(\rho, \mathtt{s}) \right] \mathrm{d}V, \tag{5.6}$$

which is a constant of the motion if the fluid is an isolated system. Furthermore, we assume that any boundary term arising from integration by parts vanishes. The Poisson bracket is calculated from the corresponding *ideal* fluid model (Euler's equations the ideal fluid) and reads

$$\begin{aligned} \{f, g\}_x = \int \mathrm{d}V \rho \left[\left(\frac{\partial}{\partial \boldsymbol{r}} \frac{\delta f}{\delta \rho} \right) \cdot \frac{\delta g}{\delta \boldsymbol{M}} - \left(\frac{\partial}{\partial \boldsymbol{r}} \frac{\delta g}{\delta \rho} \right) \cdot \frac{\delta f}{\delta \boldsymbol{M}} \right] + \\ + \boldsymbol{M} \cdot \left[\left(\mathrm{grad} \frac{\delta f}{\delta \boldsymbol{M}} \right) \frac{\delta g}{\delta \boldsymbol{M}} - \left(\mathrm{grad} \frac{\delta g}{\delta \boldsymbol{M}} \right) \frac{\delta f}{\delta \boldsymbol{M}} \right] + \\ + \mathtt{s} \left[\left(\frac{\partial}{\partial \boldsymbol{r}} \frac{\delta f}{\delta \mathtt{s}} \right) \cdot \frac{\delta g}{\delta \boldsymbol{M}} - \left(\frac{\partial}{\partial \boldsymbol{r}} \frac{\delta g}{\delta \mathtt{s}} \right) \cdot \frac{\delta f}{\delta \boldsymbol{M}} \right]. \end{aligned} \tag{5.7}$$

Other forms of the bracket may be obtained according to the procedure that is used to the scope, and they all related through integration by parts. This is the more symmetric and the one which allows for the most abstract treatment of the hydrodynamic system, which we do not talk about.

The Poisson bracket has several Casimir functions: some of them are of the form

$$C(x) = \int \rho \psi(\mathtt{s}/\rho) \mathrm{d}V, \tag{5.8}$$

among which we notice the total mass and the total entropy of the fluid

$$m(x) = \int \rho \mathrm{d}V \tag{5.9}$$

$$S(x) = \int \mathsf{s} \mathrm{d}V. \tag{5.10}$$

The latter is chosen to be the generator of the irreversible dynamics, which is governed by the dissipative bracket

$$\begin{aligned}[f,g]_x = \int \mathrm{d}V \Bigg\{ & \kappa T^2 \partial_k \left(\frac{1}{T} \frac{\delta f}{\delta \mathsf{s}} \right) \partial^k \left(\frac{1}{T} \frac{\delta g}{\delta \mathsf{s}} \right) + \\ & + T \Lambda^{ikmn} \left[\partial_i \left(\frac{\delta f}{\delta M^k} \right) - \frac{1}{T} \partial_i \left(\frac{M_k}{\rho} \right) \frac{\delta f}{\delta \mathsf{s}} \right] \left[\partial_m \left(\frac{\delta g}{\delta M^n} \right) - \frac{1}{T} \partial_m \left(\frac{M_n}{\rho} \right) \frac{\delta g}{\delta \mathsf{s}} \right] \Bigg\} . \end{aligned} \tag{5.11}$$

This may verified to be symmetric and positive-definite [Mo84]; the Hamiltonian function and the total mass are distinguished functions.

To the brackets correspond the Poisson operator[2]

$$P^\sharp_x = - \begin{pmatrix} 0 & \partial_k (\rho & 0 \\ \rho \partial_k & M_i \partial_k + \partial_i (M_k & s \partial_k \\ 0 & \partial_k (s & 0 \end{pmatrix} \tag{5.12}$$

and the friction operator

$$h^\sharp_x = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \partial_i \left(T \Lambda^{ikmn} \partial_m \right. & -\partial_i \left[\Lambda^{ikmn} \partial_n \left(\frac{M_n}{\rho} \right) \right. \\ 0 & \Lambda^{ikmn} \partial_i \left(\frac{M_k}{\rho} \right) \partial_m & -\frac{1}{T} \Lambda^{ikmn} \partial_i \left(\frac{M_k}{\rho} \right) \partial_n \left(\frac{M_m}{\rho} \right) - \frac{1}{T} \partial_i \left[\beta T^2 \partial_i \left(\frac{1}{T} \right. \right. \end{pmatrix} . \tag{5.13}$$

The last ingredients we need are the functional derivatives

$$\frac{\delta H}{\delta x} = \begin{pmatrix} \mu + \phi_{\mathrm{grav}} - \dfrac{\boldsymbol{M}^2}{2\rho^2} \\ \boldsymbol{v} \\ T \end{pmatrix} , \qquad \frac{\delta S}{\delta x} = \begin{pmatrix} 0 \\ \boldsymbol{0} \\ 1 \end{pmatrix} . \tag{5.14}$$

Putting all the pieces together, we realise that the full system of equations (5.2)

---

[2] With the open bracket we mean that what must be differentiated is not only what stands immediately on its right, but also what comes after the matrix multiplication has been performed.

takes the metriplectic form (5.5).

Equilibrium states

Equilibrium states can be found by doing entropy maximisation subjected to constraints. For example, a family of equilibrium states may be recovered if we define the potential

$$\Phi(x,\beta) = -S(x) + \beta H(x) + \lambda m(x). \tag{5.15}$$

Its functional derivatives are

$$\begin{cases} \dfrac{\delta\Phi}{\delta\rho} = \beta\left[-\dfrac{M^2}{2\rho^2} + \phi_{\text{grav}} + \mu(\rho,\mathtt{s})\right] + \lambda \\ \dfrac{\delta\Phi}{\delta\boldsymbol{M}} = \beta\dfrac{\boldsymbol{M}}{\rho} \\ \dfrac{\delta\Phi}{\delta\mathtt{s}} = \beta T(\rho,\mathtt{s}) - 1 \end{cases}, \tag{5.16}$$

and setting them to zero, we obtain

$$\begin{cases} \lambda_{\text{eq}} = -\dfrac{\mu(\rho_{\text{eq}},\mathtt{s}_{\text{eq}}) + \phi_{\text{grav}}}{T(\rho_{\text{eq}},\mathtt{s}_{\text{eq}})} \\ \boldsymbol{M}_{\text{eq}} = \mathbf{0} \\ \beta_{\text{eq}} = \dfrac{1}{T(\rho_{\text{eq}},\mathtt{s}_{\text{eq}})} \end{cases}. \tag{5.17}$$

The potential thus becomes

$$\Phi(x_{\text{eq}},\beta_{\text{eq}},\lambda_{\text{eq}}) = -S(x_{\text{eq}}) + \frac{1}{T(\rho_{\text{eq}},\mathtt{s}_{\text{eq}})}H(x_{\text{eq}}) - \frac{\mu(\rho_{\text{eq}},\mathtt{s}_{\text{eq}}) + \phi_{\text{grav}}}{T(\rho_{\text{eq}},\mathtt{s}_{\text{eq}})}m(x_{\text{eq}}), \tag{5.18}$$

i.e., the opposite of a Massieu function at equilibrium, plus a contribution from the external potential. In **Eq.** (5.17) we have already the value of the Lagrange multiplier β, but in the first equation we can only determine λ_{eq} as a function of ρ_{eq} and $\mathtt{s}_{\text{eq}}$. This means that the equilibrium is characterized by three degrees of freedom: one is the external potential energy ϕ_{grav}, and the other two are ρ_{eq} and $\mathtt{s}_{\text{eq}}$ (or, by means of a proper change of variables, T_{eq} or μ_{eq}).

Thermodynamic forces and fluxes

As briefly reported in **Subsec.** 4.4.6, Edwards and Öttinger decompose the friction operator into

$$h^\sharp = B \circ L \circ B^T, \tag{5.19}$$

where L is the phenomenological matrix of classical nonequilibrium thermodynamics. In the case under study, we have

$$B = \begin{pmatrix} 0 & 0 \\ \partial_i & 0 \\ \frac{1}{T}\partial_i\left(\frac{M_k}{\rho}\right) & \frac{1}{T}\partial_i \end{pmatrix} \tag{5.20}$$

$$B^T = \begin{pmatrix} 0 & -\partial_m & \frac{1}{T}\partial_n\left(\frac{M_m}{\rho}\right) \\ 0 & 0 & -\partial_n\left(\frac{1}{T}\right) \end{pmatrix} \tag{5.21}$$

$$L = \begin{pmatrix} -T\Lambda^{ikmn} & 0 \\ 0 & \kappa T^2 \delta^{ni} \end{pmatrix} \tag{5.22}$$

Then, thermodynamic forces are defined:

$$X \overset{\text{def}}{=} -B^T \frac{\delta S}{\delta x} = \begin{pmatrix} -\frac{1}{T}\partial_n\left(\frac{M_m}{\rho}\right) \\ -\partial_n \frac{1}{T} \end{pmatrix}, \tag{5.23}$$

so that the classical form of linear irreversible thermodynamics arises naturally:

$$\dot{S} = X(LX) \tag{5.24}$$

and the (Onsager) symmetry and non-negative definiteness of L are deduced from the analogous properties of $h^\sharp$ (cf. **Eq.** (5.19)). Defining also the thermodynamic fluxes

$$J = LX, \tag{5.25}$$

we realize that the structure of this GENERIC version carries the one of Linear Irreversible Thermodynamics. The transformation B^T basically moves some differential operator from M to B, thus to X, allowing us to recognize the forces of Classical Non-Equilibrium Thermodynamics.

5.2 Metriplectic formulation of visco-resistive magnetohydrodynamics

Another example of realisation of metriplectic (GENERIC) dynamics is the formulation of visco-resistive MHD in [MT12]. In this article, Materassi and Tassi consider a plasma interacting with the magnetic field generated by its own motion, and subjected to dissipation due to finite viscosity, electrical resistivity and thermal conductivity. The goal of the paper is to show that the resulting system of differential equations can effectively be written in metriplectic form. Here we borrow their result in order to give a concrete example of realization of **Eq.** (4.35) without entering the details of their derivation.

The state is chosen as

$$x = (\rho, \boldsymbol{v}, \boldsymbol{B}, s)\,, \tag{5.26}$$

where ρ is the mass density, $\boldsymbol{v}$ the bulk velocity of the plasma, $\boldsymbol{B}$ the magnetic field, and s the mass-specific entropy. This is slightly different choice with respect to the previous formulation of Hydrodynamics. The passage between the two can be made by means of the change of variables

$$\begin{cases} \rho(\rho, \boldsymbol{v}, s) = \rho \\ \boldsymbol{M}(\rho, \boldsymbol{v}, s) = \rho \boldsymbol{v} \\ \mathsf{s}(\rho, \boldsymbol{v}, s) = \rho s \end{cases} , \tag{5.27}$$

which implies the laws for the functional derivatives

$$\begin{cases} \dfrac{\delta F(\rho, \boldsymbol{v}, s)}{\delta \rho} = \left.\dfrac{\delta F(\rho, \boldsymbol{M}, \mathsf{s})}{\delta \rho}\right|_{(\rho, \rho\boldsymbol{v}, \rho s)} + \boldsymbol{v} \cdot \left.\dfrac{\delta F(\rho, \boldsymbol{M}, \mathsf{s})}{\delta \boldsymbol{M}}\right|_{(\rho, \rho\boldsymbol{v}, \rho s)} + \\ \qquad\qquad + s \left.\dfrac{\delta F(\rho, \boldsymbol{M}, \mathsf{s})}{\delta \mathsf{s}}\right|_{(\rho, \rho\boldsymbol{v}, \rho s)} \\ \dfrac{\delta F(\rho, \boldsymbol{v}, s)}{\delta \boldsymbol{v}} = \rho \left.\dfrac{\delta F(\rho, \boldsymbol{M}, \mathsf{s})}{\delta \boldsymbol{M}}\right|_{(\rho, \rho\boldsymbol{v}, \rho s)} \\ \dfrac{\delta F(\rho, \boldsymbol{v}, s)}{\delta s} = \rho \left.\dfrac{\delta F(\rho, \boldsymbol{M}, \mathsf{s})}{\delta \mathsf{s}}\right|_{(\rho, \rho\boldsymbol{v}, \rho s)} \end{cases} . \tag{5.28}$$

It is assumed a local thermal equilibrium, and linear relations between the viscous stress tensor and the bulk velocity:

$$\begin{cases} \tau^{ik} = \Lambda^{ikmn}\partial_m v_n \\ \Lambda^{ikmn} \overset{\text{def}}{=} -\eta\left(\delta^{ni}\delta^{mk} + \delta^{nk}\delta^{mi} - \dfrac{2}{3}\delta^{ik}\delta^{mn}\right) - \zeta\delta^{ik}\delta^{mn} \end{cases} . \tag{5.29}$$

and between heat flux and temperature gradient (Fourier Law).

The whole system of equation is the following (the magnetic permeability of vacuum has been put to zero in order to keep the notation neater):

$$\begin{cases} \partial_t\rho = -\partial_k(\rho v^k) \\ \partial_t v^i = -v^k\partial_k v^i - \dfrac{1}{\rho}\partial^i p - \dfrac{1}{2\rho}\partial^i B^2 + \dfrac{1}{\rho}B^k\partial_k B^i - \partial^i\phi_{\text{grav}} - \dfrac{1}{\rho}\partial_k\tau^{ik} \\ \partial_t B^i = B^j\partial_j v^i - B^i\partial_j v^j - v^j\partial_j B^i + \mu\partial^2 B^i \\ \partial_t s = -v^k\partial_k s + \dfrac{\tau^{ik}}{\rho T}\partial_k v_i - \dfrac{\mu}{\rho T}\varepsilon^{ikh}\varepsilon_h{}^{mn}\partial_i B_k\partial_m B_n + \dfrac{\kappa}{\rho T}\partial^2 T \end{cases} , \tag{5.30}$$

where ϕ_{grav} is the gravitational energy per unit mass. Pressure and temperature arise for the local equilibrium assumption and are not new variables, since they are functions of mass and entropy density:

$$\begin{cases} p = \rho^2\dfrac{\partial u(\rho,s)}{\partial\rho} \\ T = \dfrac{\partial u(\rho,s)}{\partial s} \end{cases} . \tag{5.31}$$

The other parameters are

η : shear viscosity
ζ : bulk viscosity
μ : plasma resistivity
κ : thermal conductivity
.

The generator of reversible dynamics is the total energy

$$H = \int\left[\rho\frac{v^2}{2} + \rho\phi_{\text{grav}} + \frac{B^2}{2} + \rho u(\rho,s)\right]\mathrm{d}V \tag{5.32}$$

and is a constant of the motion if the plasma is an isolated system. The Poisson bracket is calculated from the corresponding *ideal* MHD model and reads

$$\begin{aligned}\{f,g\}_x = -\int \mathrm{d}V \Bigg[&\frac{\delta f}{\delta \rho}\partial^i\left(\frac{\delta g}{\delta v^i}\right)+\frac{\delta g}{\delta \rho}\partial^i\left(\frac{\delta f}{\delta v^i}\right)-\frac{1}{\rho}\frac{\delta f}{\delta v^i}\varepsilon^{ikj}\varepsilon_{jmn}\frac{\delta g}{\delta v^k}\partial^m v^n+\\ &+\frac{1}{\rho}\frac{\delta f}{\delta v^i}\varepsilon^{ijk}\varepsilon_{kmn}B_j\partial^m\left(\frac{\delta g}{\delta B_n}\right)+\\ &+\frac{\delta f}{\delta B^i}\varepsilon^{ijk}\partial_j\left(\frac{1}{\rho}\varepsilon_{kmn}B^m\frac{\delta g}{\delta v_n}\right)+\frac{1}{\rho}\partial^i s\left(\frac{\delta f}{\delta s}\frac{\delta g}{\delta v^i}-\frac{\delta g}{\delta s}\frac{\delta f}{\delta v^i}\right)\Bigg].\end{aligned} \tag{5.33}$$

This has several Casimir functions: some of them are of the form

$$C(x)=\int \rho\psi(s)\mathrm{d}V, \tag{5.34}$$

among which we notice the total mass and the total entropy of the plasma

$$m(x)=\int \rho \mathrm{d}V \tag{5.35}$$

$$S(x)=\int \rho s \mathrm{d}V. \tag{5.36}$$

The latter is chosen to be the generator of the irreversible dynamics, which is governed by the dissipative bracket

$$\begin{aligned}&[f,g]_x =\\ &=\int \mathrm{d}VT\Bigg\{-\Lambda^{ikmn}\left[\partial_i\left(\frac{1}{\rho}\frac{\delta f}{\delta v^k}\right)-\frac{1}{\rho T}\partial_i v_k\frac{\delta f}{\delta s}\right]\left[\partial_m\left(\frac{1}{\rho}\frac{\delta g}{\delta v^n}\right)-\frac{1}{\rho T}\partial_m v_n\frac{\delta g}{\delta s}\right]+\\ &\qquad+\mu\varepsilon^{ikj}\varepsilon_j^{\ mn}\left[\partial_i\left(\frac{\delta f}{\delta B^k}\right)-\frac{1}{\rho T}\partial_i B_k\frac{\delta f}{\delta s}\right]\left[\partial_m\frac{\delta g}{\delta B^n}-\frac{1}{\rho T}\partial_m B_n\frac{\delta g}{\delta s}\right]+\\ &\qquad+\kappa T\partial_k\left(\frac{1}{\rho T}\frac{\delta f}{\delta s}\right)\partial^k\left(\frac{1}{\rho T}\frac{\delta g}{\delta s}\right)\Bigg\}.\end{aligned} \tag{5.37}$$

This may verified to be symmetric, positive-definite [MT12]; the Hamiltonian function and the total mass are distinguished functions.

The equilibrium states are the same as the previous section, with the additional condition $\boldsymbol{B}=\mathbf{0}$.

In this case too, the Poisson and friction operators could be worked out, as well as thermodynamic forces and fluxes may be given an expression of the type $-B^T\dfrac{\delta S}{\delta x}$.

5.3 Metriplectic and SEA formulations of the Boltzmann Equation

In this section, we shall illustrate how the two models realize themselves in Kinetic Theory: in particular, we shall see the form that the various quantities defined above assume to reproduce the Boltzmann Equation

$$\begin{aligned}\frac{\partial f(\boldsymbol{r},\boldsymbol{p};t)}{\partial t} = \left[\frac{\partial\phi(\boldsymbol{r})}{\partial\boldsymbol{r}}\cdot\frac{\partial}{\partial\boldsymbol{p}} - \frac{\boldsymbol{p}}{m}\cdot\frac{\partial}{\partial\boldsymbol{r}}\right] f(\boldsymbol{r},\boldsymbol{p};t)+ \\ + \int \mathrm{d}^3\boldsymbol{p}_1\mathrm{d}^3\boldsymbol{p}_2\mathrm{d}^3\boldsymbol{q}_1\mathrm{d}^3\boldsymbol{q}_2\delta^{(3)}(\boldsymbol{p}_1-\boldsymbol{p})w(\boldsymbol{q}_1,\boldsymbol{q}_2|\boldsymbol{p}_1,\boldsymbol{p}_2) \\ \left[f(\boldsymbol{r},\boldsymbol{q}_1;t)f(\boldsymbol{r},\boldsymbol{q}_2;t) - f(\boldsymbol{r},\boldsymbol{p}_1;t)f(\boldsymbol{r},\boldsymbol{p}_2;t)\right].\end{aligned} \tag{5.38}$$

$\phi(\boldsymbol{r})$ is an external potential, $w(\boldsymbol{q}_1,\boldsymbol{q}_2|\boldsymbol{p}_1,\boldsymbol{p}_2)$ the transition probability given by

$$\begin{aligned}&w(\boldsymbol{q}_1,\boldsymbol{q}_2|\boldsymbol{p}_1,\boldsymbol{p}_2) = \\ &\quad = \delta^{(3)}(\boldsymbol{q}_1+\boldsymbol{q}_2-\boldsymbol{p}_1-\boldsymbol{p}_2)\delta^{(3)}(\boldsymbol{q}_1^2+\boldsymbol{q}_2^2-\boldsymbol{p}_1^2-\boldsymbol{p}_2^2)\frac{8}{m}\sigma(\boldsymbol{q}_1,\boldsymbol{q}_2|\boldsymbol{p}_1,\boldsymbol{p}_2),\end{aligned} \tag{5.39}$$

and σ the differential cross section calculated in the centre-of-mass frame. This is the formulation given by Öttinger in [Ö97].

First of all, comes the choice of the state. According to the formulation by Marsden and Weinstein in [MW82] and recalling the arguments in **Par.** 4.1.2, we see the space of distribution functions as a Banach space in duality with the Lie algebra of the smooth functions (modulo constant); this space has a natural Poisson structure. Let's make it more precise.

The space of smooth functions (modulo constant) is given the structure of a Poisson algebra $(C^\infty(M), \{,\}_1)$, which – in particular – is a Lie algebra. Next, we choose a space in (strong) duality with the latter. For reasons that will become clear later, we choose the space of strictly positive Schwartz functions $\mathcal{S}_+(M)$ as the dual to $C^\infty(M)$ through the (strongly non-degenerate) pairing $\langle,\rangle\colon \mathcal{S}_+(M)\times C^\infty(M)\to\mathbb{R}$ given by the $L^2(M)$-product:

$$\langle f, a\rangle = \int_M fa\Lambda, \tag{5.40}$$

where $f\in\mathcal{S}_+(M)$, $a\in C^\infty(M)$ and $\Lambda = \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p}$ is the Liouville measure. The strengthen of the pairing allows to identify covectors in $T\mathcal{S}_+(M)\cong\mathcal{S}_+(M)$ with smooth functions in $C^\infty(M)$.

In **Par.** 4.1.2 we explained that a space in duality with a Lie algebra is naturally a Poisson manifold. Hence, the space of smooth functions on it is a Poisson algebra, whose Poisson brackets $\{,\}_2$ are often called *Lie-Poisson brackets* and take the expression

$$\{A, B\}_2(g) = \left\langle g, \left\{ \frac{\delta A}{\delta f}, \frac{\delta B}{\delta f} \right\}_1 \right\rangle,$$

where $A, B \in C^\infty(\mathcal{S}_+(M))$, $g \in \mathcal{S}_+(M)$ and the functional derivatives are defined by the usual implicit relation

$$\mathrm{d}A_f(g) = \left\langle g, \frac{\delta A}{\delta f} \right\rangle = \int_M g \frac{\delta A}{\delta f} \Lambda \qquad \left(f, g \in \mathcal{S}_+(M),\ \frac{\delta A}{\delta f} \in C^\infty(M) \right).$$

The question that still must be answered is: why did we choose the Schwartz space? The smooth functions (A) we shall consider in $C^\infty(\mathcal{S}_+(M))$ are the (mean) values of the physical observables a

$$A(f) = \langle f, a \rangle = \int_M f \mathsf{a} \Lambda, \qquad \gamma \in \mathcal{S}_+(M),\ \mathsf{a} \in C^\infty(M). \tag{5.41}$$

The physical observables a may be polynomial functions in M: examples are the linear momentum and the energy. Thus, in order for the integrals in **Eq.** (5.41) to be finite, the best choice is the space of (strictly positive) rapidly decreasing functions in M.[3]

The abstract dynamics is the motion of the state on the (infinite-dimensional) Hilbert space $\mathcal{S}_+(M)$, and is represented by a curve $\alpha : I \to \mathcal{S}_+(M)\, (I \subseteq \mathbb{R})$, solution of the following differential equation:

$$\dot{\alpha}(t) = X^H_{\alpha(t)} + Y^S_{\alpha(t)}, \tag{5.42}$$

where

$$\alpha(t) = f(\boldsymbol{r}, \boldsymbol{p}; t). \tag{5.43}$$

Let's now build the fundamental quantities of the theory, which will give rise to the dynamical equation in both SEA and GENERIC formulations. The entropy

---

[3] Indeed, a rapid decrease in momentum space would be enough, since observables are polynomial functions of momentum, so that $\mathcal{S}_+(M)$ may also be thought of as the suitable space of functions that eliminate the divergences in the correct way.

functional is

$$S(f(\boldsymbol{r},\boldsymbol{p})) = -k_B \int f(\boldsymbol{r},\boldsymbol{p}) \ln \frac{f(\boldsymbol{r},\boldsymbol{p})}{b} \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p} \overset{\text{def}}{=} \int \mathtt{s}(\boldsymbol{r},\boldsymbol{p}) f(\boldsymbol{r},\boldsymbol{p}) \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p}, \quad (5.44)$$

where b is a suitable constant. The conserved properties are the five collisional invariants (and linear combinations)

$$C^0(f(\boldsymbol{r},\boldsymbol{p})) = \int f(\boldsymbol{r},\boldsymbol{p}) \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p} = N, \quad (5.45)$$

$$C^1(f(\boldsymbol{r},\boldsymbol{p})) = \int p_x f(\boldsymbol{r},\boldsymbol{p}) \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p} = P_x, \quad (5.46)$$

$$C^2(f(\boldsymbol{r},\boldsymbol{p})) = \int p_y f(\boldsymbol{r},\boldsymbol{p}) \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p} = P_y, \quad (5.47)$$

$$C^3(f(\boldsymbol{r},\boldsymbol{p})) = \int p_z f(\boldsymbol{r},\boldsymbol{p}) \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p} = P_z, \quad (5.48)$$

$$C^4(f(\boldsymbol{r},\boldsymbol{p})) = \int \left[\frac{\boldsymbol{p}^2}{2m} + \phi(\boldsymbol{r})\right] f(\boldsymbol{r},\boldsymbol{p}) \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p} = H, \quad (5.49)$$

among which we recognize the total energy. In short,

$$C^j(f(\boldsymbol{r},\boldsymbol{p})) = \int \varphi^j(\boldsymbol{r},\boldsymbol{p}) f(\boldsymbol{r},\boldsymbol{p}) \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p}, \quad (5.50)$$

where, of course,

$$\begin{aligned}
\varphi^0(\boldsymbol{r},\boldsymbol{p}) &= 1,\\
\varphi^1(\boldsymbol{r},\boldsymbol{p}) &= p_x,\\
\varphi^2(\boldsymbol{r},\boldsymbol{p}) &= p_y,\\
\varphi^3(\boldsymbol{r},\boldsymbol{p}) &= p_z,\\
\varphi^4(\boldsymbol{r},\boldsymbol{p}) &= \frac{\boldsymbol{p}^2}{2m} + \phi(\boldsymbol{r})
\end{aligned}$$

are smooth functions. Finally, the expressions for the functional derivatives are

$$\left.\frac{\delta S}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} = -k_B \left[\ln \frac{f(\boldsymbol{r},\boldsymbol{p})}{b} + 1\right] \quad (5.51)$$

$$\left.\frac{\delta C^j}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} = \varphi^j(\boldsymbol{r},\boldsymbol{p}). \quad (5.52)$$

If the distribution were allowed to vanish, the expression (5.51) would present a divergence for values of $\boldsymbol{r}$ and $\boldsymbol{p}$ outside the support of f: it would not be a function of $C^\infty(M)$.

5.3.1 SEA

At variance with the last formulation in [BH13], here we consider an isolated system: we do not have a state for each point of space, and the quantity being maximized is the total entropy production.

In the SEA model, **Eq.** (5.42) takes the form:

$$\dot{\alpha}(t) = X^H_{\alpha(t)} + g^\sharp_{\alpha(t)} \left(\left.\frac{\delta S}{\delta x}\right|_{\alpha(t)} - \sum\nolimits_j \beta^j \left.\frac{\delta C^j}{\delta x}\right|_{\alpha(t)} \right) . \tag{5.53}$$

Thus, **Eq.** (5.53) becomes

$$\frac{\partial f(\boldsymbol{r},\boldsymbol{p};t)}{\partial t} = X^H_{f(\boldsymbol{r},\boldsymbol{p};t)} + g^\sharp_{f(\boldsymbol{r},\boldsymbol{p};t)} \left(\left.\frac{\delta S}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p};t)} - \sum\nolimits_j \beta^j \left.\frac{\delta C^j}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p};t)} \right) , \tag{5.54}$$

where

$$X^H_{f(\boldsymbol{r},\boldsymbol{p};t)} = \left[\frac{\partial \phi(\boldsymbol{r})}{\partial \boldsymbol{r}} \cdot \frac{\partial}{\partial \boldsymbol{p}} - \frac{\boldsymbol{p}}{m} \cdot \frac{\partial}{\partial \boldsymbol{r}} \right] f(\boldsymbol{r},\boldsymbol{p};t) \tag{5.55}$$

is prescribed.

The values of the Lagrange multipliers are found by solving the following system of five algebraic equations:

$$\sum_{j=0}^{4} g(\operatorname{grad} C^i, \operatorname{grad} C^j)\beta^j = g(\operatorname{grad} C^i, \operatorname{grad} S) \qquad (i \in [0,4]) . \tag{5.56}$$

In contrast to the GENERIC, which imposes restricting conditions on the friction operator $D^\sharp$, the SEA method is that it does not assume *a priori* any form for the metric tensor g. So, without a particular choice of g, the gradients remain unspecified. The subsequent effort in the SEA philosophy, is to find a metric tensor that models correctly and efficiently the collision integral of the Boltzmann Equation in the best way, so as to extend the validity of the traditional Kinetic Models, such as BGK, ES-BGK, etc. from the near-equilibrium to the far non-equilibrium domain. The problem of identifying criteria to that choice is still open. Recent numerical results [BH13] (which we shall expose in **Sec.** 5.4) show that the choice of a uniform (Fisher-Rao) metric yields poor models in this framework; more precisely, although near equilibrium it is fully equivalent to the BGK model, in the far non-equilibrium regime it selects trajectories in state space that diverge from the direction of evolution actually chosen by the full Boltzmann collision integral. It is hoped that the present analysis and perhaps Information Geometry could provides hints to find a suitable metric for this purpose.

5.3.2 GENERIC

The results of this subsection are borrowed from Öttinger's article [Ö97], where they are also justified. Here we only present them without proofs.

GENERIC for our infinite-dimensional system takes the expression:

$$\dot{\alpha}(t) = P^{\sharp}_{\alpha(t)}\left(\left.\frac{\delta H}{\delta x}\right|_{\alpha(t)}\right) + h^{\sharp}_{\alpha(t)}\left(\left.\frac{\delta S}{\delta x}\right|_{\alpha(t)}\right), \tag{5.57}$$

or, more explicitly for the case of Kinetic Theory,

$$\frac{\partial f(\boldsymbol{r},\boldsymbol{p};t)}{\partial t} = P^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p};t)}\left(\left.\frac{\delta H}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p};t)}\right) + h^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p};t)}\left(\left.\frac{\delta S}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p};t)}\right), \tag{5.58}$$

The Poisson operator at point $f(\boldsymbol{r},\boldsymbol{p})$ is

$$P^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p})} = \frac{\partial}{\partial \boldsymbol{p}} \cdot f(\boldsymbol{r},\boldsymbol{p})\frac{\partial}{\partial \boldsymbol{r}} - \frac{\partial}{\partial \boldsymbol{r}} \cdot f(\boldsymbol{r},\boldsymbol{p})\frac{\partial}{\partial \boldsymbol{p}}, \tag{5.59}$$

and the associated Poisson bracket at point $f(\boldsymbol{r},\boldsymbol{p})$

$$\begin{aligned}
&\{A,B\}_{f(\boldsymbol{r},\boldsymbol{p})} = P_{f(\boldsymbol{r},\boldsymbol{p})}\left(\left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})}, \left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})}\right) = \\
&= \left\langle \left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})}, P^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p})}\left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})}\right\rangle = \\
&= \int \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3p \left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} P^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p})}\left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} = \\
&= \int \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p} \left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} \left[\frac{\partial}{\partial \boldsymbol{p}} \cdot f(\boldsymbol{r},\boldsymbol{p})\frac{\partial}{\partial \boldsymbol{r}} - \frac{\partial}{\partial \boldsymbol{r}} \cdot f(\boldsymbol{r},\boldsymbol{p})\frac{\partial}{\partial \boldsymbol{p}}\right] \left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} = \\
&= \int \mathrm{d}^3\boldsymbol{r}\mathrm{d}^3\boldsymbol{p} f(\boldsymbol{r},\boldsymbol{p}) \left[\frac{\partial}{\partial \boldsymbol{p}}\left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} \cdot \frac{\partial}{\partial \boldsymbol{r}}\left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} - \frac{\partial}{\partial \boldsymbol{p}}\left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} \cdot \frac{\partial}{\partial \boldsymbol{r}}\left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})}\right]
\end{aligned} \tag{5.60}$$

In the case of the friction operator, we don't know how to arrive at an expression similar to **Eq.** (5.59), which was obtained by some integrations by parts.[4] Indeed, the Boltzmann Equation itself suggests that we are not dealing with a differential operator, but with an integral one. Therefore, we present only its kernel $\hat{M}(\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{p}_1)$, defined by

$$h^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p})} \left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} = \int \mathrm{d}^3\boldsymbol{p}_1 \hat{M}_{f(\boldsymbol{r},\boldsymbol{p})}(\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{p}_1) \left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p}_1)}, \tag{5.61}$$

where

$$\begin{aligned}
&\hat{M}_{f(\boldsymbol{r},\boldsymbol{p})}(\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{p}_1) = \\
&\quad = \frac{1}{k_B} \int \mathrm{d}^3\boldsymbol{q}_1 \mathrm{d}^3\boldsymbol{q}_2 \mathrm{d}^3\boldsymbol{p}_2 w(\boldsymbol{q}_1, \boldsymbol{q}_2 | \boldsymbol{p}, \boldsymbol{p}_2) \\
&\qquad \left[\delta^{(3)}(\boldsymbol{p} - \boldsymbol{p}_1) + \delta^{(3)}(\boldsymbol{p}_2 - \boldsymbol{p}_1) - \delta^{(3)}(\boldsymbol{q}_1 - \boldsymbol{p}_1) - \delta^{(3)}(\boldsymbol{q}_2 - \boldsymbol{p}_1)\right] \\
&\qquad \frac{f(\boldsymbol{r}, \boldsymbol{q}_1) f(\boldsymbol{r}, \boldsymbol{q}_2) - f(\boldsymbol{r}, \boldsymbol{p}) f(\boldsymbol{r}, \boldsymbol{p}_2)}{\ln[f(\boldsymbol{r}, \boldsymbol{q}_1) f(\boldsymbol{r}, \boldsymbol{q}_2)] - \ln[f(\boldsymbol{r}, \boldsymbol{p}) f(\boldsymbol{r}, \boldsymbol{p}_2)]}.
\end{aligned} \tag{5.62}$$

is the dissipative 'matrix' given in **Eq.** (12) of Ref. [Ö97]. The associated dissipative bracket at point $f(\boldsymbol{r}, \boldsymbol{p})$ reads

$$\begin{aligned}
[A, B]_{f(\boldsymbol{r},\boldsymbol{p})} &= h_{f(\boldsymbol{r},\boldsymbol{p})} \left(\left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})}, \left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} \right) = \\
&= \left\langle \left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})}, h^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p})} \left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} \right\rangle = \\
&= \int \mathrm{d}^3\boldsymbol{r} \mathrm{d}^3\boldsymbol{p} \left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} h^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p})} \left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} = \\
&= \int \mathrm{d}^3\boldsymbol{r} \mathrm{d}^3\boldsymbol{p} \mathrm{d}^3\boldsymbol{p}_1 \left.\frac{\delta A}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} \hat{M}_{f(\boldsymbol{r},\boldsymbol{p})}(\boldsymbol{r}, \boldsymbol{p}, \boldsymbol{p}_1) \left.\frac{\delta B}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p}_1)} \qquad (5.63)
\end{aligned} \tag{5.64}$$

It is easy but important to verify that the degeneracy requirements

$$h^{\sharp}_{f(\boldsymbol{r},\boldsymbol{p})} \left.\frac{\delta C^j}{\delta f}\right|_{f(\boldsymbol{r},\boldsymbol{p})} = 0 \qquad \forall j \tag{5.65}$$

are consequence of the symmetry property (invariance upon exchange of q_1, q_2

---

[4] The kernel of the Poisson operator, if we derive it directly from the Boltzmann Equation, contains some Dirac delta functions, so that – after some integrations by parts – the integral operator can be turned into a differential operator.

with p_1, p_2) of both the transition probabilities $w(q_1, q_2|p_1, p_2)$ and the positive semi-definite resistance matrix

$$\Gamma(\boldsymbol{q}_1, \boldsymbol{q}_2|\boldsymbol{p}_1, \boldsymbol{p}_2) = \frac{\ln[f(\boldsymbol{r}, \boldsymbol{q}_1)f(\boldsymbol{r}, \boldsymbol{q}_2)] - \ln[f(\boldsymbol{r}, \boldsymbol{p}_1)f(\boldsymbol{r}, \boldsymbol{p}_2)]}{f(\boldsymbol{r}, \boldsymbol{q}_1)f(\boldsymbol{r}, \boldsymbol{q}_2) - f(\boldsymbol{r}, \boldsymbol{p})f(\boldsymbol{r}, \boldsymbol{p}_1)} \tag{5.66}$$

whose form was suggested by the related work in [Si87] on chemical kinetics and in the present kinetic theory framework can be interpreted as a resistance matrix due to the collisions from $\boldsymbol{q}_1, \boldsymbol{q}_2$ to $\boldsymbol{p}_1, \boldsymbol{p}_2$ and vice versa. Indeed, the entropy production rate can be written as

$$\begin{split}\Sigma = k_B \int \mathrm{d}^3\boldsymbol{q}_1 \mathrm{d}^3\boldsymbol{q}_2 \mathrm{d}^3\boldsymbol{p}_1 \mathrm{d}^3\boldsymbol{p}_2 w(\boldsymbol{q}_1, \boldsymbol{q}_2|\boldsymbol{p}_1, \boldsymbol{p}_2)\Gamma(\boldsymbol{q}_1, \boldsymbol{q}_2|\boldsymbol{p}_1, \boldsymbol{p}_2) \\ [f(\boldsymbol{r}, \boldsymbol{q}_1)f(\boldsymbol{r}, \boldsymbol{q}_2) - f(\boldsymbol{r}, \boldsymbol{p}_1)f(\boldsymbol{r}, \boldsymbol{p}_2)]^2 .\end{split} \tag{5.67}$$

In the case of GENERIC, the effort has been to put the Boltzmann Equation in metriplectic form, so that the Poisson operator and the friction operator $\hat{M}$ have arisen from this procedure. The friction operator $\hat{M}_f$ given above leads exactly to the collision integral of the Boltzmann Equation. In spite of the complexity of such operators, it is hoped that knowing their explicit forms may help identify kinetic models of the Boltzmann collision integral in the same spirit of the BGK model but capable of capturing more features of the collision dynamics and of giving better approximations in the far non-equilibrium domain.

5.4 Numerical results for relaxation from non-equilibrium states

As it has been stated, the Steepest Entropy Ascent Modelling of Kinetic Theory, oppositely to the GENERIC modelling, aims at developing new model equations for the resolution of the integro-differential Boltzmann equation. The key aspect is thus the choice of the metric. As the final goal is to *create* a kinetic model, the ideal choice would be a metric as simple as possible, being, at the same time, as accurate as possible. The aim of the following paragraphs will be to illustrate the numerical realization of the SEA kinetic model proposed by Beretta and Hadjiconstantinou, that has been exposed in the recent work to which we refer by the same authors [BH13]. Beretta and Hadjiconstantinou have examined in the work two simple metrics. The organization of the present section will be as follows:

- first of all, the Steepest Entropy Ascent kinetic model will be illustrated and it will be shown that it satisfies the two fundamentals properties asked for a collision model, stated in **Eq.** 2.64 and in **Eq.** 2.65;

- secondly, it will be shown that the SEA models that have been proposed coincide with standard BGK models near equilibrium;

- finally, the *numerical setup* will be illustrated and the results for the relaxation from a highly non-equilibrium state will be shown.

5.4.1 Steepest Entropy Ascent Collision Term

In [Be13] and [BH13], the theory is slightly different in that it does not consider isolated systems, as we did in this thesis, but it picks up a state for each point of space. Without referring all the details of the differences between the formulation in those articles and ours in **Subsec.** 5.3.1, we report the result, which – indeed – is very similar to **Eq.** (5.54):

$$\frac{\partial f(\boldsymbol{r};\boldsymbol{v};t)}{\partial t} = \left[\frac{1}{m}\frac{\partial \phi(\boldsymbol{r})}{\partial \boldsymbol{r}}\cdot\frac{\partial}{\partial \boldsymbol{v}} - \boldsymbol{v}\cdot\frac{\partial}{\partial \boldsymbol{r}}\right] f(\boldsymbol{r};\boldsymbol{v};t) - g^{\sharp}_{f(\boldsymbol{r};\boldsymbol{v};t)}\Phi_{f(\boldsymbol{r};\boldsymbol{v};t)}, \tag{5.68a}$$

$$\begin{aligned}\sum_{j=0}^{4} g\left(g^{\sharp}_{f(\boldsymbol{r};\boldsymbol{v};t)}\varphi^{i}(\boldsymbol{r};\boldsymbol{v}), g^{\sharp}_{f(\boldsymbol{r};\boldsymbol{v};t)}\varphi^{j}(\boldsymbol{r};\boldsymbol{v})\right)\beta^{j}_{f(\boldsymbol{r};\boldsymbol{v};t)} &= \\ = g\left(g^{\sharp}_{f(\boldsymbol{r};\boldsymbol{v};t)}\varphi^{i}(\boldsymbol{r};\boldsymbol{v}), g^{\sharp}_{f(\boldsymbol{r};\boldsymbol{v};t)}\mathsf{S}_{f(\boldsymbol{r};\boldsymbol{v};t)}\right)&,\end{aligned} \tag{5.68b}$$

where, in this case,

$$\Phi_{f(\boldsymbol{r};\boldsymbol{v})} \overset{\text{def}}{=} -\mathsf{S}_{f(\boldsymbol{r};\boldsymbol{v})} + \sum_{j=0}^{4}\beta^{j}_{f(\boldsymbol{r};\boldsymbol{v})}\varphi^{j}(\boldsymbol{r};\boldsymbol{v}) \tag{5.69}$$

The non-degenerate pairing between smooth functions and distribution functions is the L^2-product

$$a(\boldsymbol{r}) = \langle f, \mathsf{a}\rangle = \frac{m}{\rho(\boldsymbol{r})}\int \mathrm{d}^3\boldsymbol{v}\, f(\boldsymbol{r};\boldsymbol{v})\mathsf{a}(\boldsymbol{r};\boldsymbol{v}),$$

which gives the local field $a(\boldsymbol{r})$. Since, in its formulation, Beretta takes as state the square root of the distribution function, in analogy with the finite-dimensional case, the Fisher metric is the uniform one. By expressing this fact in terms of f

in lieu of its square root, the SEA collision term reads

$$J_{\text{SEA,Fisher}}(f) = Y^S_{f(\boldsymbol{r};\boldsymbol{v})} = -\frac{1}{k_B\mathcal{T}} f(\boldsymbol{r};\boldsymbol{v})\Phi_{f(\boldsymbol{r};\boldsymbol{v})}. \tag{5.70}$$

In order to make the model well-behaved near equilibrium, as a slightly more complex expression, a positive function $w = w(\boldsymbol{v}, f)$ is added:

$$J_{\text{SEA,w}}(f) = Y^S_{f(\boldsymbol{r};\boldsymbol{v})} = -\frac{1}{k_B\mathcal{T}} f(\boldsymbol{r};\boldsymbol{v})w(\boldsymbol{v}, f(\boldsymbol{r};\boldsymbol{v}))\Phi_{f(\boldsymbol{r};\boldsymbol{v})} \tag{5.71}$$

and, in the successive numerical simulations, is assumed to take on two specific values.

It may be verified that the SEA kinetic model satisfies the two fundamental properties listed before. Indeed, for each collision invariant φ^i, we have that:

$$\begin{aligned}\frac{m}{\rho}\int \mathrm{d}^3\boldsymbol{v}\varphi^i(\boldsymbol{r};\boldsymbol{v})J_{\text{SEA,w}}(f) &= -\frac{m}{\rho}\int \mathrm{d}^3\boldsymbol{v}\varphi^i(\boldsymbol{r};\boldsymbol{v})g^\sharp_{f(\boldsymbol{r};\boldsymbol{v})}\Phi_{f(\boldsymbol{r};\boldsymbol{v})} = \\ = -\left\langle g^\sharp_{f(\boldsymbol{r};\boldsymbol{v})}\Phi_{f(\boldsymbol{r};\boldsymbol{v})}, \varphi^i(\boldsymbol{r};\boldsymbol{v})\right\rangle &= -g\left(g^\sharp_{f(\boldsymbol{r};\boldsymbol{v})}\Phi_{f(\boldsymbol{r};\boldsymbol{v})}, g^\sharp_{f(\boldsymbol{r};\boldsymbol{v})}\varphi^i(\boldsymbol{r};\boldsymbol{v})\right) = 0,\end{aligned} \tag{5.72}$$

where the last equality comes from the fact that the β^i satisfy the system of equations in **Eq.** 5.68b.

The kinetic model satisfies also the H-theorem:

$$\begin{aligned}\frac{m}{\rho}\int_{\Omega_c} \mathrm{d}^3\boldsymbol{v}\mathrm{S}_{f(\boldsymbol{r};\boldsymbol{v})}J_{\text{SEA,w}}(f) &= -\frac{m}{\rho}\int \mathrm{d}^3\boldsymbol{v}\mathrm{S}_{f(\boldsymbol{r};\boldsymbol{v})}g^\sharp_{f(\boldsymbol{r};\boldsymbol{v})}\Phi_{f(\boldsymbol{r};\boldsymbol{v})} = \\ = -\left\langle g^\sharp_{f(\boldsymbol{r};\boldsymbol{v})}\Phi_{f(\boldsymbol{r};\boldsymbol{v})}, \mathrm{S}_{f(\boldsymbol{r};\boldsymbol{v})}\right\rangle &= g\left(g^\sharp_{f(\boldsymbol{r};\boldsymbol{v})}\Phi_{f(\boldsymbol{r};\boldsymbol{v})}, g^\sharp_{f(\boldsymbol{r};\boldsymbol{v})}\Phi_{f(\boldsymbol{r};\boldsymbol{v})}\right) \geq 0.\end{aligned} \tag{5.73}$$

The mathematical steps are the same as in **Sec.** 4.2.

In the numerical setup that follows, two precise choices for the value of the *weight function* w will be made: in the first case, $w = 1$ will be considered, while, in the second case, the choice will be $w = \xi C^\alpha$ with $C = \sqrt{2\varphi^4}$, $\alpha = (\sqrt{21} - 1)/2$ and $\xi = 1$. The first choice corresponds to the one of a Fisher metric, while the second choice is related to the fact that it gives the correct Prandtl number for a monoatomic gas.

5.4.2 Compatibility with BGK models near equilibrium

The Steepest Entropy Ascent kinetic model coincides with a standard BGK model near equilibrium [BH13]. In order to prove this, it is necessary to define a *pseudo-*

Maxwellian distribution f_P that differs slightly from the *local Maxwellian* used in a standard BGK model. Based on f_P, it is possible to define the deviation of a certain probability density function f from equilibrium:

$$\phi = \frac{f - f_P}{f_P} \tag{5.74}$$

The *pseudo-Maxwellian* is then defined by

$$\mathtt{S}_P = -k_B \ln(b f_P) = \sum_{j=0}^{4} \beta_j^P \varphi^j, \tag{5.75}$$

where the coefficients β_j^P are given by

$$\left\langle f_p, w\varphi^j \right\rangle = \left\langle f, w\varphi^j \right\rangle \tag{5.76}$$

From the definition, the physical meaning of the *pseudo-Maxwellian* may be understood: whereas the local Maxwellian of a standard BGK model is the Maxwellian function having the same values of the density ρ, the macroscopic velocity $\boldsymbol{w}$ and the temperature T that the probability density function f has, the local pseudo-Maxwellian is the Maxwellian that has the same values of these macroscopic quantities *averaged* over the weight function w. Indeed, if a uniform Fisher is chosen ($w = 1$), then the pseudo-Maxwellian reduces to the standard local Maxwellian.

Multiplying **Eq.** 5.75 by $w\varphi^i$ and then averaging over the probability density function f, the following system of equations is obtained:

$$\sum_{j=0}^{4} \left\langle f, w\varphi^i\varphi^j \right\rangle \beta_j^P = \left\langle f, w\mathtt{S}_P\varphi^i \right\rangle \tag{5.77}$$

with $i = 0, ..., 4$. It may also be noted that, using the definition of the local pseudo-Maxwellian, **Eq.** 5.74, the local average may be written as

$$\langle f, \mathtt{a} \rangle = \langle f_p, (1 + \phi)\mathtt{a} \rangle \tag{5.78}$$

The entropy $\mathtt{S}(f)$ may then be written as

$$\mathtt{S} = \mathtt{S}_\phi + \mathtt{S}_P \tag{5.79}$$

with

$$\mathtt{S}_\phi = -k_B \ln\left(1+\phi\right) \tag{5.80}$$

Subtracting then **Eq.** 5.77 from **Eq.** 5.68b and using the expressions in **Eq.** 5.78 and **Eq.** 5.79, we obtain that

$$\sum_{j=0}^{4} \left\langle f_P, w(1+\phi)\varphi^i\varphi^j \right\rangle \left(\beta_j - \beta_j^P\right) = \left\langle f_P, w(1+\phi)\mathtt{S}_\phi \varphi^i \right\rangle \tag{5.81}$$

with $i = 0, ..., 4$.

For near-equilibrium situations, the deviation of the probability density function f from the local pseudo-Maxwellian goes to zero ($\phi \to 0$) and, thus, expressions in ϕ may be approximated with a Taylor expansion. Stopping the expansion to first order, from **Eq.** 5.80, we have that $\mathtt{S}_\phi = -k_B\phi$ because $\ln(1+x) \to x$ for $x \to 0$. The right-hand side of **Eq.** 5.81 may then be written as

$$-k_B \left\langle f_P, w\phi\varphi^i \right\rangle = -k_B \left(\left\langle f, w\varphi^i \right\rangle - \left\langle f_P, w\varphi^i \right\rangle\right) = 0 \tag{5.82}$$

where the first equality comes from the relationship between averages on f and averages on f_P illustrated in **Eq.** 5.78, while the second equality comes from **Eq.** 5.76.

As a consequence of this, it may be concluded from **Eq.** 5.81 that $\beta_j = \beta_j^P$ if the approximation is arrested to first order in ϕ. Using **Eq.** 5.79 and **Eq.** 5.75, the function Φ contained in the Steepest Entropy Ascent collision term may be re-written as

$$\Phi = \mathtt{S}_\phi - \sum_{j=0}^{4} (\beta_j - \beta_j^P)\varphi^j \tag{5.83}$$

Using the first-order approximations that have been obtained before, the first term becomes $\mathtt{S}_\phi = -k_B\phi$, while the second term goes to zero. The SEA kinetic model thus becomes

$$J_{\mathrm{SEA,w}}(f) = \frac{1}{k_B\tau} w f \Phi \longrightarrow -\frac{1}{\tau} w f_P \phi = \frac{1}{\tau} w (f_P - f). \tag{5.84}$$

This final expression is the one of a standard BGK model if $w = 1$ and is the one of a $\nu(\boldsymbol{v})$-BGK model if $w = \tau\nu(\boldsymbol{v})$. In this way, it is also proved that, for near equilibrium situations, the w-BGK models are equivalent to w-SEA models for any weight function w.

5.4.3 Numerical setup and results

The numerical simulation that was set up in [BH13] is based on the homogeneous relaxation from a highly non-equilibrium initial distribution. The initial distribution is indeed characterized by the mixture of two Maxwellians with the same density n_0, the same temperature T_0 and opposite initial velocities in the x-direction (with initial velocities in the other two coordinate directions being equal to zero) [BH13]. The initial Maxwellian f_0 may then be written as:

$$f_0 = \frac{1}{2}\left[f^{\text{eq}}(n_0, (v_i, 0, 0), T_0) + f^{\text{eq}}(n_0, (-v_i, 0, 0), T_0)\right] \tag{5.85}$$

where the initial velocity in the x-direction v_i is given by

$$v_i = \text{Ma}\sqrt{\frac{5k_BT_0}{3m}} \tag{5.86}$$

In the following **Fig.** 5.1 and **Fig.** 5.2, the two initial distributions f_0 for Ma $= 4$ and Ma $= 0.2$, respectively, are shown. It may be noted that the initial distribution in the first case is much farther from equilibrium than in the second case.

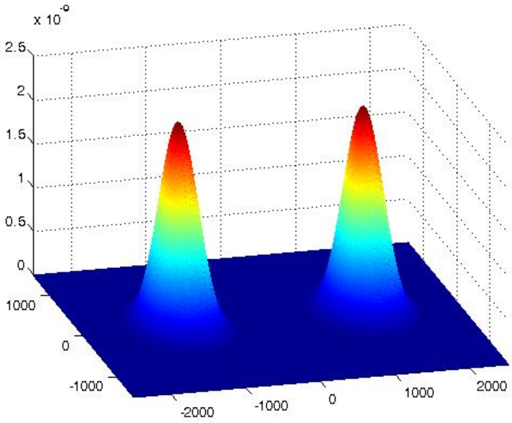

Fig. 5.1: Initial distribution f_0 with Ma $= 4$, given by the sum of two Maxwellians with a consistent difference in the macroscopic velocity in the x-direction; the graph is plotted at constant v_z and the Maxwellians are considered having unitary density ρ.

From the conservation requirements it may be deduced that the final state corresponds to a Maxwellian having density n_0, null macroscopic velocity and

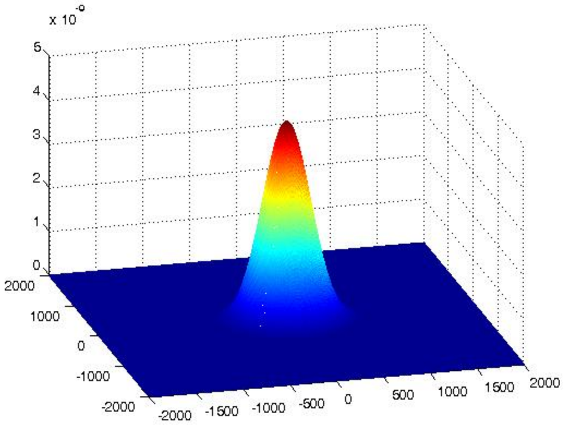

Fig. 5.2: Initial distribution f_0 with Ma $=0.2$, given by the sum of two Maxwellians with a small difference in the macroscopic velocity in the x-direction; the graph is plotted at constant v_z and the Maxwellians are considered having unitary density ρ.

a temperature $T_\infty = \frac{1+5\mathrm{Ma}^2}{9} T_0$. The values that have been considered for the constants are $k_B = 1.38 \cdot 10^{-23}$ J/K, $m = 6.63 \cdot 10^{-26}$ kg, $T_0 = 273$ K. In the following **Fig.** 5.3 and **Fig.** 5.4, the final equilibrium distributions for the situations previously illustrated are shown.

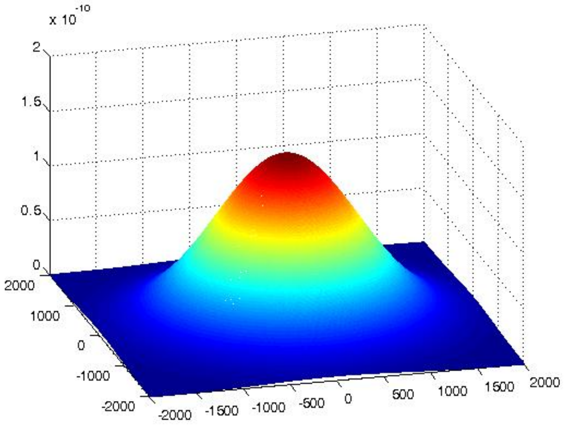

Fig. 5.3: Final equilibrium distribution for the initial situation with Ma $=4$: the final equilibrium temperature is about one order of magnitude higher than the initial one; the graph is plotted at constant v_z and the Maxwellian is considered having unitary density ρ.

As the two final Maxwellians may seem at first sight very similar (and, qual-

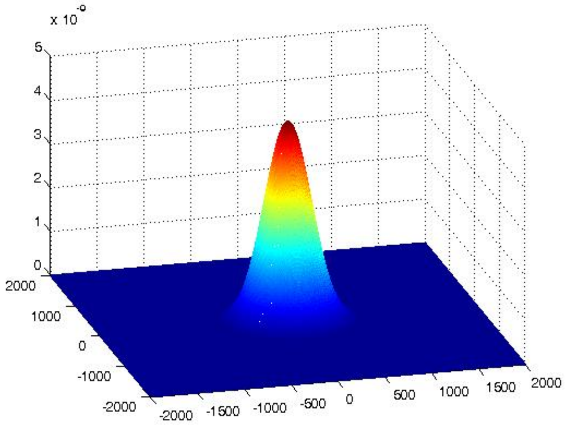

Fig. 5.4: Final equilibrium distribution for the initial situation with Ma $= 0.2$: the final equilibrium temperature is of the same order of magnitude of the initial one; the graph is plotted at constant v_z and the Maxwellian is considered having unitary density ρ.

itatively, indeed they are), it may be more useful for the reader to notice the difference between the initial and final distributions in the two examined cases: it may be seen in **Fig.** 5.5 and in **Fig.** 5.6 that the initial and final distributions are much closer to each other for the case with the lower initial Mach number.

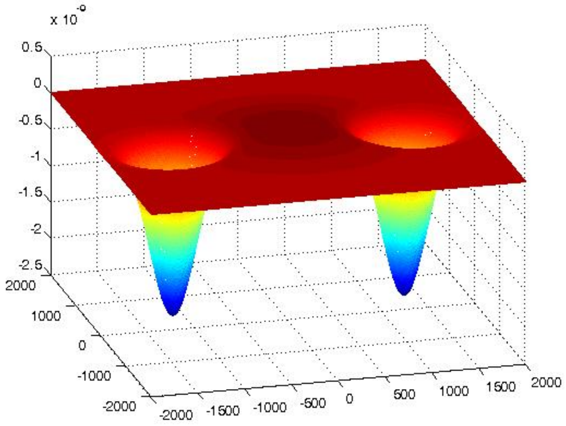

Fig. 5.5: Difference between final and initial distributions $f_\infty - f_0$ in the case with Ma $= 4$; the graph is plotted at constant v_z and the Maxwellians are considered having unitary density ρ.

As it has been shown by Struchtrup [St97], in the variable-collision frequency BGK-model where $\nu(\boldsymbol{v}) = \zeta C^\alpha$ with $C = |\boldsymbol{v} - \boldsymbol{w}|$, the correct Prandtl number for

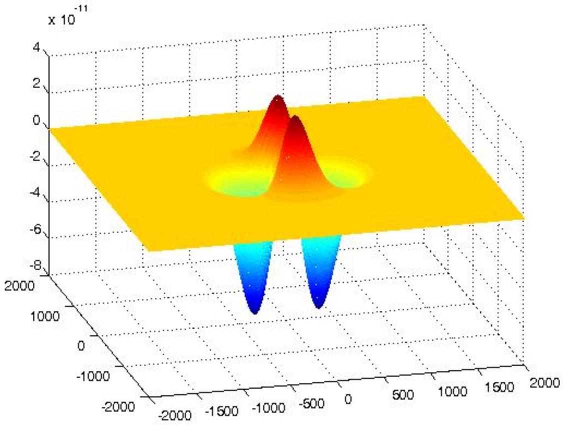

Fig. 5.6: Difference between final and initial distributions $f_\infty - f_0$ in the case with $\mathrm{Ma} = 0.2$; the graph is plotted at constant v_z and the Maxwellians are considered having unitary density ρ.

a monoatomic gas $Pr = 2/3$ is recovered with $\zeta = 1$ and $\alpha = \frac{\sqrt{21}-1}{2}$. Because of this, in the numerical simulations, the weight function is firstly assigned the value $w = 1$ (Fisher metric) and then the value $w = C^\alpha$ with $\alpha = \frac{\sqrt{21}-1}{2}$. In the following figures, the results that have been obtained with the numerical simulations are shown: **Fig.** 5.7 to **Fig.** 5.10 refer to the case with an initial condition with $\mathrm{Ma} = 4$ in **Eq.** 5.86, while **Fig.** 5.11 to **Fig.** 5.14 refer to the case with $\mathrm{Ma} = 0.2$ in **Eq.** 5.86.

In each one of the figures that are reported in the present paragraph and that refer to the recent work by Beretta and Hadjiconstantinou, the results for the two SEA kinetic models are compared with the results for the standard BGK-model, the velocity-dependent collision frequency BGK-model and the assumed *exact solution*, obtained by Professor Hadjiconstantinou of MIT with a Direct Simulation Monte Carlo (DSMC) code using Hard-Sphere (HS) dynamics [Bi94]. The results for the first four models have been obtained using Matlab and, more specifically, using the *ode45* solver. During the simulation, the symmetry of the problem has been exploited to simplify it in the velocity-space from a three-dimensional one into a two-dimensional one: $v_r = \sqrt{v_y^2 + v_z^2}$. The following discretisations were adopted: v_x was discretised in the range $\pm 10\sqrt{2k_B T_0/m}$ using 30 cells, while c_r was discretised in the range $0 \leq v_r \leq 10\sqrt{2k_B T_0/m}$ using at least 180 cells.

In the following figures [BH13], significant quantities and moments obtained from the probability density function f are plotted:

- in **Fig.** 5.7 and **Fig.** 5.11, an adimensionalized entropy is plotted against an adimensionalized time;
- in **Fig.** 5.8 and **Fig.** 5.12, an adimensionalized *parallel temperature* (temperature related to the square of the velocity in the x-direction) is plotted against an adimensionalized entropy;
- in **Fig.** 5.9 and **Fig.** 5.13, an adimensionalized fourth moment of the velocity in the x-direction is plotted against an adimensionalized entropy;
- in **Fig.** 5.10 and **Fig.** 5.14, an adimensionalized fourth moment of the velocity in the x-direction is plotted against an adimensionalized *parallel temperature.*

The numerical results show some important evidence:

- both BGK models, the standard one and the variable-collision-frequency one, show a good agreement with the *exact solution* obtained through the Montecarlo method;
- Steepest Entropy Ascent models show a significantly lower agreement with the *exact solution* with respect to the corresponding BGK-models;
- between the two Steepest Entropy Ascent models, better agreement with the Hard-Sphere solution and with the BGK-models is given by the one with $w = C^{\alpha}$, while a significantly poorer performance is given by the uniform-metric one;
- the differences between the various models are consistently reduced for the relaxation from a near-equilibrium situation ($\mathrm{Ma} = 0.2$) with respect to the relaxation from a far-equilibrium situation ($\mathrm{Ma} = 4$);
- the numerical results prove that, for near-equilibrium situations, the two SEA kinetic models coincide with the respective BGK-models, as the time-evolution is shown to be the same.

Some considerations may be then made on the result of the numerical experiment, taking in consideration also the first part of the Chapter, related to the interpretation of the Boltzmann Equation by the Steepest Entropy Ascent dynamic model and the GENERIC model.

- First of all, it is shown that SEA dynamics, reformulated from a Differential Geometry standpoint in the previous Chapter, actually provides consistent kinetic models, as both families of SEA models that have been proposed satisfy the fundamental conditions required for a kinetic model, that is, the conservation of the collision invariants and the verification of the H-theorem.

- It is further shown the SEA dynamics applied to the Boltzmann Equation is characterized by the presence of one 'degree of freedom', related to the choice of the metric. From a geometric standpoint, the choice of the metric is related to the velocity of evolution in state-space, while, from an applicative standpoint in the Boltzmann Equation kinetic model, the choice of the metric is related to the 'coefficient' multiplying the term ruling the dissipative evolution.

- In the SEA models that have been analyzed, two precise choices for the metric have been made, the uniform one and the variable-collision-frequency one; these choices have been dictated by two factors: on one hand, simplicity for the calculations and, on the other hand, precise calibration to the corresponding BGK models near equilibrium (it was indeed shown analytically that, near equilibrium, the SEA models converge to the corresponding BGK models).

- However, it is shown in the calculations by Beretta and Hadjiconstantinou that these two choices for the metric do not yield satisfying results far from equilibrium (while, near equilibrium, as it was shown analytically, they are coincident with the corresponding BGK model).

- In our opinion, this is not an element against the idea of a Steepest Entropy Ascent evolution of a system governed by the Boltzmann Equation because the geometric equivalence of SEA and GENERIC and the fact that GENERIC has identified an exact metric that reproduces the Boltzmann Equation, suggest that there is a precise metric that makes the collision term of the Equation a SEA term.

- In our opinion, then, the real problem does not reside in the validity of the idea of a SEA evolution of the system, but in the choice of a metric, that, in the examined case, has been limited and partial and has to satisfy two opposing needs: being able of reproducing correctly real behaviours and being sufficiently simple to generate a 'facilitation' in the calculations, also considering more complex cases than the one that has been presented, for which spatial dependence does not exist.

- Thus, the next step should be to examine metrics that are calibrated on models and methods of approximate solution that are valid also far from equilibrium, such as the Chapman-Enskog expansion.

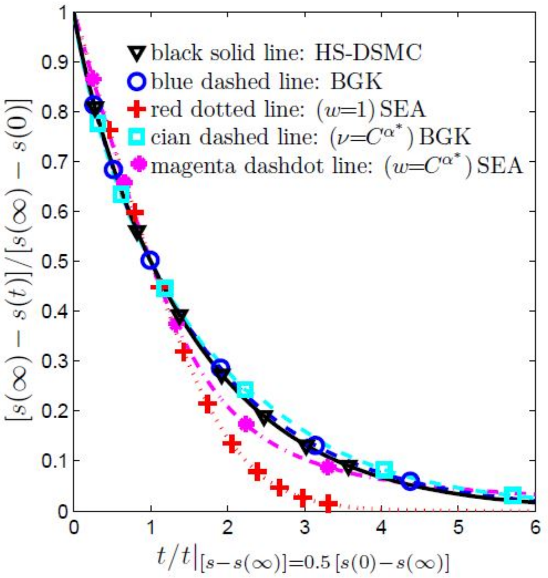

Fig. 5.7: Adimensionalized entropy as a function of an adimensionalized time for the different collision operators for relaxation from the initial non-equilibrium distribution with Ma = 4. The time is scaled on the basis of the time where the entropy change is half of the total change due to the relaxation.

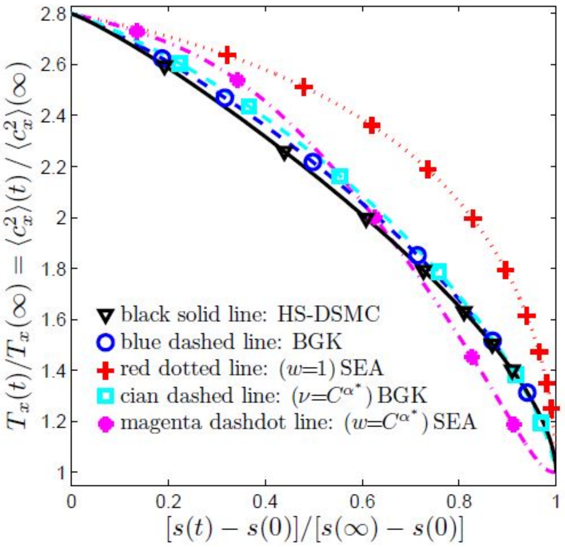

Fig. 5.8: Adimensionalized *parallel temperature* as a function of an adimensionalized entropy for different collision operators for relaxation from the non-equilibrium distribution with $\mathrm{Ma} = 4$.

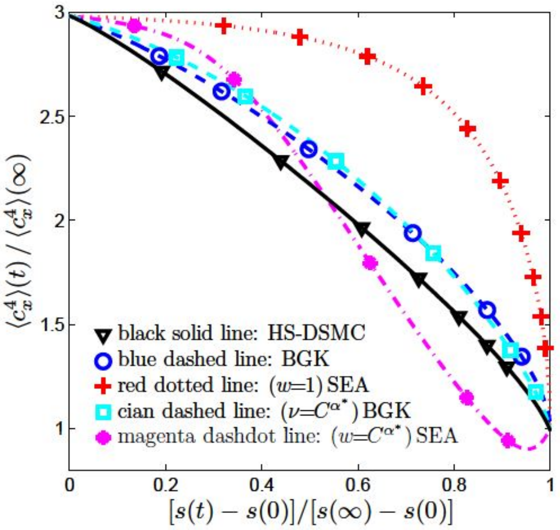

Fig. 5.9: Adimensionalized fourth moment of the velocity in the x-direction as a function of an adimensionalized entropy for different collision operators for relaxation from the initial non-equilibrium state with $\mathrm{Ma} = 4$.

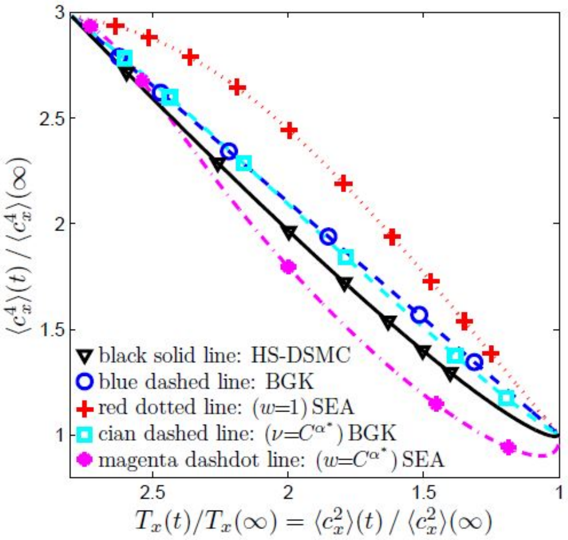

Fig. 5.10: Adimensionalized fourth moment of the velocity in the x-direction as a function of an adimensionalized *parallel temperature* for different collision operators for relaxation from the initial non-equilibrium distribution with Ma = 4.

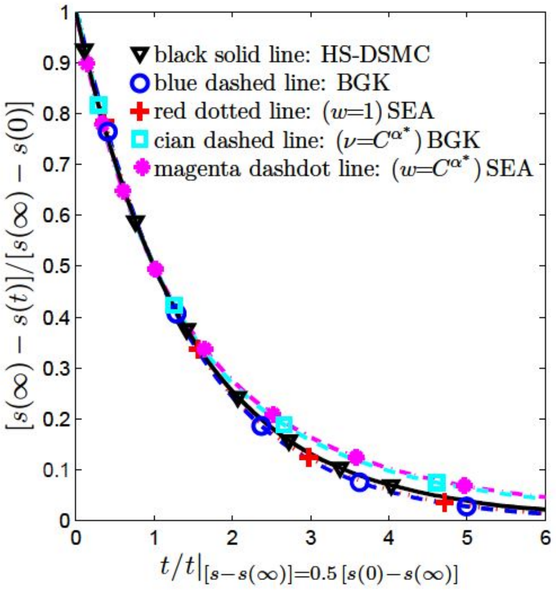

Fig. 5.11: Normalized entropy as a function of an adimensionalized time for the different collision operators for relaxation from the initial non-equilibrium distribution with Ma = 0.2. The time is scaled on the basis of the time where the entropy change is half of the total change due to the relaxation.

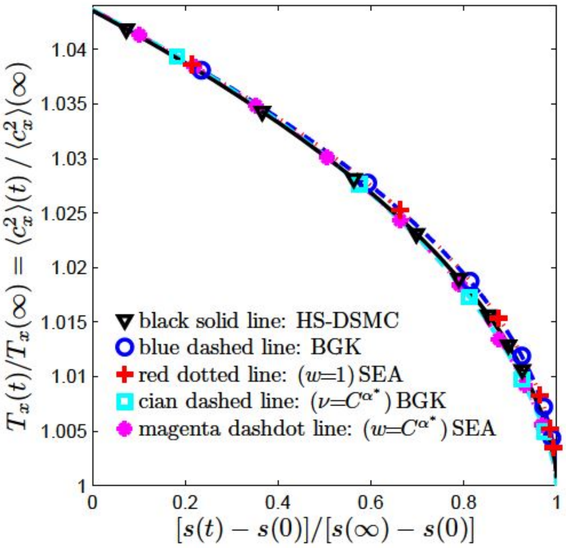

Fig. 5.12: Adimensionalized *parallel temperature* as a function of an adimensionalized entropy for different collision models for relaxation from the non-equilibrium distribution with $\mathrm{Ma} = 0.2$.

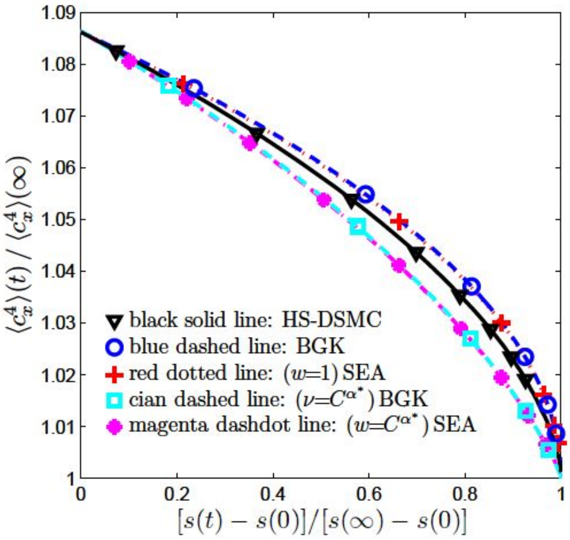

Fig. 5.13: Adimensionalized fourth moment of the velocity in the x-direction as a function of an adimensionalized entropy for different collision operators for relaxation from the initial non-equilibrium state with $\mathrm{Ma} = 0.2$.

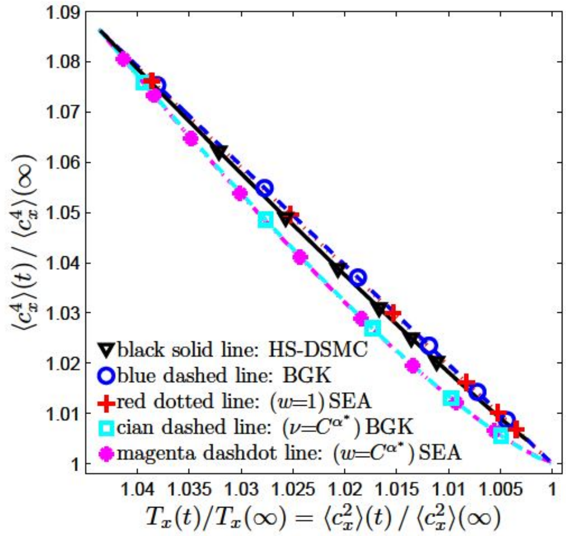

Fig. 5.14: Adimensionalized fourth moment of the velocity in the x-direction as a function of an adimensionalized *parallel temperature* for different collision operators for relaxation from the initial non-equilibrium distribution with $\mathrm{Ma} = 0.2$.

References

[Be13] G. P. Beretta. “Steepest-Entropy-Ascent and Maximal-Entropy-Production Dynamical Models of Irreversible Relaxation to Stable Equilibrium from Any Non-Equilibrium State. Unified Treatment for Six Non-Equilibrium Frameworks”. In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013*. Ed. by M. Pilotelli and G. P. Beretta. Brescia: Cartolibreria Snoopy, 2013, pp. 100–109 (cit. on p. 133).

[BH13] G. P. Beretta and N. G. Hadjiconstantinou. “Steepest Entropy Ascent Models of the Boltzmann Equation. Comparisons with Hard-Sphere Dynamics and Relaxation-Time Models for Homogeneous Relaxation from Highly Non-Equilibrium States”. In: *Proceedings of the ASME 2013 International Mechanical Engineering Congress and Exposition, IMECE 2013*. 2013 (cit. on pp. 129, 132–134, 137, 140, 142).

[Bi94] G. A. Bird. *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*. Oxford: Clarendon, 1994 (cit. on p. 140).

[E98] B. J. Edwards. “An analysis of single and double generator thermodynamic formalisms for the macroscopic description of complex fluids”. In: *Journal of Non-Equilibrium Thermodynamics* 23.4 (1998), pp. 301–333 (cit. on p. 122).

[Mo84] P. J. Morrison. *Some Observations Regarding Brackets and Dissipation*. Research rep. PAM-228. Center for Pure and Applied Mathematics, University of California, Berkeley, 1984 (cit. on pp. 117, 120).

[MT12] M. Materassi and E. Tassi. “Metriplectic framework for dissipative magneto-hydrodynamics”. In: *Physica D: Nonlinear Phenomena* 241.6 (2012), pp. 729–734 (cit. on pp. 123, 125).

[MW82] J. E. Marsden and A. Weinstein. “The Hamiltonian structure of the Maxwell-Vlasov equations”. In: *Physica D: Nonlinear Phenomena* 4.3 (1982), pp. 394–406 (cit. on p. 126).

[Ö05] H. C. Öttinger. *Beyond Equilibrium Thermodynamics*. 1st ed. Hoboken (New Jersey): Wiley, 2005 (cit. on pp. 117, 122).

[Ö97] H. C. Öttinger. “GENERIC formulation of Boltzmann’s kinetic equation”. In: *Journal of Non-Equilibrium Thermodynamics* 22.4 (1997), pp. 386–391 (cit. on pp. 126, 130–131).

[Si87] S. Sieniutycz. “From a least action principle to mass action law and extended affinity”. In: *Chemical Engineering Science* 42.11 (1987), pp. 2697–2711 (cit. on p. 132).

[St97] H. Struchtrup. “The BGK-model with velocity-dependent collision frequency”. In: *Continuum Mechanics and Thermodynamics* 9.1 (1997), pp. 23–31 (cit. on p. 139).

6

Conclusions and future developments

6.1 Conclusions

6.1.1 Development of the Work

The work developed in the present thesis is focused on the dynamical modelling of thermodynamic systems, namely on the branch of Thermodynamics that aims at describing the causes and the effects of the motion of a system. As highlighted in the introduction, a dynamical law of evolution such as Newton's Law for Classical Mechanics or Schrödinger's Equation for Quantum Mechanics, does not exist in the field of Thermodynamics. According to the view of Gyftopoulos and Beretta, such a law should yield as corollaries the statements of the two Principles of Thermodynamics, which are proved features of dynamical evolution. Such a law would not only be a milestone in the field of Physics, but it would help in modelling the time evolution of thermodynamic systems. In particular, through its application in Kinetic Theory it might help to find an easier model for the collision term of the Boltzmann Equation, thus generating a tangible improvement in the many fields where the Boltzmann Equation is used to schematize physical situations: rarefied gas dynamics, the description of electrons in a conductor and phonons in an insulator, the time-evolution of neutrons in a nuclear reactor and the description of plasmas.

In this context, the objective of the Thesis is threefold:

- first of all, to carry out a systematic review of some different formulations of one of the most frequently proposed dynamical principles for the evolution of thermodynamic systems, that is, the Maximum Entropy Production Principle;

- second, to identify and discuss similarities and differences between the Steepest Entropy Ascent (SEA) dynamical model, initially proposed by Beretta [B81] in a quantum framework and recently adapted to meso- and macroscopic systems [B13], and the GENERIC (General Equation for Non-Equilibrium Reversible-Irreversible Coupling) formalism, developed, among others, by Grmela and Öttinger [GÖ97];

- third, to understand how these two models apply to Kinetic Theory and how the different terms of the Boltzmann Equation are interpreted using the *building blocks* of the two theories, in order to develop new kinetic models based on these principles.

Chapter 2 illustrates Classical Non-Equilibrium Thermodynamics and Kinetic Theory, that is, the *theories* needed to understand the successive chapters and *used* to *validate* the dynamic models that are presented in the thesis.

Chapter 3 deals with the first one of the goals listed above. First of all, the link between the Maximum Entropy Production Principle and Prigogine's Minimum Entropy Production Principle is clarified: if we consider its interpretation in SEA, the first one is usually understood to be a principle governing the evolution of a thermodynamic system *locally*, at a fixed instant of time, when the thermodynamic system, among all possible directions, chooses the one that *maximizes* the entropy production. On the other hand, the Principle of Minimum Entropy Production introduced by Prigogine [P47] refers *globally* to the whole system under study and the *minimization* of the entropy production has to be understood in the sense that the total entropy production of the system decreases in time (time is not fixed). Moreover, while the Maximum Entropy Production Principle appears to be valid with no particular restrictions, the Minimum Principle has been proved to hold under restrictive hypotheses such as linearity of the relations between forces and fluxes, validity of Onsager's reciprocal relations, constancy of the phenomenological coefficients and time-independence of the constraints on the border of the body.

Successively, Ziegler's theory, which may be considered the first complete statement of the Maximum Entropy Production Principle, is illustrated. It constitutes an attempt to *geometrize* Non-Equilibrium Thermodynamics: indeed, if entropy production is maximized subject to the constraint of being the product of forces times fluxes, the final output is that the gradient of the entropy production with respect to the fluxes should be parallel to the forces. This condition, for reasons that are illustrated in the Chapter, is called *orthogonality condition*. It is further shown that Ziegler's principle and linear relations between forces and fluxes entail Onsager's reciprocity. In the final part of the Chapter, Edelen's theory is illustrated and it is shown how, by proposing a *decomposition* of the flux vector in two parts, his theory may be seen as a precursor of GENERIC's two-generator formalism that is analyzed in the successive Chapter.

Chapter 4 deals with the second one of the objectives listed above: the comparison between SEA and GENERIC, two structures that have been motivated by the search for Non-Equilibrium Thermodynamics formulations that are fully compatible with the Second Law of Thermodynamics. The link with the previous Chapter lies in the fact that the Steepest Entropy Ascent dynamic model is of the *entropy-gradient* type, that is, it states that a thermodynamic system evolves in the direction of maximum-entropy-production compatible with the conservation requirements. Moreover, SEA aims at *geometrizing* the time-evolution of the system. Because of this, the Steepest Entropy Ascent model has been compared to the GENERIC model, which is, up to date, among the most structured approaches towards the *geometrization* of Non-Equilibrium Thermodynamics. The first part of the Chapter deals with the study of the rationalization, through the formalism of Differential Geometry, of the equations of Hamiltonian Mechanics. This is a field that has developed in the XX Century, taking the name of Geometric Mechanics, and is set in the 'arenas' of symplectic manifolds and of their generalisation, that is, Poisson manifolds. This study has been particularly useful for the development of the second part of the Chapter for two reasons:

- first of all, it has allowed us to understand how Differential Geometry represents a consistent formalism and setting for a rigorous mathematical formulation of physical theories;
- secondly, it has allowed us to understand concepts, such as the one of Poisson

manifolds, that are at the base of the 'differential geometric formulation' of the GENERIC formalism.

6.1.2 Reformulation of SEA and comparison with GENERIC

With this background, Differential Geometry has been considered as the ideal setting to carry out the comparison between GENERIC and SEA. Steepest Entropy Ascent dynamics was then reformulated in a more abstract and mathematical way using the 'differential geometric language', making it more immediately comparable to the formalism developed, among others, by Öttinger e Grmela, which had already been expressed in these terms. In our opinion, this constitutes one of the most innovative outputs of the present Thesis.

It has thus been showed that the two dynamic models show similar patterns as both may be considered as belonging to the *maximum-entropy-producing* or the *entropy-gradient* type. That is, the time evolution of the thermodynamic system, in a suitable manifold, goes in the direction of the entropy gradient. However, specific differences must be pointed out and are listed as follows.

- GENERIC is characterized by a more explicit separation between the reversible and irreversible terms, while SEA only considers explicitly the irreversible part of the dynamics.

- SEA chooses a non-degenerate Riemannian metric tensor that regulates the evolution of the dissipative part of the equation, while GENERIC is set in a more structured environment, as it uses a Poisson structure to describe the reversible part of the dynamics and a degenerate co-Riemannian structure to describe the irreversible part of the dynamics;

- As SEA dynamics is set in a less-structured manifold, the constraints related to the conservation laws must be imposed at a later phase (by forcing the thermodynamic process to be locally orthogonal to the gradients of the conserved properties), while, in the GENERIC formalism, conservation laws are already *built-in* in the degeneracy of the two structures that characterize the manifold.

- In SEA dynamics, the choice of a non-degenerate metric allows to univocally define gradients, while in the GENERIC formalism, the choice of a degenerate metric makes it impossible to define a metric and, thus, a gradient, unless a further condition (the *Leibniz identity*) on the dissipative bracket is imposed.

- By imposing this further condition on the dissipative bracket, a metric and, thus, gradients may be defined on *metric leaves*, that is, *surfaces* where the dissipative GENERIC evolution of a system takes place. These surfaces are characterized by constant values of the energy and the conserved quantities and it is shown that the dynamics that takes place on them is of the *entropy-gradient* type because the dissipative vector points in the direction of maximum entropy increase.

As a concluding remark, we may state that, after the reformulation of Steepest Entropy Ascent dynamics using Differential Geometry formalism, it is shown that SEA is more general because it is less structured than GENERIC. Moreover, by imposing the *Leibniz identity* on the dissipative bracket, GENERIC becomes Steepest Entropy Ascent on *metric leaves*.

6.1.3 Realizations of the dynamical models

Chapter 5 illustrates the realizations of the two dynamic models in different physical theories. It is shown how the governing equations of classical hydrodynamics and magnetohydrodynamics are interpreted by the GENERIC framework. Considering also the GENERIC *interpretation* of the Boltzmann equation, the main goal of the GENERIC formalism emerges clearly: reproducing known equations of motion of known physical theories by casting them in a single abstract form. The authors claim that through this formalism, a deep insight of the physical phenomena that are described may be gained and the solution of important problems may be facilitated. On the other hand, in its realization in Kinetic Theory, SEA shows that its goal is not reproducing the equation in its *exact* form through its building blocks, but developing kinetic models for its resolution that depend on the choice of the metric (that is, on the specific choice that is made for the Riemannian metric tensor).

In the last part of the Chapter, numerical results obtained by Beretta and Hadjiconstantinou for the relaxation from two non-equilibrium states through the Boltzmann Equation are shown [BH13]. The *exact* solution, calculated through a Montecarlo method, is compared with the time-evolutions obtained through two different BGK models and two different SEA Kinetic Models (that differ for the choice of the metric). The SEA models are shown to satisfy the usual requirements needed for kinetic models, that is, the conservation of collision invariants and the verification of the H-theorem. Moreover, they are proved to converge to the correspondent BGK models in a near-equilibrium approximation. The following considerations may be made on the numerical results:

- BGK models show a good agreement with the *exact* solution both for near-equilibrium and far-equilibrium relaxations;
- SEA models do not show a good agreement with the *exact* solution for far-equilibrium relaxation, while in near-equilibrium, as they converge to the correspondent BGK model, results are more satisfactory;
- between the two SEA models, the one characterized by the variable-collision-frequency is the best, while the uniform metric model clearly yields poor results.

In the light of these results and of the conclusions of Chapter 4, some comments may be made:

- the key aspect in the SEA models, that were shown to be consistent with the requirements for kinetic models, is the 'degree of freedom' related to the choice of the metric, which, in practical terms, influences the 'coefficient' multiplying the term ruling the dissipative evolution;
- in the two models that have been examined in the numerical examples by Beretta and Hadjiconstantinou, the choice of the uniform and variable-collision-frequency metrics has been dictated by the reasons of simplicity (they are simply a numerical coefficient in front of the dissipative term) and by a precise *calibration* with BGK models near-equilibrium;
- in our opinion, the poor results that are yielded far from equilibrium are not an element against the idea of a Steepest Entropy Ascent evolution of a system governed by the Boltzmann Equation because the geometric equivalence of SEA and GENERIC and the fact that GENERIC has identified an exact metric that reproduces the Boltzmann Equation, suggest that there is a precise metric that makes the collision term of the Equation a SEA term;
- in our opinion, then, the real problem does not reside in the validity of the idea of a general SEA evolution of the system, but in the choice of a metric and, thus, in the choice of a precise family of SEA methods; in the examined cases, the choice of the metric, which has two main needs, being able of reproducing correctly real behaviours and being sufficiently simple to

generate a facilitation in the calculations, has been limited to only two cases which turned out to be not satisfactory.

6.2 Future developments

Many of the topics that have been considered in the present thesis, especially regarding the *geometrical comparison* of SEA and GENERIC and the SEA kinetic models for the Boltzmann Equation, may either give spark to further ideas or deserve a deeper analysis.

Among the open problems, there are the following ones:

- a systematic identification of the conserved properties of a system, a topic not examined in this work and related to the choice of the particular *level of description* [BZ11], still does not exist and is one of the most important points to be addressed in Kinematics;

- the modelling of interactions still remains an open issue, as mathematical devices needed to perform such a modelling are under study or have been recently proposed, but there is not a consolidated theory like for isolated systems;

- regarding the SEA modelling of the Boltzmann Equation, whose numerical results are unsatisfactory, a better choice of the metric is needed in order to improve SEA kinetic models.

The following is instead a list of the ideas sparked by the present work:

- the reversible-part formalism that GENERIC borrows from Geometric Mechanics might be 'transferred' to SEA in order to have a more complete model that explicitly considers Hamiltonian dynamics;

- imposing the dissipative bracket in GENERIC to satisfy the further condition of the *Leibniz identity* in order to have a non-degenerate metric on the metric leaves and to be able to define a horizontal gradient is an idea that, in practical realizations in equations, could be tested by a computer algorithm as Kröger, Hütter, and Öttinger did for the Jacobi identity in [KHÖ01], as it is analytically too burdensome to perform;

- suggestions for this *improved* metric for an efficient kinetic model have not been found in the GENERIC modelling as it simply aims at reproducing the exact equations governing the time evolution of the system, while help could be found in the emerging field of Information Geometry or by *calibrating* the metric with respect to approximate methods of solution for the Boltzmann Equation that are valid farther from equilibrium with respect to the BGK methods.

References

[B13] G. P. Beretta. “Steepest-Entropy-Ascent and Maximal-Entropy-Production Dynamical Models of Irreversible Relaxation to Stable Equilibrium from Any Non-Equilibrium State. Unified Treatment for Six Non-Equilibrium Frameworks”. In: *Proceedings of the 12th Joint European Thermodynamics Conference, JETC 2013*. Ed. by M. Pilotelli and G. P. Beretta. Brescia: Cartolibreria Snoopy, 2013, pp. 100–109 (cit. on p. 152).

[B81] G. P. Beretta. “On the general equation of motion of quantum thermodynamics and the distinction between quantal and nonquantal uncertainties”. thesis. 1981 (cit. on p. 152).

[BH13] G. P. Beretta and N. G. Hadjiconstantinou. “Steepest Entropy Ascent Models of the Boltzmann Equation. Comparisons with Hard-Sphere Dynamics and Relaxation-Time Models for Homogeneous Relaxation from Highly Non-Equilibrium States”. In: *Proceedings of the ASME 2013 International Mechanical Engineering Congress and Exposition, IMECE 2013*. 2013 (cit. on pp. 155–156).

[BZ11] G. P. Beretta and E. Zanchini. “Rigorous and General Definition of Thermodynamic Entropy”. In: *Thermodynamics*. Ed. by M. Tadashi. InTech, 2011, pp. 23–50 (cit. on p. 157).

[GB05] E. P. Gyftopoulos and G. P. Beretta. *Thermodynamics: Foundations and Applications*. 2nd ed. (1st ed., Macmillan, 1991). Mineola (New York): Dover Publications, 2005 (cit. on p. 151).

[GÖ97] M. Grmela and H. C. Öttinger. “Dynamics and thermodynamics of complex fluids. I. Development of a general formalism”. In: *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics* 56.6 (1997), pp. 6620–6632 (cit. on p. 152).

[KHÖ01] M. Kröger, M. Hütter, and H. C. Öttinger. “Symbolic test of the Jacobi identity for given generalized 'Poisson' bracket”. In: *Computer Physics Communications* 137 (2001), pp. 325–340 (cit. on p. 157).

[P47] I. Prigogine. *Étude thermodynamique des phénomènes irréversibles.* Liège: Desoer, 1947 (cit. on p. 152).

A

Compendium of Differential Geometry

The purpose of the present appendix is to briefly introduce the basic concepts of Differential Geometry that are needed for the comprehension of the part of the thesis regarding the geometrical formulation of the Steepest Entropy Ascent and GENERIC theories. Differential Geometry is usually not present in regular Master of Science curricula in engineering and it has been thus considered necessary to present at least a smattering of the subject. The following pages do not pretend to be exhaustive and we anticipate that, when considered necessary, mathematical rigour will be sidelined in favour of a more accessible and colloquial presentation because of the need, on one hand, to summarize such a wide variety of topics and, on the other, to offer an immediate mean of comprehension to a reader that is assumed to be totally unfamiliar with the subject. The following treatise, including notations and figures, refers to the book *Introduction to Smooth Manifolds* by John M. Lee.

A.1 Basic concepts

Differential Geometry is defined as the "*study of smooth manifolds endowed with some structure and of their properties that are invariant under structure-*

preserving maps".

A.1.1 Manifolds and charts

The basic idea behind smooth manifolds is that they are spaces that *locally look like* a Euclidean space and are endowed with a *smooth structure* that allows to make sense of derivatives. Basically, as derivations and integrations cannot be made directly *on* the manifold, the manifold has to be *mapped* to a Euclidean space where these calculations are perfectly defined. This is not a forced choice: in place of Euclidean space, one may choose another vector space where differential operations are defined. In general, this vector space takes the name of ***model space***. Hence, one has the notions of:

- *Fréchet manifolds*, whose model space is a Fréchet space;
- *Banach manifolds* (a particular case of Fréchet manifolds), whose model space is a Banach space;
- *Hilbert manifolds* (a particular case of Banach manifolds), whose model space is a Hilbert space.

In the following discussion, we shall consider only the Euclidean case, which gives rise to *topological manifolds* and is a subcase of the Hilbert one. However, all definition which do not directly involve the dimension of the model space can be extended to Hilbert manifolds with minor modifications. This is not an unimportant point, since – in our thesis – we have considered some cases in infinite dimensions.

Moreover, although the model space may also be a complex vector space, or a quaternionic vector space, etc., only the real case will be considered.

In order for a topological space M to be a ***topological manifold of dimension n***, it has to be characterized by three properties:

- M has to be an ***Hausdorff space***: for every pair of distinct points $p, q \in M$, there are disjoint open subsets $U, V \subseteq M$ such that $p \in U$ and $q \in V$, that is, a space for which two disjoint neighbourhoods may be defined for two different points.
- M is ***second-countable***: there exists a countable basis for the topology of M.

- M is ***locally Euclidean of dimension n***: each point $p \in M$ possesses a neighbourhood which is homeomorphic to an open subset of $\mathbb{R}^n$.

The third property means that, for each $p \in M$, there is:

- an open subset $U \subseteq M$ containing p;
- an open subset $\hat{U} \subseteq \mathbb{R}^n$;
- a homeomorphism $\varphi\colon U \to \hat{U}$.

The key to the *mapping* of a manifold to the *equivalent* Euclidean space is a ***coordinate chart***.

Definition A.1. Let M be a topological n-manifold. A ***coordinate chart*** (or just a ***chart***) on M is a pair (U, φ), where U is an open subset of M and $\varphi\colon U \to \hat{U}$ is a homeomorphism from U to an open subset $\hat{U} = \varphi(U) \subseteq \mathbb{R}^n$ (**Fig.** A.1). Given a chart (U, φ), we call the set U a ***coordinate domain***, or a ***coordinate neighbourhood*** of each of its points. The map φ is called a ***(local) coordinate map***, and the components functions $(x^1, \dots, x^n)$ of φ, defined by $\varphi(p) = (x^1(p), \dots, x^n(p))$ are called ***local coordinates*** on U.

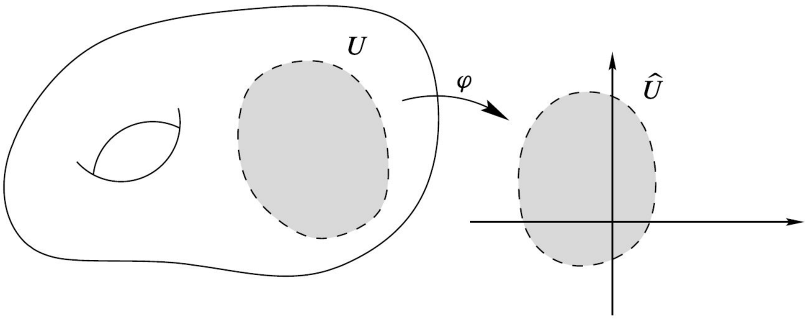

Fig. A.1: Coordinate chart on a manifold: homeomorphism from an open subset of the manifold M to an open subset of the Euclidean space $\mathbb{R}^n$.

A.1.2 Smooth structures

The definition of topological manifold that has been given is sufficient to study its topological structure, but not to make sense of the instruments of calculus on it. Thus, a new kind of manifold needs to be introduced: a *smooth manifold*, which is a topological manifold with an extra structure. In the present treatise we assume the concept of ***smoothness*** to be synonymous with ***infinitely differentiable***, that is $\boldsymbol{C^\infty}$. Let M be a topological n-manifold. If (U, φ), (V, ψ) are two charts such that $U \cap V \neq \emptyset$, the composite map $\psi \circ \varphi^{-1} \colon \varphi(U \cap V) \to \psi(U \cap V)$ is called the ***transition map from φ to ψ*** (**Fig.** A.2). Two charts (U, φ) and (V, ψ) are said to be ***smoothly compatible*** if either $U \cap V = \emptyset$ or the transition map $\psi \circ \varphi^{-1}$ is a ***diffeomorphism***, that is, it is a bijective map and has a smooth inverse map. Since (U, φ) and (V, ψ) are open subsets of $\mathbb{R}^n$, smoothness of this map is to be interpreted in the sense of having continuous partial derivatives of all orders.

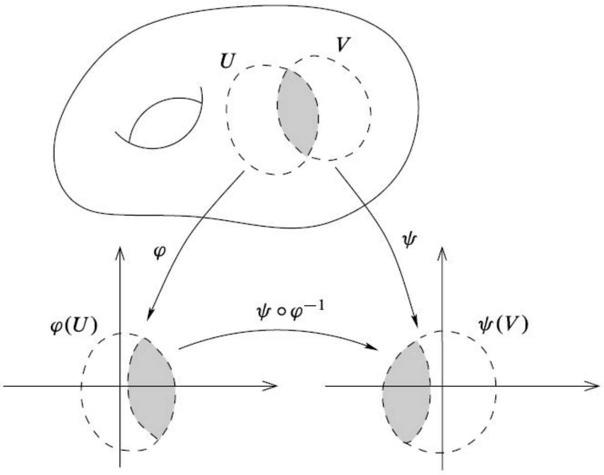

Fig. A.2: Transition map between two open subsets of $\mathbb{R}^n$: if it is a diffeomorphism, the two charts are smoothly compatible.

An ***atlas*** is then defined as a collection of charts whose domains cover all M. The *atlas* is called a ***smooth atlas*** if any two charts are *smoothly compatible* with each other. The idea is then to define a *smooth structure* on the manifold by assigning a smooth atlas and then to define a function $f \colon M \to \mathbb{R}^n$ to be *smooth* if and only if the composition $f \circ \varphi^{-1}$ is smooth in the ordinary sense for every coordinate chart of the atlas. In this way, it is possible to define differentiability for the functions on the manifold. A problem that arises is due to the fact that there are many possible atlases that determine the same smooth structure: among these, an atlas A is defined as ***maximal*** if it is not properly contained in any larger smooth atlas.

At this point an important concept may be defined: if M is a topological manifold, a ***smooth structure on $\boldsymbol{M}$*** is a maximal smooth atlas. A ***smooth manifold*** is then a pair (M, A), where M is a topological manifold and A is a smooth structure on M. Then, if M is a smooth manifold, any chart (U, φ) contained in the maximal smooth atlas is called ***smooth chart***, and the relative coordinate map φ is called ***smooth coordinate map***.

Once the concept of *smooth structure* has been clarified, also ***smooth functions*** and ***smooth maps*** may be defined: even if the two words are sometimes used as synonyms, we refer to *functions* as maps whose codomain is $\mathbb{R}^n$ and *maps* as, more generally, maps between manifolds.

If M is a smooth n-manifold, k is a non-negative integer, and $f\colon M \to \mathbb{R}^k$ is any function, we say that f is a ***smooth function*** if, for every $p \in M$, there exists a smooth chart (U, φ) for M whose domain contains p and such that the composite function $f \circ \varphi^{-1}$ is smooth on the open subset $\hat{U} = \varphi(U) \subseteq \mathbb{R}^n$ (**Fig.** A.3).

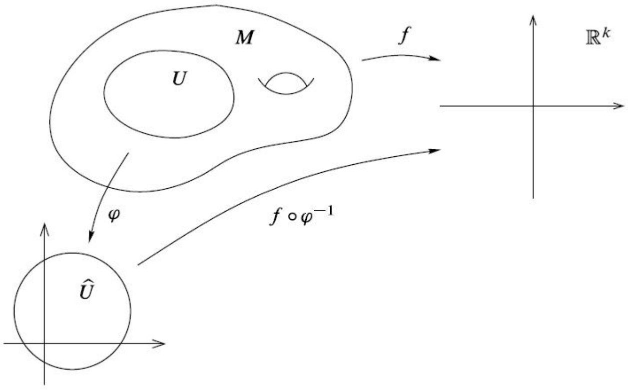

Fig. A.3: Definition of smooth function.

If M and N are a smooth manifolds and $F\colon M \to N$ is a map, we say that F is a ***smooth map*** if, for every $p \in M$, there exists smooth charts (U, φ) containing p and (V, ψ) containing $F(p)$ such that $F(U) \subseteq V$ and the composite map $\psi \circ F \circ \varphi^{-1}$ is smooth from $\varphi(U)$ to $\psi(V)$ (**Fig.** A.4).

It is then possible to define a diffeomorphism between manifolds: if M and N are smooth manifolds, a ***diffeormophism from M to N*** is a smooth bijective map $F\colon M \to N$ that has a smooth inverse.

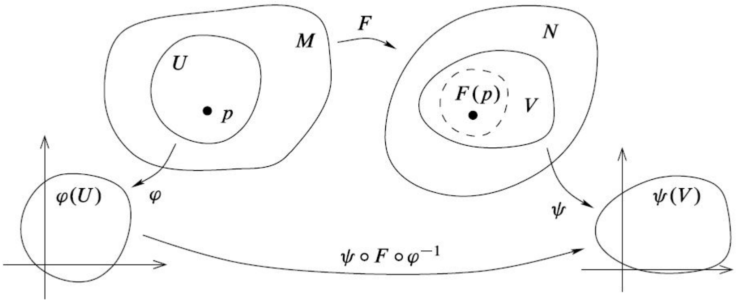

Fig. A.4: Definition of smooth map.

A.2 Tangent vectors and vector fields

A.2.1 Tangent spaces

One of the key concepts of Differential Geometry is the one of *tangent vectors.* We introduce it by giving a geometrical intuition of the subject. When we talk about elements of $\mathbb{R}^n$, we either think of them as *points*, whose only relevant property is location, or *vectors*, whose relevant properties are magnitude and direction, independently of location. In Differential Geometry, in order to define a *vector*, both location, i.e. the point of application, and the *sum* of magnitude and direction are important: a vector in $\mathbb{R}^n$ in a point a may be seen as belonging to a copy of $\mathbb{R}^n$ located in that precise point. A vector of the same magnitude and direction, applied in a different point of $\mathbb{R}^n$, is a different entity.

Moreover, the second relevant property of the concept of vector in Differential Geometry is that it acts as a *derivation*, that is, it satisfies the *product rule.* This is linked to the fact that a vector at a point can be bijectively associated to the directional derivative of a generic function at that point. For example, any vector $v_a \subseteq \mathbb{R}^n_a$ (space $\mathbb{R}^n$ at point a) yields a map $D_v|_a \colon C^\infty(\mathbb{R}^n) \to \mathbb{R}$, which takes the directional derivative in the direction v at a:

$$D_v|_a f = D_v f(a) = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t=0} f(a + tv). \tag{A.1}$$

We can thus define tangent vectors on manifolds. Let M be a smooth manifold and p be a point of M. A linear map $v \colon C^\infty(M) \to \mathbb{R}$ is called a ***derivation at p*** if it satisfies

$$v(fg) = f(p)vg + g(p)vf \tag{A.2}$$

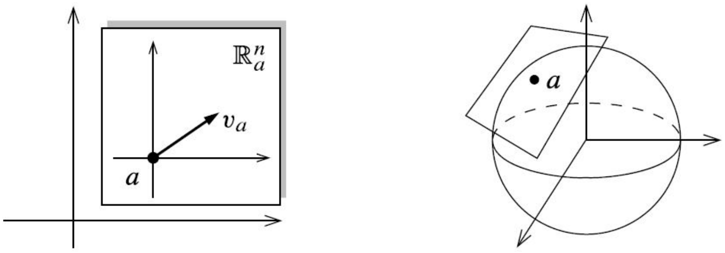

Fig. A.5: Geometric tangent space and tangent space to a sphere in $\mathbb{R}^3$.

for all $f, g \in C^\infty(M)$. The set of all derivations of C^∞ at p, denoted by T_pM, is a vector space called the ***tangent space to M at p***. An element of T_pM is called ***tangent vector at p***.

When talking about tangent spaces, it is important to understand how a smooth map affects tangent vectors. This is related to the action of a linear map between tangent spaces called ***differential*** of the function. If M and N are smooth manifolds and $F\colon M \to N$ is a smooth map, for each $p \in M$, we define a map

$$\mathrm{d}F_p\colon T_pM \to T_{F(p)}N \tag{A.3}$$

called the ***differential of F at p*** (**Fig.** A.6), as follows. Given $v \in T_pM$, we let $\mathrm{d}F_p(v)$ be the derivation at $F(p)$ that acts on $f \in C^\infty(N)$ by the rule

$$\mathrm{d}F_p(v)(f) = v(f \circ F) \tag{A.4}$$

It may be proved that this operator is linear and satisfies the product rule. In common words, the differential operator associates to a vector in the tangent space of a point of the domain, a vector in the tangent space of the corresponding point in the codomain. It may be proved that tangent spaces have the same dimensions

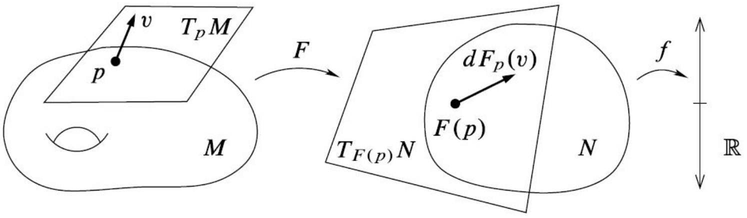

Fig. A.6: Differential on a manifold.

of the manifolds they are tangent to.

It is often useful to consider the set of all tangent vectors at all points of a manifold. Given a smooth manifold M, we define the ***tangent bundle of M***, denoted by TM, to be the disjoint union of the tangent spaces at all points of M:

$$TM = \coprod_{p \in M} T_p M. \tag{A.5}$$

A tangent bundle comes equipped with a natural ***projection map*** $\pi\colon TM \to M$, which sends each vector in T_pM to the point p at which it is tangent: $\pi(p, v) = p$. Just as an example, in the case $M = \mathbb{R}^n$, as the tangent space to each point may be identified with $\mathbb{R}^n$, the tangent bundle is the union of all these spaces and is thus the Cartesian product of $\mathbb{R}^n$ with itself. For any n-dimensional smooth manifold, its tangent bundle is a $2n$-dimensional smooth manifold.

In Differential Geometry it is also important to define velocity vectors to curves in a manifold. First of all, we define a curve: if M is a manifold, a ***curve in M*** is a continuous map $\gamma\colon J \to M$, where $J \subseteq \mathbb{R}$ is an interval. Practically, a curve is a map that associates a point on the manifold to each real number, which may be interpreted as a particular instant of time. At this point, given a smooth curve $\gamma\colon J \to M$ and $t_0 \in J$, the ***velocity of γ at t_0*** (**Fig.** A.7), denoted by $\gamma'(t_0)$, is the vector

$$\gamma'(t_0) = \mathrm{d}\gamma \left(\left. \frac{\mathrm{d}}{\mathrm{d}t} \right|_{t_0} \right) \in T_{\gamma(t_0)}M, \tag{A.6}$$

where $\mathrm{d}/\mathrm{d}t|_{t_0}$ is the standard coordinate basis vector in $T_{t_0}\mathbb{R}$. In other words, the velocity vector is the one associated by the differential of γ to the standard tangent vector in the space of real numbers, which is the regular derivation in the 'positive direction'. The latter belongs to the tangent space in t_0, while the former belongs to the tangent space in $\gamma(t_0)$.

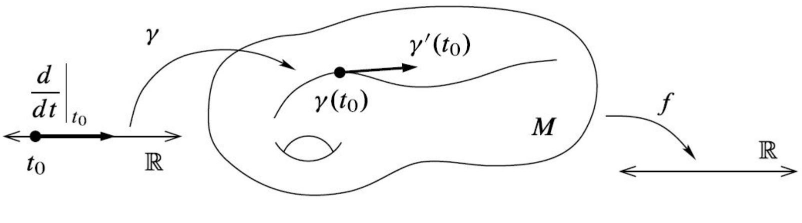

Fig. A.7: The velocity of a curve in a manifold.

A.2.2 Vector fields and flows

In Euclidean geometry, a vector field is an assignment of a vector to each point in the Euclidean space. The concept of vector field in Differential Geometry is quite similar: if M is a smooth manifold, a ***vector field on M*** is a continuous map $X\colon M \to TM$, usually written $p \mapsto X_p$, with the property that $X_p \in T_pM$ for each $p \in M$. A vector field on a manifold may be visualized in exactly the same way as in a Euclidean space: as an arrow attached to each point in M, belonging to the local tangent space and varying continuously from point to point (**Fig.** A.8). Vector fields may be also seen as endomorphisms in $C^\infty(M)$. ***Smooth vector fields*** are smooth maps from the manifold M to the manifold TM (the vector bundle).

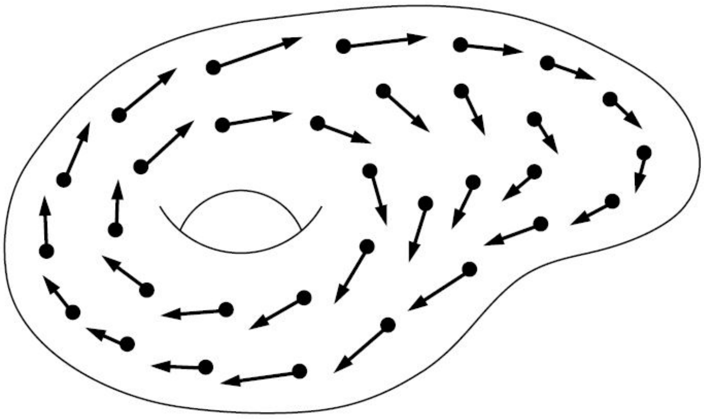

Fig. A.8: A vector field in a manifold.

Velocity vectors to a curve are vectors belonging to the tangent space at that point of the curve on the manifold. On the other hand, it may be possible to seek a curve whose tangent vector at each point is equal to a given vector: if V is a vector field on M, an ***integral curve of V*** is a differentiable curve $\gamma\colon J \to M$ whose velocity at each point is equal to the value of V at that point (**Fig.** A.9):

$$\gamma'(t) = V_{\gamma(t)} \qquad \forall t \in J. \tag{A.7}$$

If $0 \in J$, the point $\gamma(0)$ is called the ***starting point of γ***.

A very important concept based on integral curves is the one of ***flows***. A ***global flow*** on M is continuous map $\theta\colon \mathbb{R} \times M \to M$, satisfying the following

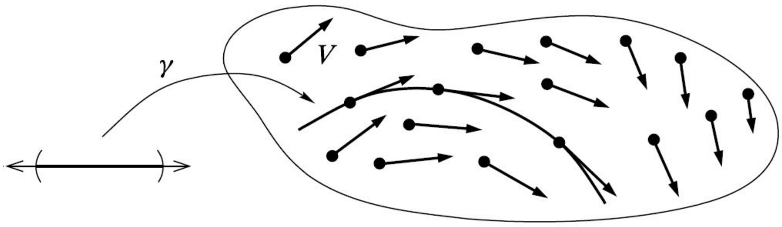

Fig. A.9: The integral curve of a vector field.

properties for all $s, t \in \mathbb{R}$ and $p \in M$:

$$\begin{cases} \theta(0,p) = p \\ \theta(t, \theta(s,p)) = \theta(t+s, p) \end{cases} . \tag{A.8}$$

This defines, for each $t \in \mathbb{R}$, a continuous map $\theta_t \colon M \to M$ by $\theta_t(p) = \theta(t,p)$ and, for each $p \in M$, a curve $\theta^{(p)} \colon \mathbb{R} \to M$ by $\theta^{(p)}(t) = \theta(t,p)$. More concretely, given an integral curve passing through a point, a flow is a map that assigns to each point $p \in M$ on the curve the point obtained by following for time t the integral curve starting at p: that is, the map θ_t *slides* the manifold along the integral curves for time t. Every smooth global flow gives rise to a smooth vector field whose integral curves are the curves defined by the flow; this vector field is called the ***infinitesimal generator of θ***. On the other hand, every smooth vector field is the infinitesimal generator of a smooth flow who is not necessarily global because it might not extend its domain to all $\mathbb{R} \times M$ (the difference between *global* and *local* is in the domain of definition).

A vector field might also depend on time: a ***time-dependent vector field on M*** is a continuous map $V \colon J \times M \to TM$, where $J \subseteq \mathbb{R}$ is an interval, such that $V(t,p) \in T_pM$ for each $(t,p) \in J \times M$. This means that for each instant of time $t \in J$, the map $V_t \colon M \to TM$ defined by $V_t(p) = V(t,p)$ is a vector field on M. If V is a time-dependent vector field on M, an ***integral curve of V*** is a differentiable curve $\gamma \colon J_0 \to M$, where J_0 is an interval contained in J, such that $\gamma'(t) = V(t, \gamma(t))$ $\forall t \in J_0$ (this means that *integral curves* are the same as *pathlines* in fluid dynamics, that is, curves that show the trajectory that an individual particle has followed).

An important way of combining two vector fields to obtain a third vector field is represented by *Lie brackets*. Given two vector fields X and Y, we define an operator $[X, Y] : C^\infty(M) \to C^\infty(M)$, called the ***Lie bracket of X and Y***, by

$$[X, Y]\, f = XYf - YXf. \tag{A.9}$$

The last concept introduced in this subsection is a very deep one, and related to the very rich theory of *Lie groups.* Here we limit ourselves to define a ***Lie algebra***, which is a vector space $\mathfrak{g}$ endowed with a map

$$\begin{array}{rccc} [,] & : & \mathfrak{g} \times \mathfrak{g} & \to & \mathfrak{g} \\ & & (X, Y) & \mapsto & [X, Y] \end{array}$$

satisfying the following properties $\forall X, Y, Z \in \mathfrak{g}$:

- *bilinearity*: for $a, b \in \mathbb{R}$,

$$\begin{aligned} [aX + bY, Z] &= a[X, Z] + b[Y, Z], \\ [X, aY + bZ] &= a[X, Y] + b[X, Z]; \end{aligned}$$

- *antisymmetry*:

$$[X, Y] = -[Y, X];$$

- *Jacobi identity*:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

A linear subspace $\mathfrak{h} \subseteq \mathfrak{g}$ is called a ***Lie subalgebra of*** $\mathfrak{g}$ if it closed under brackets.

Lie brackets satisfy precisely the axioms of a Lie algebra.

A.3 Vector bundles

It has already been said that the disjoint union of all the tangent spaces in the points of a manifold make up another smooth manifold structure, called the ***tangent bundle***. This concept may be generalized: a collection of vector spaces, one for each point of the manifold, might be *glued* together to make up a *vector bundle.* The rigorous definition is the following one:

Definition A.2. Let M be a topological space. A ***vector bundle of rank k over M*** is a topological space E together with a surjective continuous map $\pi \colon E \to M$ satisfying the following conditions:

- For each $p \in M$, the fibre $E_p = \pi^{-1}(p)$ over p is endowed with the structure of a k-dimensional real vector space.

- For each $p \in M$, there exists a neighbourhood U of p in M and a homeomorphism $\Phi\colon \pi^{-1}(U) \to U \times \mathbb{R}^k$, (called a ***local trivialization of E over U***), satisfying the following conditions (**Fig.** A.10):
  - $\pi_U \circ \Phi = \pi$ (where $\pi_U\colon U \times \mathbb{R}^k \to U$ is the projection);
  - for each $q \in U$, the restriction of Φ to E_q is a vector space isomorphism from E_q to $\{q\} \times \mathbb{R}^k \cong \mathbb{R}^k$.

The space E is usually called the ***total space of the bundle***, M is called its ***base***, and π is its ***projection***.

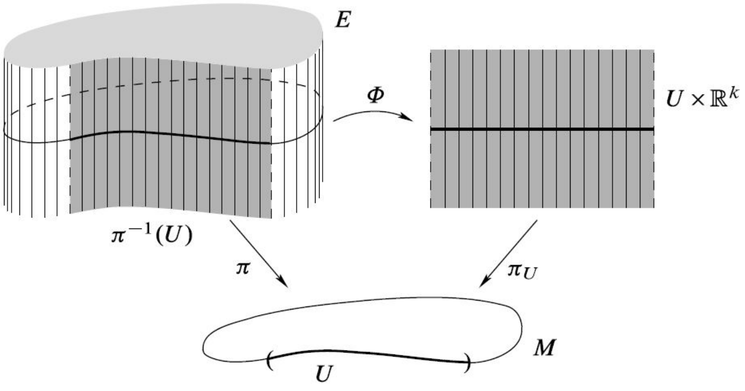

Fig. A.10: Local trivialization of a vector bundle.

Let $\pi\colon E \to M$ be a vector bundle. A ***section of E*** (sometimes called a ***cross section***) is a section of the map π, that is, a continuous map $\sigma\colon M \to E$ satisfying $\pi \circ \sigma = Id_M$. This means that $\sigma(p)$ is an element of the fibre E_p for each $p \in M$. There is a further distinction between ***local*** (**Fig.** A.11) and ***global sections*** according to whether the domain is a subset of the manifold or the whole manifold. We denote by $\Gamma(E)$ the vector space of all smooth global sections of E. Thus, vector fields are nothing but global sections of the tangent bundle; according to the usual notation, this section takes the special symbol $\mathfrak{X}(M) \overset{\text{def}}{=} \Gamma(TM)$.

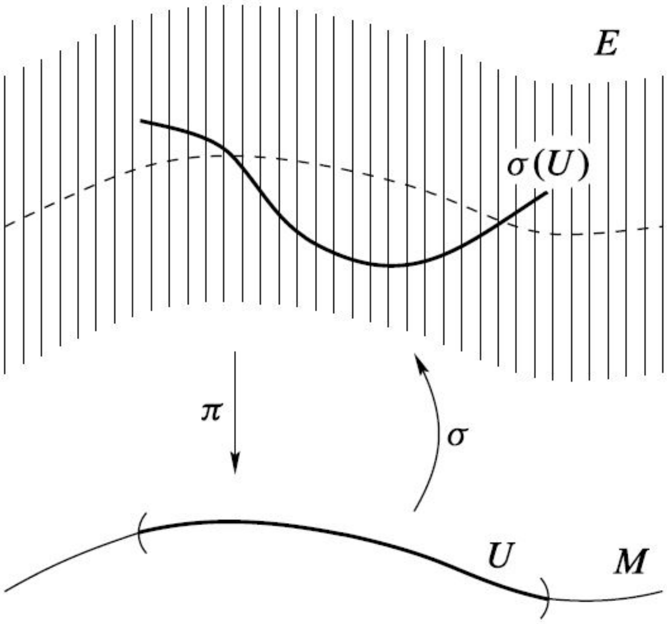

Fig. A.11: Local section of a vector bundle.

A.4 Covectors and tensors

A.4.1 Covectors

Another very important concept in Differential Geometry is the one of ***covector***. Let V be a finite-dimensional vector space. A ***covector on V*** is a real-valued linear functional on V, that is, a linear map $\omega\colon V \to \mathbb{R}$. The space of all covectors on V is itself a real vector space under the operations of pointwise addition and scalar multiplication. It is denoted by $\boldsymbol{V^*}$ and called the ***dual space of V***. It may be proved that the dimensions of a covector space are the same as those of the vector space it is associated to.

Very often there is confusion about the notation used to indicate vectors and covectors. We assume that basis covectors are written with upper indices and the components of a covector with lower indices. On the other hand, basis vectors are written with lower indices and components of a vector with higher indices. The notation will be illustrated and will be more clear in the successive **Subsec.** A.4.4.

Each map that links two vector spaces has a ***dual map*** between covector spaces. Let V and W be vector spaces and $A\colon V \to W$ is a linear map. We define a linear map $A^*\colon W^* \to V^*$, called the ***dual map*** or ***transpose of $\boldsymbol{A}$***, by

$$(A^*\omega)(v) = \omega(Av) \tag{A.10}$$

for $\omega \in W^*$, $v \in V$. It is also possible to define a ***second dual space*** $V^{**} = (V^*)^*$, which is the dual space of the dual space. For a finite-dimensional vector space $V^{**} = V$.

Once the concept of covector has been introduced, it is possible to define the ***cotangent space***. Given a smooth manifold M, for each $p \in M$, the ***cotangent space at p***, denoted by T_p^*M, is the dual space to T_pM:

$$T_p^*M \overset{\text{def}}{=} (T_pM)^*. \tag{A.11}$$

The elements of T_p^*M are called ***tangent covectors at p***, or just ***covectors at p***. For any smooth manifold M, the disjoint union

$$T^*M = \coprod_{p\in M} T_p^*M \tag{A.12}$$

is called the ***cotangent bundle of M***. The cotangent bundle is thus the collection of all cotangent spaces at the various points of a manifold. The cotangent bundle of smooth n-dimensional manifold has a natural smooth structure that make it into a smooth rank-n vector bundle over M. It is hard to visualize covector fields from a graphic point of view; as they are collections of linear functionals, the idea is that a covector ω_p at a certain point p may be identified by a hyperplane containing the endpoints of vectors v for which $\omega_p(v) = 1$. In **Fig.** A.12, two hyperplanes per each point are illustrated, where the second one is the hyperplane parallel to the previously mentioned one, passing through the origin. It represents the kernel of the covector.

If f is a smooth real-valued function on a smooth manifold M, it is possible to implicitly define the covector field $\mathrm{d}f$, called the ***differential of f***, as

$$\mathrm{d}f_p(v) = vf \tag{A.13}$$

for $v \in T_pM$.

The concept of differential allows the extension of some ideas commonly found in calculus to Differential Geometry. First of all, a smooth covector field ω on a smooth manifold M is said to be ***exact*** on M if there is a function $f \in C^\infty(M)$

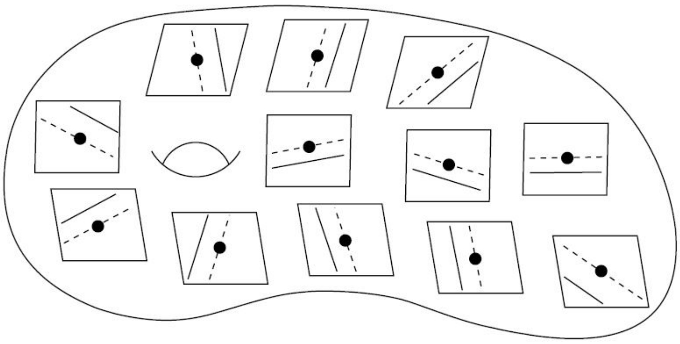

Fig. A.12: Graphical representation of a covector field.

such that $\omega = \mathrm{d}f$. The function f is called a ***potential for $\boldsymbol{\omega}$***.

Analogously to calculus, a smooth covector field ω is ***conservative*** if the line integral of ω over every piecewise smooth closed curve segment is zero. The definition of line integral on a manifold is here omitted for the sake of brevity. It may be proved that, on a smooth manifold, a smooth covector field is conservative if and only if it is exact. Also the idea of irrotationality might be extended to the field: in Differential Geometry, it is associated to the concept of ***closedness***. Without defining exactly a closed covector field (we shall come back to this point later), it is sufficient to say that the two concepts match and, as every conservative vector field in a Euclidean space is irrotational, in the same way every exact covector field on a manifold is closed.

As a smooth map yields a linear map between tangent spaces, in the same way, there is a sort of dual of this map, that is, a linear map between covector spaces that goes in the opposite direction.

Let $F\colon M \to M$ be a smooth map between smooth manifolds with or without boundary, and let $p \in M$ be arbitrary. The differential $\mathrm{d}F_p\colon T_pM \to T_F(p)N$ yields a dual linear map

$$\mathrm{d}F_p^*\colon T^*_{F(p)}N \to T^*_pM \tag{A.14}$$

called the ***(pointwise) pullback by $\boldsymbol{F}$ at $\boldsymbol{p}$***, or the ***cotangent map of $\boldsymbol{F}$***. Exploiting the definition, we may see that

$$\mathrm{d}F_p^*(\omega)(v) = \omega(\mathrm{d}F_p(v)) \tag{A.15}$$

with $\omega \in T^*_{F(p)}N$ and $v \in T_pM$.

It may be proved that covector fields are always *pulled back* to covector fields.

Given a smooth map $F\colon M \to N$, and a covector field ω on N, we define the covector field $F^*\omega$ on M, called the ***pullback of $\boldsymbol{\omega}$ by $\boldsymbol{F}$***, by

$$(F^*\omega)_p = \mathrm{d}F_p^*(\omega_{F(p)}) \tag{A.16}$$

It acts on a vector $v \in T_pM$ by

$$(F^*\omega)_p(v) = \omega_{F(p)}(\mathrm{d}F_p(v)). \tag{A.17}$$

A.4.2 Tensors

After illustrating *linear maps*, we generalize and illustrate *multilinear maps*, that is, maps that take several vectors as inputs and depend linearly on each one of them separately. ***Tensors*** in Differential Geometry are, in their simplest interpretation, *real-valued multilinear functions* of one or more variables. A covector is, for example, a tensor of rank one.

To be more precise, if $V_1, \ldots, V_k$, and W are vector spaces, a map

$$F\colon V_1 \times \cdots \times V_k \to W$$

is said to be ***multilinear*** if it is linear as a function of each variable separately when the others are held fixed. That is, for each i,

$$F(v_1, \ldots, av_i + a'v'_i, \ldots, v_k) = aF(v_1, \ldots, v_i, v_k) + a'F(v_1, \ldots, v'_i, \ldots, v_k).$$

The set $L(V_1, \ldots, V_k; W)$ of all multilinear maps from $V_1, \ldots, V_k$ to W is a vector space under the operations of pointwise addition and scalar multiplication. Two generic tensors give another tensor as a result of ***tensorial multiplication***: let $V_1, \ldots, V_k, W_1, \ldots, W_l$ be real vector spaces, and suppose $F \in L(V_1, \ldots, V_k; \mathbb{R})$ and $G \in L(W_1, \ldots, W_l; \mathbb{R})$. The function $F \otimes G\colon V_1 \times \cdots \times V_k \times W_1 \times \cdots \times W_l \to \mathbb{R}$ is defined as

$$F \otimes G(v_1, \ldots, v_k, w_1, \ldots, w_l) = F(v_1, \ldots, v_k)G(w_1, \ldots, w_l). \tag{A.18}$$

As F and G are both multilinear, also $F \otimes G$ depends linearly on each argument and is thus an element of the space $L(V_1, \ldots, V_k, W_1, \ldots, W_l; \mathbb{R})$, called ***tensor product of $\boldsymbol{F}$ and $\boldsymbol{G}$***.

A distinction can be made among tensors: they are defined as ***covariant*** or

contravariant according to the input that they take. More rigorously, we have the following definitions:

Definition A.3. Let V be a finite-dimensional vector space. If k is a positive integer, a ***covariant k-tensor on V*** is an element of the k-fold tensor product $V^* \otimes \cdots \otimes V^*$, which may be thought as a real-valued multilinear function of k elements of V. $\alpha\colon V \otimes \cdots \otimes V \to \mathbb{R}$ (with V replicated k times).

The number k is called the ***rank of α***. The vector space of all covariant k-tensors on V is denoted by $T^k(V^*) = V^* \otimes \cdots \otimes V^*$.

On the other hand, we have:

Definition A.4. Let V be a finite-dimensional vector space. If k is a positive integer, a ***contravariant tensor on V of rank k*** is an element of the k-fold tensor product $V \otimes \cdots \otimes V$, which may be thought as a real-valued multilinear function of k elements of V^*. $\alpha\colon V^* \otimes \cdots \otimes V^* \to \mathbb{R}$ (with V^* replicated k times).

The vector space of all contravariant k-tensors on V^* is denoted by $T^k(V) = V \otimes \cdots \otimes V$.

In general, when any two arguments of a tensor are exchanged, it is not possible to predict the effect on the outcome of the operation. However, there are two particular categories of tensors for which the exchange of any two arguments either does not change the value or changes its sign. A covariant k-tensor α on V is said to be ***symmetric*** if its value is unchanged by interchanging any pair of arguments. A covariant k-tensor α on V is said to be ***alternating*** (or ***antisymmetric*** or ***skew-symmetric***) if it changes sign whenever two of its arguments are interchanged. As it will be explained later, alternating tensor fields are called ***differential forms***.

It is also possible to define ***tensor bundles***. The ***bundle of covariant k-tensors on M*** is

$$T^k T^* M = \coprod_{p \in M} T^k(T_p^* M), \tag{A.19}$$

while the ***bundle of contravariant k-tensors on M*** is

$$T^k T M = \coprod_{p \in M} T^k(T_p M). \tag{A.20}$$

Sections of tensor bundles are called ***tensor fields***. A ***smooth tensor field*** is a smooth section of a tensor bundle; in other words, it is a smooth choice of a tensor for each point of an open set of the manifold. In particular, smooth covariant tensor fields are most important and deserve the shorthand notation

$$\mathfrak{T}^k(M) \overset{\text{def}}{=} \Gamma(T^kT^*M).$$

The pullback operation (A.16) is easily generalized to covariant tensors. Let $A \in \mathfrak{T}^k(N)$ be a covariant tensor field on N and $F\colon M \to N$ a smooth map. Then, the ***pullback of A by F*** is the covariant tensor field on M

$$(F^*A)_p = \mathrm{d}F_p^*(A_{F(p)}), \tag{A.21}$$

which acts on vectors $v_1, \ldots, v_k \in T_pM$ by

$$(F^*A)_p(v_1, \ldots, v_k) = A_{F(p)}(\mathrm{d}F_p(v_1), \ldots, \mathrm{d}F_p(v_k)). \tag{A.22}$$

For contravariant tensors, the definition works, with slight modifications, only if F is a diffeomorphism.

A.4.3 Lie derivative

A canonical way of computing directional derivatives of a tensor field on a manifold passes through the concept of *Lie derivative.* Since there is no canonical way of comparing tensors in different points, the idea is to 'slide' a tensor field along the flow of a vector field, as shown in **Fig.** A.13.

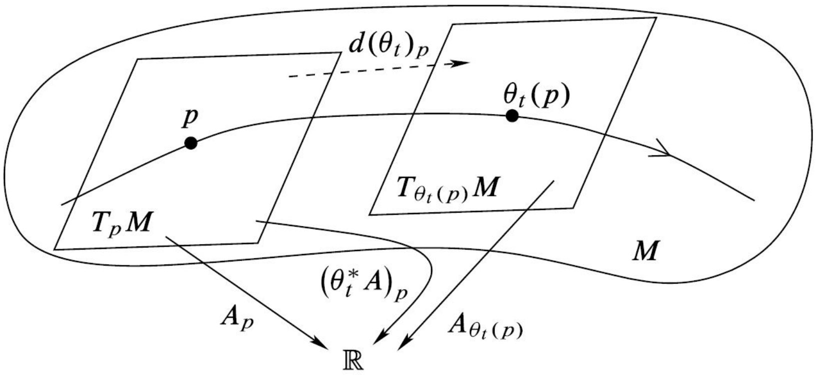

Fig. A.13: The Lie derivative of a tensor field.

The ***Lie derivative of the covariant tensor field A with respect to V*** is thus the following tensor field:

$$\left(\pounds_V A\right)_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} (\theta_t^* A)_p = \lim_{t\to 0} \frac{\mathrm{d}(\theta_t)_p^* \left(A_{\theta_t(p)}\right) - A_p}{t} \tag{A.23}$$

The very same definition of Lie derivative as in **Eq.** (A.23) applies to vector fields and, more generally, to contravariant tensor fields (with the condition to correctly interpret the various terms).

For vector fields, we have

$$\begin{aligned}\left(\pounds_V W\right)_p = \left.\frac{\mathrm{d}}{\mathrm{d}t}\right|_{t_0} (\theta_t^* W)_p = \lim_{t\to 0} \frac{\mathrm{d}(\theta_t)_p^* \left(W_{\theta_t(p)}\right) - W_p}{t} = \\ = \lim_{t\to 0} \frac{\mathrm{d}(\theta_{-t})_{\theta_t(p)} \left(W_{\theta_t(p)}\right) - W_p}{t},\end{aligned} \tag{A.24}$$

which is illustrated in **Fig.** A.14.

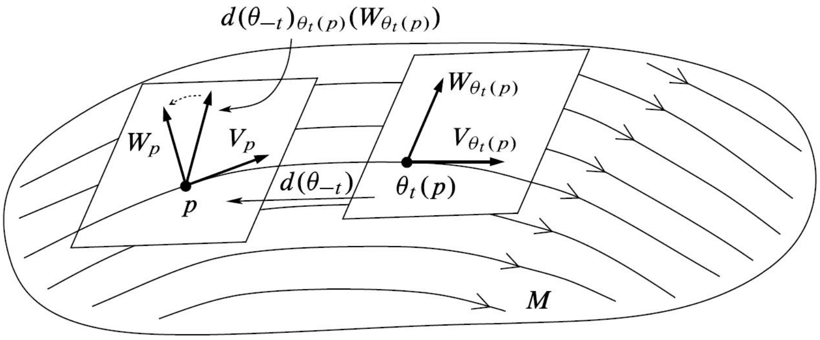

Fig. A.14: The Lie derivative of a tensor field.

Moreover, one could verify that

$$\pounds_X Y = [X, Y]\,. \tag{A.25}$$

A further important property is Leibniz rule with respect to contraction (we take a contravariant tensor field as example, because we used this case in Chapter 4):

$$\begin{aligned}&\pounds_X(A(\omega_1, \omega_2, \ldots, \omega_k)) = (\pounds_X A)(\omega_1, \omega_2, \ldots, \omega_k) + \\ &+ A(\pounds_X \omega_1, \omega_2, \ldots, \omega_k) + A(\omega_1, \pounds_X \omega_2, \ldots, \omega_k) + A(\omega_1, \omega_2, \ldots, \pounds_X \omega_k).\end{aligned} \tag{A.26}$$

A.4.4 Riemannian manifolds

In Euclidean spaces, geometric concepts such as distances and angles are defined thanks to the inner product. In Differential Geometry, in order to accomplish a similar feature, the appropriate structure is a ***Riemannian metric***. The subject of Riemannian Geometry is very deep and was illustrated by Lee [L97]. In the following paragraphs, we will only give some basic definitions.

In Euclidean geometry an inner product is a symmetric tensor as it has two vectors as arguments and gives back a real number and is invariant upon an exchange of arguments. In much the same way, if M is a smooth manifold, a ***Riemannian metric on M*** is a smooth symmetric covariant 2-tensor field on M that is positive definite at each point. A ***Riemannian manifold*** is a pair (M, g), where M is a smooth manifold and g is a Riemannian metric on M. If g is a Riemannian metric on M, then for each $p \in M$, the 2-tensor g_p is an inner product on T_pM. Because of this, the notation $\langle v, w\rangle_g$ is used to denote the real number $g_p(v, w)$ for $v, w \in T_pM$. One of the simplest examples of a Riemannian metric is the *Euclidean metric* on $\mathbb{R}^n$, that is, the operator that associates to any pair of vectors the sum of the products of the corresponding components of the two vectors.

Metrics that have significantly different geometric properties may be defined on the same manifold. Thus, the choice of a metric is a particularly delicate one. Different geometric concepts may be defined on a Riemannian manifold:

- the ***length*** or ***norm*** of a tangent vector $v \in T_pM$ is defined to be $||v||_g = \sqrt{\langle v, v\rangle_g} = \sqrt{g_p(v, v)}$;
- the ***angle*** between two nonzero tangent vectors $v, w \in T_pM$ is the unique $\theta \in [0, \pi]$ satifying $\cos\theta = \dfrac{\langle v, w\rangle_g}{||v||_g ||w||_g}$;
- tangent vectors $v, w \in T_pM$ are said to be ***orthogonal*** if $\langle v, w\rangle_g = 0$.

One of the most important aspects of a Riemannian metric is that it provides a natural isomorphism, that is a one-to-one correspondence, between the tangent and the cotangent bundles. Given a Riemannian metric g on a smooth manifold M, a *vector bundle homomorphism* (sometimes we say *vector bundle map*) $g^\flat : TM \to T^*M$ is defined as follows. For each $p \in M$ and each $v \in T_pM$, we let $g^\flat_p(v) \in T^*_pM$ be the covector defined by

$$g^\flat_p(v)(w) = g_p(v, w) \qquad \forall w \in T_pM \tag{A.27}$$

The components of the covector field are usually written as:

$$g^\flat(X) = X_j dx^j, \qquad X_j = g_{ij}X^i. \tag{A.28}$$

Because of this, one says that $g^\flat(X)$ is obtained from X by ***lowering an index***. Thus, the symbol $\flat$ is used because, in musical notation, it indicates that a tone is to be lowered. As $[g_{ij}]$ is the matrix of the isomorphism $g^\flat$, it is invertible and its inverse is the matrix of the inverse map $g^\sharp \colon T_p^*M \to T_pM$, indicated as $[g^{ij}]$. The inverse matrix is symmetric, too, and it is said that $g^\sharp$ ***raises an index***; because of this, the symbol $\sharp$ is used, as, in musical notation, it indicates that a tone is to be raised. These two isomorphisms are thus called ***musical isomorphisms***.

Thanks to the *sharp* operation, it is possible to reinstate the ***gradient*** as a vector field on Riemannian manifolds:

$$(\operatorname{grad} f)_p = g_p^\sharp \mathrm{d}f_p. \tag{A.29}$$

Thus, the gradient is the unique vector field that satisfies

$$\langle \operatorname{grad} f, X\rangle_g = Xf \qquad \forall X. \tag{A.30}$$

It may thus be seen that the definition of a gradient in Differential Geometry is not independent from the choice of metric: it is upon the choice of a particular type of metric that the gradient is determined.

A.4.5 Differential forms

As it has been stated when introducing tensors, *symmetric* and *alternating* tensors are tensors that behave specifically under the exchange of any two arguments. The importance of a particular example of symmetric tensor fields, that is Riemannian metrics, has been stated; however, in Differential Geometry, as surprising as it might be, the study and the use of alternating tensors turns out to be more useful than the ones of symmetric tensors. Particularly important is the idea of ***differential forms***. The algebra that results after defining the *wedge product*, the *exterior differential* and the *interior product* gives a powerful generalization to any dimension of the familiar concepts of the differential operators gradient, divergence and curl, of the cross product, and allows to develop the theory of orientation and integration on manifolds.

Given a smooth manifold M, T^kT^*M is the bundle of covariant k-tensors on the manifold. The subset of the bundle consisting of alternating tensors is denoted by Λ^kT^*M:

$$\Lambda^kT^*M = \coprod_{p\in M} \Lambda^k(T_p^*M). \tag{A.31}$$

A section of Λ^kT^*M is called a ***differential k-form***, or just a ***k-form***; it is thus a continuous tensor field whose value at each point is an alternating tensor. The integer k is called the ***degree*** of the form. Smooth k-forms make up a vector space indicated by $\Omega^k(M) \overset{\text{def}}{=} \Gamma(\Lambda^kT^*M)$.

There is a notion of product between differential forms that turns their space into a very useful algebra. Let's outline the main facts.

Given a k-tensor α, we define the ***alternation operator*** $\operatorname{Alt}\colon T^k(V^*) \to \Lambda(V^*)$ as follows:

$$(\operatorname{Alt}\alpha)(v_1, v_2, \dots, v_k) = \frac{1}{k!}\sum_{\sigma\in S_k}(\operatorname{sgn}\sigma)\alpha(v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(k)}),$$

where S_k is the symmetric group on k elements (the group of all permutations of the set $\{1, \dots, k\}$), $\operatorname{sgn}\sigma = 1$ if σ is *even* (i.e., can be written as a composition of an *even* number of transpositions), and $\operatorname{sgn}\sigma = -1$ otherwise. This means that $\operatorname{Alt}\alpha$ is a k-form.

At this point, we can introduce the ***wedge product*** (or ***exterior product***). Given $\alpha \in \Lambda^k(V^*)$ and $\beta \in \Lambda^\ell(V^*)$, their product is the following $(k+\ell)$-form:

$$\alpha \wedge \beta = \frac{(k+\ell)!}{k!\ell!}\operatorname{Alt}(\alpha\otimes\beta). \tag{A.32}$$

The coefficient is justified by the simplicity of certain calculations which we do not discuss here. For instance, in the case of two 1-forms, **Eq.** (A.32) reads:

$$\omega\wedge\eta = 2\cdot\frac{1}{2}(\omega\otimes\eta - \eta\otimes\omega) = \omega\otimes\eta - \eta\otimes\omega.$$

The wedge product is *bilinear*, *associative* and *anticommutative*. This means that, if we define the vector space

$$\Lambda(V^*) = \bigoplus_{k=0}^{n}\Lambda^k(V^*)$$

the wedge product turns $\Lambda(V^*)$ into an *associative and anticommutative graded* *algebra*, called ***exterior algebra*** (or ***Grassmann algebra***) ***of*** $\boldsymbol{V}$.

An ***exterior derivative*** is a differential operator on smooth forms that constitutes a generalization of the concept of differential of a function. Without entering into definitions and mathematical details, the exterior derivative d of a differential form of degree k is a differential form of degree $k+1$ that satisfies certain properties that are here omitted for the sake of brevity. In the case of a function f, that is, a zero-form, the exterior derivative is its differential $\mathrm{d}f$. It is said that a differential form $\omega \in \Omega^k(M)$ is ***closed*** if $\mathrm{d}\omega = 0$ and ***exact*** if there exists a smooth $(k-1)$-form η on M such that $\omega = \mathrm{d}\eta$. Every exact form is automatically closed for the property of the exterior differential $\mathrm{d} \circ \mathrm{d} = 0$. Vice versa, a closed form is exact if it is defined over a simply connected domain; hence, a closed form is always *locally exact*. This is equivalent to the familiar Poincaré lemma, which says that every irrotational vector field is conservative in a simply connected domain.

On the other hand, it is possible to define the ***interior product***, which is an important operation that relates vectors with alternating tensors. If V is a finite-dimensional vector space, for each $v \in V$, we define a linear map $\iota_v \colon \Lambda^k(V^*) \to \Lambda^{k-1}(V^*)$, called ***interior product by*** $\boldsymbol{v}$, as follows:

$$\iota_v\omega(w_1, w_2, \ldots, w_{k-1}) = \omega(v, w_1, w_2, \ldots, w_{k-1}) \tag{A.33}$$

In common words, $\iota_v\omega$ is obtained from ω by inserting v into the first slot. This is also written as

$$v \lrcorner\, \omega = \iota_v\omega. \tag{A.34}$$

Finally, we report a very useful theorem to calculate the Lie derivative of differential form.

Theorem A.5 (***Cartan's Magic Formula***)**.** *Let* V *a smooth vector field and* ω *a smooth differential form on a smooth manifold. Then,*

$$\pounds_V\, \omega = V \lrcorner\, \mathrm{d}\omega + \mathrm{d}(V \lrcorner\, \omega). \tag{A.35}$$

The most important consequence is the

Corollary A.6. *If* V *is a smooth vector field and* ω *is a smooth differential form, then the Lie derivative and exterior differential operators commute:*

$$\pounds_V(\mathrm{d}\omega) = \mathrm{d}(\pounds_V\, \omega). \tag{A.36}$$

A.5 (Regular) distributions and foliations

For the discussion we would need to say too many things of topology and submanifolds. Since we don't have the space to that, we try to give only notions as intuitive as possible.

A ***(regular) distribution on M of rank k*** is a rank-k subbundle of TM, that is, a vector bundle $\pi_D\colon D \to M$, where

- D is a topological subspace of TM;
- π_D is the restriction of $\pi_{TM}\colon TM \to M$;
- for each $p \in M$, the subset $D_p = D \cap T_pM$ is a linear subspace (and inherits the vector space structure) of T_pM.

Vectors in D are called ***horizontal***.

A distribution D is a ***smooth distribution*** if each point $p \in M$ has a neighbourhood U on which there are smooth vector fields $X_1, \dots, X_k\colon U \to TM$ such that $X_1|_p, \dots, X_k|_p$ form a basis for D_p at each $p \in U$. In this case, we say that D is the distribution ***(locally) spanned by the vector fields $\boldsymbol{X_1, \dots, X_k}$***. We write $X_1, \dots, X_k \in \mathfrak{X}^D_{\mathrm{loc}}(M)$.

Suppose $D \subseteq TM$ is ***smooth (regular) distribution***. A nonempty immersed submanifold $N \subseteq M$ is called an ***integral submanifold of D*** if $T_pN = D_p$ for each p in N. A smooth (regular) distribution D on M is said to be ***integrable*** if each point of M is contained in an integral submanifold of D.

Theorem A.7 (**Frobenius Theorem**). *Let D a smooth (regular) distribution on M and $\Gamma(D) \subseteq \mathfrak{X}(M)$ denote the space of smooth global sections of D. The following statements are equivalent:*

- *D is integrable;*
- *$\Gamma(D)$ is a Lie subalgebra of $\mathfrak{X}(M)$ (*involutivity *condition).*

An integral submanifold of D is called *maximal* if it is not contained in any strictly larger integral submanifold of D. For a smooth integrable distribution, every point is contained in a unique maximal integral submanifold. The collection of all maximal integral submanifolds form a ***foliation***[1] ***of M induced by the***

---

[1] A foliation is defined independently from a distribution; however, we are not interested in its definition, but only in the case where its leaves are the maximal integral submanifolds of a distribution.

integrable distribution $\boldsymbol{D}$, and each maximal integral manifold is called a ***leaf of the foliation***.

References

[L13] J. M. Lee. *Introduction to Smooth Manifolds*. 2nd ed. (1st ed., 2003). New York: Springer, 2013 (cit. on p. 161).

[L97] J. M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. 1st ed. New York: Springer, 1997 (cit. on p. 180).

B

Helmholtz-Hodge decomposition

Helmholtz decomposition for vector fields in $\mathbb{R}^3$ was first developed by Hermann L. F. von Helmholtz in [vH58], and it is based on using two potentials to extract the irrotational and divergenceless components from a vector field defined in all three-dimensional space. The same decomposition often takes the name of *Helmholtz-Hodge decomposition* when it is generalized (see [BNPB13] for an interesting survey). For bounded domains, either

- by specifying the tangential component of the irrotational part or the normal component of the divergenceless one,
- or by adding a third harmonic part and specifying both tangential and normal components of the other two parts at the same time.

The two or three components are assured to be unique because they are mutually $L^2(\mathbb{R}^3)$-orthogonal.

A further generalization can be built on a compact, boundaryless, oriented, Riemannian manifold, and reads

$$\begin{aligned} \omega &= \omega_\alpha + \omega_\beta + \omega_\gamma \\ &= \mathrm{d}\alpha + \mathrm{d}^*\beta + \gamma, \end{aligned} \tag{B.1}$$

where $\alpha \in \Omega^{k-1}(M)$ (i.e., ω_α is the *exact part* of ω), $\beta \in \Omega^{k+1}(M)$ (ω_β is the *co-exact part* of ω) and $\gamma \in H^k(M)$. d is the exterior derivative, d^* its adjoint with respect to the Riemannian metric, and $H^k(M)$ is the space of harmonic forms on M. Harmonic forms on a Riemannian manifold are defined through the Laplace-Beltrami operator $\Delta = \mathrm{d}\mathrm{d}^* + \mathrm{d}^*\mathrm{d}$, which is a generalization of the Laplacian for a p-form. In terms of vector spaces, decomposition (B.1) is

$$\Omega(M) = \mathrm{d}\Omega^{k-1}(M) \oplus \mathrm{d}^*\Omega^{k+1}(M) \oplus H^k(M).$$

This may also be generalized to the case of manifold with boundary.

In the context of **Subsec.** 3.3.4, the ambient space is $\mathbb{R}^n$, but we are interested only in the fact that any p-form has always an *exact* component; the remaining component is left unspecified. The decomposition is thus used in the form

$$\omega = \mathrm{d}\alpha + y, \tag{B.2}$$

where $y \in \left[\mathrm{d}\Omega^{k-1}(\mathbb{R}^n)\right]^\perp = \mathrm{d}^*\Omega^{k+1}(\mathbb{R}^n) \oplus H^k(\mathbb{R}^n)$.

If ω is a 1-form (a covector), $\alpha \in \Omega^0(\mathbb{R}^n) = C^\infty(\mathbb{R}^n)$ is a smooth function, and the same decomposition may also be expressed in terms of vectors through the tangent-cotangent isomorphism $g^\sharp$:

$$\boldsymbol{j} = g^\sharp\omega = g^\sharp(\mathrm{d}\alpha + y) = \operatorname{grad}\alpha + \boldsymbol{y}.$$

We denote by $\mathfrak{X}_{\mathrm{cons}}(M) = \mathrm{d}C^\infty(M)$ the space of *conservative* vector field and by $[\mathfrak{X}_{\mathrm{cons}}(M)]^\perp$ its orthogonal complement.

References

[BNPB13] H. Bhatia, G. Norgard, V. Pascucci, and P.-T. Bremer. “The Helmholtz-Hodge Decomposition–A Survey”. In: *IEEE Transactions on Visualization and Computer Graphics* 19.8 (2013), pp. 1386–1404 (cit. on p. 187).

[vH58] H. L. F. von Helmholtz. “Über Integrale der hydrodynamischen Gleichungen, welche den Wirbelbewegungen entsprechen.” In: *Journal für die reine und angewandte Mathematik* 55 (1858), pp. 25–55 (cit. on p. 187).

Index