



HAL
open science

Inférence des propriétés physiques de régions d'ionisation stellaires au moyen de glitches sismiques

Pierre Houdayer

► **To cite this version:**

Pierre Houdayer. Inférence des propriétés physiques de régions d'ionisation stellaires au moyen de glitches sismiques. Astrophysique [astro-ph]. Université Paris Cité, 2022. Français. NNT : 2022UNIP7285 . tel-04473876

HAL Id: tel-04473876

<https://theses.hal.science/tel-04473876>

Submitted on 22 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS CITÉ

ECOLE DOCTORALE ASTRONOMIE ET ASTROPHYSIQUE D'ÎLE DE FRANCE (ED 127)

THÈSE DE DOCTORAT D'ASTROPHYSIQUE

Inférence des propriétés physiques de régions d'ionisation stellaires au moyen de glitches sismiques

par

Pierre HOUDAYER

et dirigée par

Marie-Jo GOUPIL & Daniel REESE

en vue de l'obtention du titre de DOCTEUR EN ASTROPHYSIQUE

au

LABORATOIRE D'ÉTUDES SPATIALES ET D'INSTRUMENTATION EN ASTROPHYSIQUE (LESIA)

Présentée et soutenue publiquement le

19 Octobre 2022

devant un jury composé de

Marie-Christine ANGININ

Achim WEISS

Orlagh CREEVEY

Sébastien DEHEUVELS

Valérie VAN GROOTEL

Andrea MIGLIO

Marie-Jo GOUPIL

Daniel REESE

Pr. des Universités, Observatoire de Paris

Professor, Max-Planck-Institut, Garching

Astronome, Observatoire de la Côte d'Azur

Maître de conférence, Université de Toulouse III

Chercheuse au FRS-FNRS, Université de Liège

Professor, Università di Bologna

Astronome, Observatoire de Paris

Astronome-adj., Observatoire de Paris

PRÉSIDENTE DU JURY

RAPPORTEUR

RAPPORTRICE

EXAMINATEUR

EXAMINATRICE

EXAMINATEUR

DIRECTRICE DE THÈSE

CO-DIRECTEUR DE THÈSE

« L'homme n'est pas un cercle à un seul centre ; c'est une ellipse à deux foyers. Les faits sont l'un, les idées sont l'autre. »

VICTOR HUGO

« — N'oublie jamais, celui qui croit savoir n'apprend plus. »

PIERRE BOTTERO

Résumé

INFÉRENCE DES PROPRIÉTÉS PHYSIQUES DE RÉGIONS D'IONISATION STELLAIRES AU MOYEN DE GLITCHS SISMIQUES

Si l'inférence sismique a de maintes fois démontré sa plus-value en ce qui concerne la détermination des paramètres stellaires, leur estimation peut être sujette à des biais et ainsi demeurer incertaine. Un exemple frappant est celui de la dégénérescence masse-hélium, où les nombreuses possibilités pour modéliser la physique interne causent à la fois une mauvaise détermination de la masse de l'étoile et de son contenu en hélium. Dans ce contexte, une estimation indépendante de l'abondance d'hélium est d'un double intérêt : en plus de réduire considérablement notre incertitude à propos de la masse de l'étoile, elle pourrait nous apporter de fortes contraintes sur sa physique interne. À cet égard, une approche prometteuse consiste à exploiter le glitch de l'ionisation, une déviation des fréquences d'oscillation due aux variations locales de structure rencontrées dans la région d'ionisation. Cependant, de nombreuses méthodes reposent aujourd'hui sur des modèles de perturbation ad-hoc et proposent une détermination de l'abondance d'hélium après calibration sur d'autres modèles, ce qui lui ôte son caractère indépendant.

Dans ce manuscrit, on propose une approche alternative pour étudier le glitch de l'ionisation en dérivant un modèle cohérent de la région d'ionisation. On s'appuie sur une approximation analytique du premier exposant adiabatique dans une zone convective avant d'obtenir une structure complète de la région à l'équilibre hydrostatique. Cette structure se trouve alors dépendre de trois propriétés physiques fondamentales que sont l'abondance d'hélium en surface, la dégénérescence électronique dans la zone convective et l'étendue de la région d'ionisation. Une caractérisation du glitch de l'ionisation fondée sur un tel modèle nous permettra non seulement d'estimer l'abondance d'hélium sans recours à la calibration, mais également d'exploiter les contraintes sur la physique interne fournies par ces propriétés additionnelles. On propose ainsi un modèle du glitch faisant directement intervenir ces grandeurs et l'on évalue ses capacités d'inférence sur des données synthétiques. L'information retrouvée sur les propriétés physiques est importante, et l'on montre comment adapter la méthode pour gérer une contamination du glitch par des effets de surface en incluant un processus gaussien dans le modèle.

MOTS-CLÉS : *Astérosismologie; Physique stellaire; Glitches sismiques; Théorie; Modélisation; Inférence bayésienne*

Abstract

INFERRING THE PHYSICAL PROPERTIES OF STELLAR IONISATION REGIONS THROUGH SEISMIC GLITCHES

Although seismic inference has proven to be valuable on many occasions when determining stellar parameters of solar-like pulsators, estimating these very parameters can be subject to biases and thus remain uncertain. A particularly important example is the helium-mass degeneracy, where the uncertainties on the internal physics cause a poor determination of both the mass and surface helium content. In this context, an independent estimate of the helium abundance is beneficial in two ways : in addition to considerably reducing the uncertainty on the mass of the star, it may also lead to strong constraints on its internal physics. A promising way to obtain such an estimate is to exploit the so-called ionisation glitch, a deviation from the asymptotic oscillation frequency pattern caused by the rapid structural variation within the ionisation zone. However, although becoming more and more sophisticated, current approaches for modelling glitches face problems such as the need for calibration by realistic stellar models.

Accordingly, we focused on an alternative approach to studying glitches using a physically motivated model of the ionisation region. An analytical approximation of the first adiabatic exponent in a convective zone is derived before obtaining a complete structure of the region at hydrostatic equilibrium. This structure is then found to depend on three fundamental physical properties, namely the surface helium abundance, the electronic degeneracy in the convective zone and the extent of the ionisation region. A characterisation of the ionisation glitch based on such a model will not only allow us to estimate the helium abundance without resorting to calibration, but also to exploit the constraints on the internal physics provided by these additional properties. We therefore propose a glitch model directly involving these quantities and evaluate its inference capabilities on synthetic data. A substantial amount of information on the physical properties may be recovered. Furthermore, we show how to adapt the method to handle a contamination of the glitch by surface effects by including a Gaussian process in the model.

KEYWORDS : *Asteroseismology; Stellar physics; Seismic glitches; Theory; Modelling; Bayesian inference*

Remerciements

Les pages qui suivent concluent trois années de travail mais aussi un chapitre de ma vie. Si j'en suis l'auteur, elles ne pourraient être sans le concours de nombreuses personnes, par des échanges précieux, leur soutien ou leur présence à mes côtés sur cette piste parfois sinueuse que peut être la thèse. Par ces quelques lignes, j'aimerais ainsi remercier celles et ceux ayant permis l'écriture de ce manuscrit et, de façon plus générale, la réalisation de cette thèse.

Mes premières pensées se tournent vers les personnes à qui je dois la supervision et sans doute l'aboutissement de ce travail de thèse. Je remercie profondément Marie-Jo de m'avoir accepté en thèse sur un sujet qui m'est cher et dont j'ai découvert les prémices il y a maintenant plus de 5 ans. Nos échanges, animés par une soif de compréhension partagée, autant que ses conseils m'auront guidé du début à la fin de cette thèse. Elle aura su m'apporter à la fois soutien et recul lorsqu'il le fallait et je me considère chanceux d'avoir pu apprendre auprès d'elle. J'aimerais ensuite remercier tout particulièrement Daniel pour son investissement et l'enthousiasme qu'il aura su insuffler à mon travail. Nos nombreuses discussions (digressions ?) sur des thèmes aussi nombreux que variés auront énormément contribué à l'accomplissement de cette thèse et les moments passés au bureau 106 me laisseront (ainsi qu'au tableau velleda) un souvenir indélébile. Je le remercie également pour son admirable vaillance, qu'il s'agisse de relire un anglais parfois hésitant ou de vérifier la quantité ahurissante d'équations ponctuant mes textes. Je remercie aussi vivement Yveline, à qui je dois la découverte du monde de la recherche et dont la disponibilité et la bienveillance m'apportent beaucoup encore aujourd'hui.

Je tiens également à remercier les membres du LESIA dont j'ai eu la chance de partager le quotidien ces trois dernières années et qui savent amener détente et légèreté lorsqu'il le faut : Kévin, Jordan, Guillaume, Mina, Fernando, David, Ethan en plus de ceux qui m'accompagnent depuis mon premier stage dans ce laboratoire : Réza, Caroline, Benoît, Kévin, Rhita, Coralie, Claude, Eric, Richard ainsi que les stagiaires et étudiants que j'ai pu côtoyer. Ils auront contribué au plaisir quotidien que j'éprouve à travailler ici et je les remercie pour leur soutien et leurs conseils.

J'aimerais également exprimer ma reconnaissance vis-à-vis du jury pour avoir consenti à la lecture du manuscrit et les remercier pour leur investissement dans cette tâche.

J'aimerais adresser mes derniers remerciements à ma famille et plus particulièrement à ma mère Cathy pour son soutien indéfectible, sa tendresse et ses encouragements depuis que je me suis engagé dans cette voie ainsi qu'à mon frère Guillaume pour tous les bons moments que l'on partage et qui valent tant. Je remercie aussi tendrement mes grand-parents, Arlette et Claude, à la fois pour leur affection et pour m'offrir tous les ans cette parenthèse reposante dont je ne saurai me passer. Je remercie enfin Naïs, avec qui le présent s'illumine et le futur n'effraie plus.

Table des matières

Résumé	iii
Abstract	iv
Remerciements	v
Table des matières	vii
Table des figures	xi
Liste des tableaux	xiii
I Introduction	1
1 Introduction	3
1.1 Prologue	3
1.2 L'astérosismologie	4
1.3 La dégénérescence masse-hélium	8
1.4 Levée de la dégénérescence	11
1.5 La région d'ionisation et le glitch sismique	12
1.6 Objectif et contenu du manuscrit	13
1.7 Purpose and content of the manuscript [en]	14
2 Théorie	15
2.1 * Préambule	16
2.1.1 Le fluide	16
2.1.2 La densité de probabilité à N particules	17
2.1.3 Grandeur physique	18
2.1.4 Le développement en <i>cluster</i>	18
2.1.5 La densité de probabilité à 1 particule	18
2.1.6 Forme explicite de l'équation de Liouville	20
2.1.7 La hiérarchie BBGKY	20
2.1.8 L'Équation de Boltzmann	21
2.1.9 Invariant collisionnel	21
2.1.10 Le théorème H	22
2.1.11 La distribution localement maxwellienne	22
2.1.12 La limite maxwellienne	24
2.1.13 Développement de Chapman-Enskog	24
2.1.14 L'équilibre thermodynamique local	25
2.1.15 Bilan : les différentes échelles	26
2.2 * De la théorie cinétique aux équations dynamiques	27
2.2.1 Évolution d'une grandeur moyenne	28
2.2.2 Évolution d'un invariant collisionnel	29
2.2.3 L'équation de continuité	29
2.2.4 L'évolution de la quantité de mouvement	29

2.2.5	Formes conservative et convective, formalismes eulérien et lagrangien	31
2.2.6	La conservation de l'énergie	32
2.2.7	Le transfert radiatif	32
2.2.8	Fermeture des équations dynamiques	35
2.2.9	Un exemple de fermeture : l'équilibre thermodynamique local	36
2.2.10	La fermeture de Navier-Stokes	41
2.3	Des équations dynamiques à la structure stellaire	43
2.3.1	Formes stationnaire et statique des équations fluides	43
2.3.2	Un modèle simple de structure à l'équilibre statique	44
2.3.3	Les étoiles	45
2.3.4	Composition et diffusion	46
2.3.5	Stabilité du système d'équations	49
2.3.6	Les régions stellaires	53
2.3.7	L'ionisation	56
2.3.8	La modélisation stellaire	59
2.3.9	Conditions aux limites	60
2.4	Les perturbations stationnaires et l'astérosismologie	63
2.4.1	Structure à l'équilibre	63
2.4.2	Perturbations eulériennes et lagrangiennes	65
2.4.3	Linéarisation des équations	66
2.4.4	Perturbations adiabatiques	68
2.4.5	Oscillations stationnaires	69
2.4.6	Décomposition radiale et angulaire des perturbations	71
2.4.7	Conditions aux limites en surface	72
2.4.8	Géométrie particulière du champ de déplacement	73
2.4.9	Séparabilité des solutions	74
2.4.10	L'équation d'onde	76
2.4.11	Analyse JWKB	78
2.4.12	Relation asymptotique des fréquences propres	80
2.5	La théorie des perturbations	82
2.5.1	Symétrie de l'opérateur \mathcal{L}	82
2.5.2	Forme intégrale des fréquences propres	84
2.5.3	Principe variationnel	84
2.5.4	Perturbation des fréquences	87
2.5.5	Inversion sismique	88
2.5.6	Le glitch sismique	89
 II Partie principale		93
 3 Le glitch sismique		95
3.1	Les origines des glitches sismiques	95
3.2	Le glitch de l'ionisation	97
3.3	Modélisation de la perturbation de structure	99
3.4	Expressions usuelles du glitch	100
3.5	Modèle de référence et diagnostic sismique	104
3.6	Limites et discussion	106
3.7	Limits and discussion [en]	110
 4 Modéliser la région d'ionisation		115
4.1	Properties of the ionisation glitch. I. Modelling the ionisation region	116
 5 Modélisation sismique du glitch de l'ionisation		135
5.1	Properties of the ionisation glitch. II. Seismic signature of the structural perturbation	136

6	Conclusion et perspectives	151
6.1	Modélisation de la région d'ionisation	151
6.2	Modélisation du glitch de l'ionisation	152
6.3	Perspectives	154
7	Conclusion and perspectives [en]	157
7.1	Modelling of the ionisation region	157
7.2	Modelling the ionisation glitch	158
7.3	Perspectives	160
III	Annexes	163
A	Compléments thermodynamiques	165
A.1	Définition de l'entropie	165
A.2	Expression des potentiels chimiques	167
A.3	Relations thermodynamiques locales	168
A.4	Variables d'états	169
A.5	Quantités thermodynamiques pour le fluide partiellement ionisé	170
A.5.1	Différentielles atomiques à l'équilibre chimique	170
A.5.2	Potentiel thermodynamique, énergie et énergie libre	171
A.5.3	Premier exposant adiabatique	173
A.6	Perturbation sous contrainte	175
B	Compléments sismiques	177
B.1	Dérivation de la perturbation de l'énergie gravitationnelle	177
B.2	Dérivation des kernels	180
B.3	Obtention de l'équation d'onde	184
B.3.1	Cas général	184
B.3.2	Approximation de Cowling	187
C	Compléments probabilistes	189
C.1	Contexte et notations	189
C.2	L'inférence bayésienne	190
C.2.1	Loi de probabilité, loi conditionnelle, loi marginale et indépendance	190
C.2.2	Formalisme de l'inférence	191
C.2.3	La plausibilité à priori	191
C.2.4	La vraisemblance	192
C.2.5	La hiérarchie bayésienne	193
C.2.6	La forêt de modèles	194
C.2.7	La prédiction de données	197
C.2.8	La théorie bayésienne de la décision	197
C.3	Les processus gaussiens	199
C.3.1	Les processus	199
C.3.2	Propriétés des processus gaussiens	200
C.3.3	Application au formalisme de l'inférence	201
C.3.4	Kernels, "paramètres" et hyperparamètres d'un processus gaussien	202
C.3.5	Ajout d'une base de fonctions à un processus gaussien	206
	Bibliographie	213

Table des figures

1.1	Représentation schématique de modes p & g dans une étoile.	5
1.2	Visibilité d'un mode d'oscillation stellaire en fonction de son degré ℓ	7
1.3	Spectre de puissance issue de la courbe de lumière de Kepler-93.	8
1.4	Illustration schématique de quelques biais de modélisation classiques	9
1.5	Illustration de la dégénérescence $M - Y_0$	10
1.6	Illustration schématique d'une levée de dégénérescence $M - Y_0$	11
2.1	Représentation schématique de la densité de probabilité à 1 particule	27
2.2	Évolution temporelle d'un champ moyen $\bar{\varphi}$	30
2.3	Illustration du critère de stabilité sur la stratification stellaire.	50
2.4	Évolution de la structure interne d'une étoile de $1.3M_{\odot}$	55
2.5	Trajectoires évolutives de modèles stellaires en fonction de leur masse.	61
2.6	Représentation schématique des contraintes s'appliquant à la propagation d'une onde dans un intérieur stellaire	70
2.7	Schéma illustrant les cavités (p) et (g) de plusieurs modes propres.	77
2.8	Illustration des écarts entre des perturbations de structure en fonction du choix de variable utilisé.	85
2.9	Illustration d'une inversion sismique du profil de vitesse du son pour le Soleil.	90
2.10	Illustration du glitch sismique engendré par une perturbation de structure qui prend la forme d'un Dirac.	91
3.1	Représentation des régions à même de causer un glitch sismique pour une étoile de type solaire	96
3.2	Variations du premier exposant adiabatique, Γ_1 , dans la zone d'ionisation.	98
3.3	Idéalisation de l'impact d'une variation d'hélium δY sur le premier exposant adiabatique Γ_1	100
3.4	Trois exemples de modèles de perturbations de structure à l'origine d'un glitch.	101
3.5	Exemples de trois glitches ayant pour origine un profil de perturbation gaussien.	104
3.6	Illustration du compromis à trouver pour le choix d'un modèle de référence "lisse".	109
3.7	Illustration of the trade-off when choosing a 'smooth' reference model.	113
C.1	Illustration schématique de la distribution sur la température effective obtenue à partir des modèles photométrique et spectroscopique	195
C.2	Ajustement d'un jeu de données par un processus gaussien.	203
C.3	Matrice de covariance d'un processus gaussien après ajustement.	205

Liste des tableaux

3.1	Quelques hypothèses sur lesquelles reposent l'exploitation du glitch d'ionisation	110
3.2	Some of the assumptions on which the exploitation of the ionisation glitch is based	114
6.1	Comparaison des hypothèses et limites entre les approches précédentes et celle introduite dans cette thèse pour l'exploitation du glitch de l'ionisation	154
7.1	Comparison of the assumptions and limits of previous approaches with the one introduced in this thesis for the exploitation of the ionisation glitch	160

Première partie

Introduction

Chapitre 1

Introduction

Contenu du chapitre

1.1	Prologue	3
1.2	L'astérosismologie	4
1.3	La dégénérescence masse-hélium	8
1.4	Levée de la dégénérescence	11
1.5	La région d'ionisation et le glitch sismique	12
1.6	Objectif et contenu du manuscrit	13
1.7	Purpose and content of the manuscript [en]	14

1.1 Prologue

La majeure partie du ciel ne se dévoile que par la lumière qui nous parvient. Observatoires et missions spatiales scrutent cette immensité, rendue toujours plus vaste par des instruments optiques toujours plus perfectionnés, dans le but d'en saisir la nature. Cependant, aussi loin que porte leur regard, l'œil des télescopes n'a de cesse de se poser sur un obstacle, une barrière qui semble toujours se substituer à la nature des choses et ne laissant s'échapper que quelques bribes de photons pour message. Le ciel se trouve ainsi peuplé, d'objets opaques, fermement hermétiques à notre seul vecteur d'information. Alors même qu'elles sont le lieu abritant la plus grande richesse de phénomènes physiques, les strates inférieures des étoiles paraissent alors hors d'atteinte. Sur bien des aspects cependant, il serait hâtif de conclure que ces astres échappent à notre compréhension.

Les étoiles se trouvent être gouvernées par des lois, tantôt bien établies, souvent mal comprises. Derrière une surface ne peut donc se dissimuler n'importe quelle structure. Si les possibilités restent immensément nombreuses, des principes physiques nous permettent de modéliser ces structures et, notamment, de les caractériser par un faible nombre de données que l'on appellera *paramètres stellaires*. Ces paramètres sont d'une importance fondamentale pour déterminer la structure et l'évolution future de l'étoile. Parmi les plus déterminants, on pourra évidemment évoquer la masse, M , ou le rayon, R , de l'astre. Puisque l'intérieur de l'étoile agit de manière déterminante sur l'allure de sa surface, l'évaluation de ces paramètres peut se faire de manière indirecte. Ainsi pourront être considérés comme plus plausibles les paramètres menant à des grandeurs de surface plus proches de celles qui sont observées. Ces grandeurs, visibles de par leur influence sur la lumière de l'étoile, sont (bien-)nommées *observables* et l'on y retrouve notamment les quantités issues du spectre lumineux de l'astre : la luminosité, L , la température effective T_{eff} ou encore la métallicité $[\text{Fe}/\text{H}]$.

Évidemment, le problème de la caractérisation stellaire ne saurait se résumer si sommairement. Au-delà du fait que similarité des conséquences ne fait pas parenté des causes, j'évoquais plus haut que toutes les lois régissant la structure ne sont pas toujours connues. Ainsi, loin d'un cadre idéal où, compte tenu de nos observables, on cherche des paramètres stellaires par le biais

d'une relation univoque, on dispose plutôt d'une multitude de modèles candidats et d'autant plus de possibilités concernant leurs paramètres. À cet égard, je me permettrai la remarque qui suit. Il me semble assez répandu de penser la détermination des observables comme un moyen plutôt qu'une fin, et ainsi que peu seraient intéressés par l'obtention de toutes les caractéristiques d'une lumière stellaire si cela devait se payer par l'impossibilité de les exploiter. En un sens, j'aurais tendance à penser que la détermination des paramètres d'une étoile est, lorsque érigée en objectif, tout aussi vaine. Si la satisfaction que procure l'estimation de quantités inobservables est indéniable, beaucoup n'ont de sens qu'entre les limites d'un modèle. Leurs déterminations n'apportent donc que peu de compréhension si elles ne servent pas un but plus global. Ainsi, à bien des égards, il me semble que l'on place intuitivement notre motivation dans la compréhension des lois sous-jacentes elles-mêmes, *la physique interne* de l'étoile. Cette connaissance apporte d'abord la capacité de prédire toute structure interne, issue du plus plausible au plus hypothétique jeu de paramètres. Mais surtout, par l'agencement des relations entre les objets que cette physique manipule, elle constitue sans doute ce qui s'approche le plus d'une description d'une réalité à jamais hors d'atteinte.

Il semble qu'on touche là à une contradiction. Moi-même, lorsque interrogé sur la nature et le but de mon travail, ai (trop souvent) tendance à m'attarder sur la détermination de tel ou tel paramètre stellaire plutôt que sur la compréhension de la physique interne de l'étoile. Après tout, mon l'objectif de thèse ne se trouve-t-il pas être l'inférence de paramètres? Pour faire écho à la remarque soulevée plus haut, il semble ainsi que le moyen se substitue souvent à la fin. En effet, la recherche de modèles physiques adéquats ne peut être découplée de la détermination des paramètres stellaires. Comme élégamment formulé par [BROWN et al., 1994](#),

In general, we find that stars obeying different physics succeed remarkably well in masquerading as stars that merely have different parameters.

L'inférence de ces paramètres constitue en réalité une étape essentielle à la compréhension et l'établissement des modèles stellaires. Les deux problèmes se recouvrent et l'on en donnera un exemple fameux dans la suite de cette introduction.

1.2 L'astérosismologie

Parmi les outils à notre disposition, l'astérosismologie se caractérise par l'étude des modes propres d'oscillation au sein d'un astre. Si l'idée de sonder une structure en étudiant ses résonances n'est pas novatrice en soi, son attrait est flagrant dans notre situation. Bien des informations vis-à-vis des oscillations d'une étoile se trouvent être visibles en surface, en effet. De par l'étude des variations temporelles de luminosité, on dispose notamment de renseignements sur les amplitudes, fréquences, ou encore longévités de nombreux modes. Ces grandeurs peuvent être inférées de l'analyse fréquentielle de la lumière reçue par l'étoile et seront nommées ici observables sismiques. Je me contenterai donc de mentionner que le problème de l'inférence sismique peut être mené dans un cadre bien plus large que ne le fait ce manuscrit, qui ne s'intéressera qu'à l'exploitation des fréquences propres d'oscillation. De façon intuitive, on peut y voir la volonté d'étudier le problème sous un angle idéal, en se concentrant uniquement sur la structure de ces modes et en mettant de côté les mécanismes d'excitation ou d'amortissement des ondes à leur origine. On verra déjà que ce seul pan du problème reste assez riche et complexe pour mériter toute notre attention.

Indépendamment, donc, de toute notion d'émergence ou de dissipation du mode, il est courant de faire la distinction entre les deux principaux mécanismes à l'origine de sa survie. Tout d'abord, une oscillation peut résulter d'un incessant combat entre zones de surpression et de souspression. Le gradient de pression tend à faire diminuer la pression où elle est la plus élevée, comprimant à son tour son environnement où elle était plus faible. Il s'agit du mécanisme à l'œuvre dans la propagation des ondes sonores dans l'air ambiant par exemple. On appelle les modes propres résultant d'un tel mécanisme les *modes p* (pour pression). Ces derniers sont à distinguer des *modes g* (pour gravité), lesquels sont causés par l'action de la gravité. Dans l'étoile,

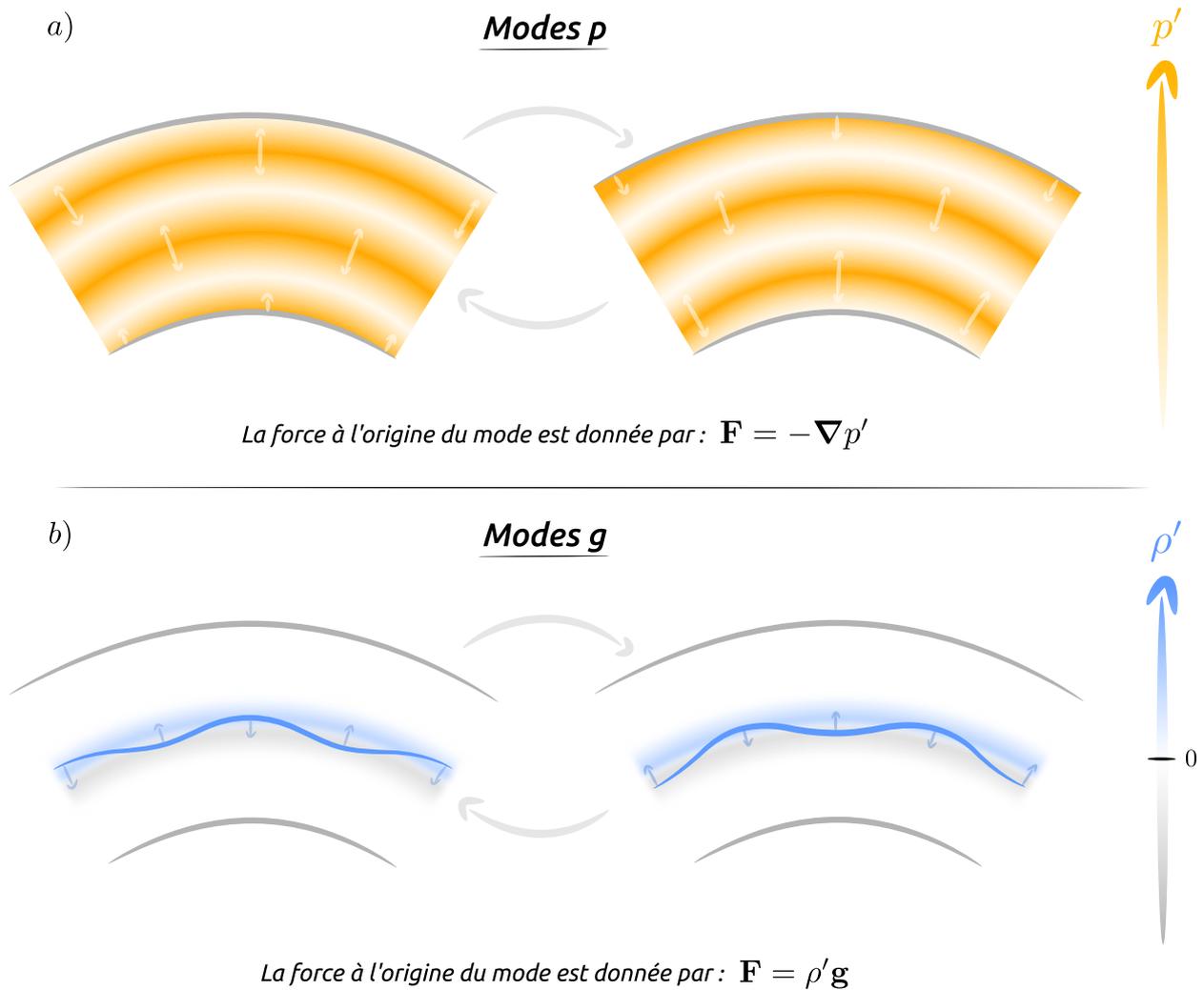


FIGURE 1.1 – Représentation schématique de modes p & g dans une étoile. Chaque schéma illustre la coupe d'une structure stellaire à deux instants avec les forces y prenant place (cf. flèches). **a)** Modes de pression (p). La force qui régit la structure du mode est la force de pression relative, $\mathbf{F} = -\nabla p'$ (ici le ' indique la valeur d'une grandeur par rapport à son environnement). Celle-ci tend à causer des surpressions locales, lesquelles vont disparaître à l'instant suivant. Notons que, d'un instant à l'autre, ce mode donne l'impression de "se déplacer" radialement dans la structure. **b)** Modes de gravité (g). La force à l'action est cette fois-ci la gravité relative vis-à-vis de l'environnement, $\mathbf{F} = \rho' \mathbf{g}$. Par rapport à une situation d'équilibre, les particules formant la ligne bleue vont tantôt se retrouver plus dense que leur entourage (la gravité va les faire plonger) ou moins dense que ce dernier (la gravité va les faire remonter). Cette fois-ci, le mode donne l'impression d'une "propagation" horizontale.

des éléments de matières peuvent être amenés dans des régions où ils sont plus denses que leur environnement. La gravité causera alors leur chute vis-à-vis de cet environnement. Ainsi entraînés, ils peuvent finir par se retrouver dans une zone qui les dépasse en densité. La gravité relative va alors avoir tendance à les faire remonter, entraînant le mouvement d'oscillation. Ces deux mécanismes sont schématiquement illustrés en FIGURE 1.1. Les modes p et g sont évidemment loin de donner une représentation complète de la variété d'oscillation que l'on peut trouver dans une étoile. Certains modes peuvent tout d'abord exhiber des propriétés propres à chacune des deux catégories, on parle alors de *modes mixtes*. De façon plus générale, bien d'autres phénomènes dynamiques peuvent survenir dans la structure d'une étoile (rotation, magnétisme, convection, ...), lesquels sont entre autres à l'origine des ondes inertielles, d'Alfvén ou de convection (GOUGH, 1993). Dans la faune de structures oscillatoires qui en émerge, on réduira encore le cadre de

discussion à l'étude des modes de pression purs dans ce qui suit. Comme expliqué plus haut, le mécanisme à leur origine est assez peu sensible à la façon dont la structure est agencée, propriété qui leur permet de subsister dans de nombreuses régions stellaires. Ces modes sont ainsi à l'origine de la plupart des fréquences d'oscillation détectées à la surface d'étoiles similaires au soleil. On appelle ces étoiles les *pulsateurs de type solaire*, et elles constituent le principal objet d'étude de ce manuscrit. Similaires en termes de structure, elles exhibent des propriétés sismiques proches du soleil mais différent pourtant de ce dernier sur un aspect essentiel : la distance qui nous sépare. Ce point fait une différence sur la qualité et la quantité des informations extraites de l'observation et participe à la distinction qui est souvent faite entre les sciences d'héliosismologie (l'étude sismique du soleil) et d'astérosismologie (l'étude sismique des astres, terme parfois employé pour désigner celle des étoiles, soleil exclu). Pour le comprendre, on peut succinctement s'attarder sur les propriétés des fréquences d'oscillation.

Il est commode et donc d'usage courant d'identifier les fréquences d'oscillation d'un astre par la spécification de deux nombres¹ : un ordre radial, n , et un degré, ℓ . Ceux-ci sont intimement liés à la géométrie du mode correspondant. Ainsi, le premier caractérise une périodicité radiale, un nombre de nœuds dans cette direction, tandis que le second spécifie plutôt une périodicité méridionale (d'un pôle à l'autre). Ce qui fait la différence entre le soleil et les autres étoiles est la résolution de sa surface. Celle-ci permet d'étudier indépendamment le mouvement² de nombreuses positions à la surface et ainsi de résoudre des modes de faible longueur d'onde et donc de degré ℓ élevé ($\ell > 1000$, CHRISTENSEN-DALSGAARD, 2002). À l'inverse, les variations de luminosité qui nous viennent des autres étoiles sont en fait intégrées sur la demi-sphère qui nous fait face. Les mouvements observés doivent donc être tels qu'ils ne s'annulent pas en moyenne, et ainsi d'une longueur d'onde suffisamment grande (cf. FIGURE 1.2, BROWN et GILLILAND, 1994). Ceci explique que les fréquences d'oscillation exploitées en astérosismologie soient associées aux degrés $\ell = 0, 1, 2$ et parfois 3.

En dépit de ce nombre de fréquences moindre, il est maintenant bien établi que l'exploitation de données sismiques peut beaucoup apporter à la compréhension d'une structure stellaire. En ne reposant que sur deux observables sismiques, la grande séparation $\Delta\nu$ et le pic du spectre en fréquence ν_{\max} (cf. FIGURE 1.3), il est déjà possible d'obtenir des estimations de la masse et du rayon d'une étoile par des relations d'échelle (ULRICH, 1986; BROWN et GILLILAND, 1994; KJELDSEN et BEDDING, 1995; BELKACEM et al., 2011). Ces évaluations possèdent l'avantage de nécessiter peu de fréquences en plus d'être déjà bien plus fines que ce qui pouvait être déterminé à partir des observables spectroscopiques seules (SILVA AGUIRRE et al., 2013; LEBRETON et GOUPIL, 2014; DAVIES et MIGLIO, 2016). Si davantage de données sismiques se trouvent être disponibles, il est alors possible de tirer profit de chaque fréquence propre. Bien entendu, il serait faux de considérer que la mesure d'une vingtaine de fréquences d'oscillation nous en apprendrait vingt fois plus qu'une observable classique comme la luminosité de l'astre. Leur grande régularité (cf. FIGURE 1.3) fait que l'information contenue dans ces fréquences possède une certaine part de redondance. Celle-ci s'explique notamment par l'existence d'une relation linéaire que les fréquences propres vérifient asymptotiquement pour $\ell \ll n$ (VANDAKUROV, 1968; TASSOUL, 1980) :

$$\nu_{n\ell} \simeq (n + \ell/2 + N_{\text{eff}}/2 + 1/4)\Delta\nu. \quad (1.2.1)$$

avec N_{eff} une quantité liée à la structure de l'ordre de l'unité.

Si cette relation décrit plutôt fidèlement la répartition des fréquences mesurées pour les pulsateurs de type solaire, l'information détenue par chaque mode réside davantage dans ce que l'équation (1.2.1) ne parvient pas à expliquer. Intuitivement, l'écart d'une fréquence à une relation supposée trahit la présence d'une échelle spatiale sur laquelle la structure diffère de notre prédiction. L'intérêt d'une exploitation simultanée de toutes les fréquences (on parle parfois de sismologie "à la carte", LEBRETON et GOUPIL, 2014) se trouve donc dans la possibilité de faire

1. Comme on le verra par la suite un troisième nombre, l'ordre azimutal m , peut être nécessaire.
2. Que ce soit par des variations de luminosité ou bien par spectrométrie Doppler.

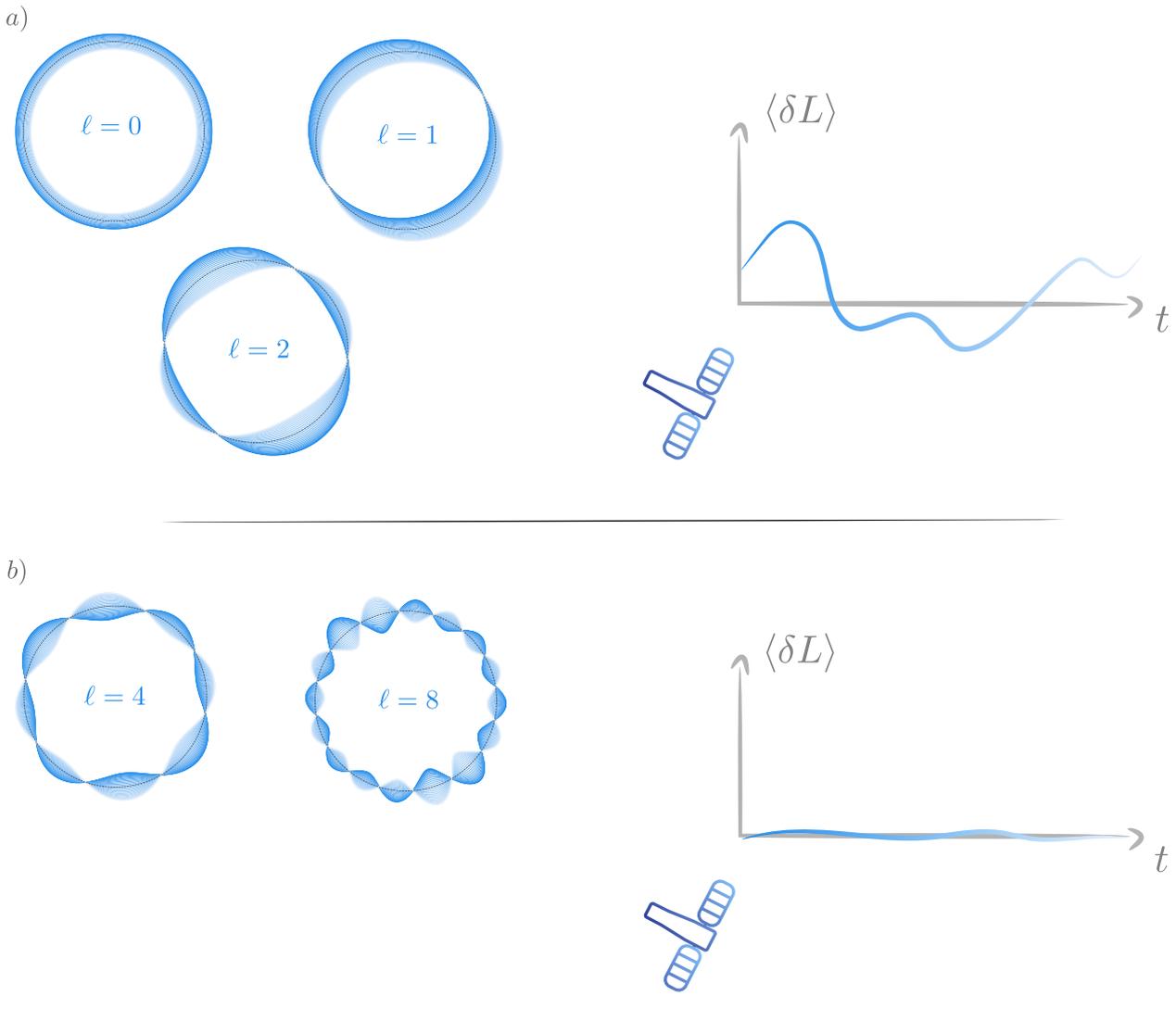


FIGURE 1.2 – Visibilité d'un mode d'oscillation stellaire en fonction de son degré ℓ . Ici, chaque diagramme circulaire représente une coupe méridionale de l'astre avec sa surface en l'absence d'oscillation (en pointillé noirs) et la déformation causée par des modes de degrés ℓ divers (en bleu, elle est bien sûr amplement accentuée) au cours du temps (du plus foncé au plus clair). La variation de luminosité causée par l'oscillation, δL , peut être pensée comme variant avec l'amplitude du mode en chaque point. **a)** Modes de bas degré. Les variations de luminosité intégrées sur la face observée par l'instrument, $\langle \delta L \rangle$, sont visibles en fonction du temps t , car de moyenne non nulle. **b)** Modes de plus haut degré. À chaque instant, les variations de luminosité moyennes sont quasi-nulles, ce qui rend la caractérisation du mode presque impossible. Notons que pour une étoile dont la surface est résolue, les variations de luminosité en un point donné sont pourtant détectables.

ressortir des différences locales par rapport à un modèle de l'étoile. Cette capacité à sonder différemment une structure est donc un outil essentiel pour tester la physique de ces modèles, dont les effets sur la structure dépendent largement des diverses régions de l'étoile. Si l'on a conjecturé (GOUGH et KOVOCHEV, 1993; KOVOCHEV, 1993; BROWN et al., 1994) puis de maintes fois démontré la précision (MATHUR et al., 2012; METCALFE et al., 2014; LEBRETON et GOUPIL, 2014; NSAMBA et al., 2018) de l'approche globale ainsi définie, celle-ci a également permis de mettre en lumière les biais de nos modèles stellaires (BASU et ANTIA, 2004; ASPLUND et al., 2009; SERENELLI et al., 2009; SERENELLI, HAXTON et PEÑA-GARAY, 2011; LEBRETON et GOUPIL, 2014; BULDGEN et al., 2019b).

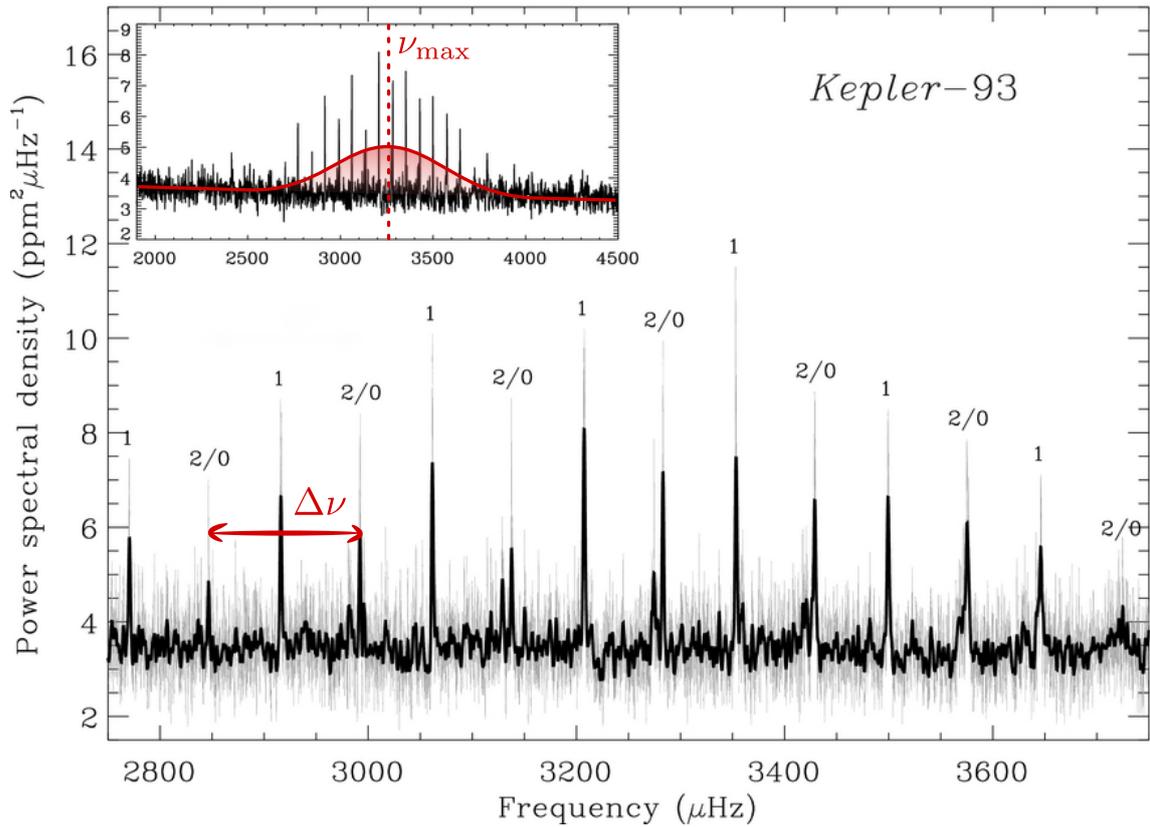


FIGURE 1.3 – Issu de BALLARD et al., 2014. Spectre de puissance issue de la courbe de lumière de Kepler-93. Les fréquences propres de l'étoile (pics du spectre) apparaissent clairement auquel ont été ajoutés les degrés des modes correspondant ($\ell = 0, 1$ ou 2). L'étendue de la grande séparation, $\Delta\nu$, définie par l'écart entre deux modes successifs de même degré, est indiquée. Le cadre du haut donne l'allure de l'ensemble du spectre, dont le maximum global définit la valeur de ν_{\max} .

1.3 La dégénérescence masse-hélium

Un modèle, qu'il vise à simuler la structure d'une étoile ou non, peut être sujet à de nombreux biais (cf. FIGURE 1.4). Cela peut être dû aux données auxquelles il est exposé mais également à sa propre conception (complexité, paramètres utilisés, phénomènes ignorés, ...).

Pour ce qui est de la modélisation stellaire, un biais déterminant a été pointé par LEBRETON et GOUPIL, 2014. Comme évoqué plus haut, on dispose d'un ensemble de processus physiques que l'on peut inclure ou non dans nos calculs, menant à un grand nombre de possibilités pour modéliser la structure d'une étoile. Si certains phénomènes ne vont causer que de faibles modifications, parfois non perceptibles, d'autres seront au contraire essentiels à la reproduction des données observées. Bien sûr, il peut être difficile de savoir quoi inclure à l'avance et, chaque processus venant avec ses propres paramètres, il est courant de ne pouvoir tous les contraindre à la fois. Certains paramètres se voient ainsi fixés sans pouvoir être ajustés, que ce soit empiriquement ou non, de façon à pouvoir estimer les quantités les plus importantes comme la masse M , le rayon R , l'âge de l'étoile A ou les abondances en hélium et en métaux à la formation de l'étoile, Y_0 et Z_0 . L'étude effectuée par LEBRETON et GOUPIL, 2014 visait notamment à estimer l'impact de la physique et des paramètres laissés fixes sur la détermination des grandeurs sus-citées. De nombreuses variantes ont ainsi été étudiées, portant à la fois sur la *microphysique* (équation d'état, opacités, taux de fusion nucléaire, ...) et la *macrophysique* (diffusion, convection, overshoot, rotation, atmosphère, composition, ...) des modèles. Toutes les notions juste évoquées seront introduites dans la section 2.3, on peut néanmoins déjà apprécier le nombre d'ajustements que cette liste procure.

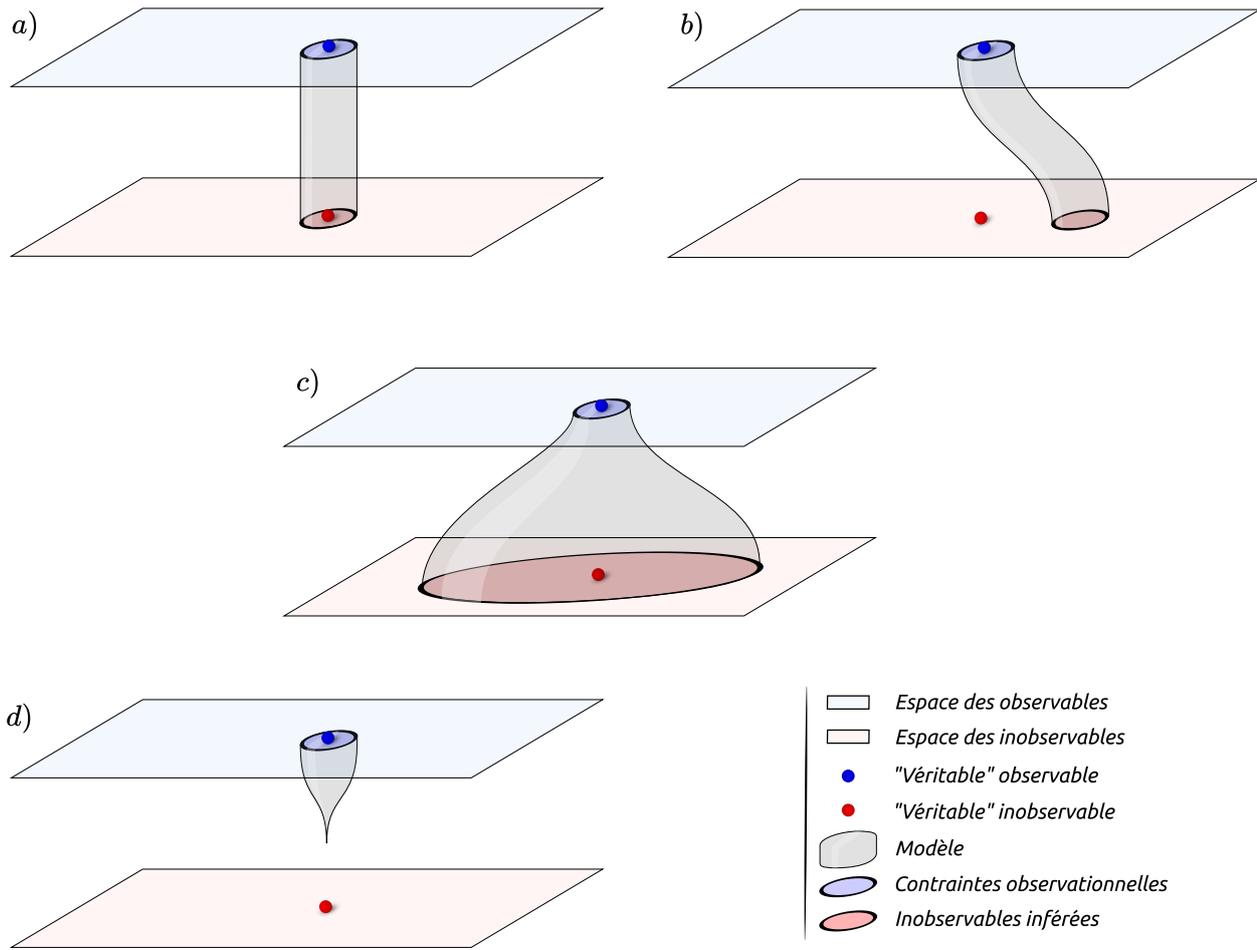


FIGURE 1.4 – Illustration schématique de quelques biais de modélisation classiques. Appliqué au problème de l'inférence de paramètres stellaires, l'espace des observables pourrait contenir les grandeurs spectroscopiques observées ainsi que les fréquences d'oscillation de l'étoile. L'espace des inobservables contiendrait, quant à lui, les paramètres fondamentaux comme la masse, M , et le rayon, R , de l'étoile mais aussi toutes les quantités qu'il convient de spécifier pour générer un modèle stellaire (voir texte). **a)** Scénario idéal dans lequel, pour des contraintes observationnelles adéquates, le modèle fournit des contraintes non biaisées sur les paramètres. **b)** Manifestation possible de processus physiques manquants ou mal pris en compte dans le modèle. Il en résulte des paramètres stellaires mal déterminés bien que les observations soient idéales. **c)** Exemple de dégénérescence. Les paramètres peuvent largement varier et quand même reproduire les observables alors que celles-ci sont fortement contraintes. Cela peut notamment venir d'un jeu d'observable mal choisi (redundant ou inadapté à l'information que l'on souhaite obtenir) ou encore d'une trop grande complexité du modèle, qui ne peut être contraint malgré des observations pertinentes. **d)** Cas dans lequel aucun jeu de paramètres ne peut reproduire les observations (du moins à une précision plausible). Cela témoigne généralement d'un processus physique manquant dans le modèle.

Offertes par ces nombreuses possibilités, plusieurs estimations de la masse M et de l'abondance initiale en hélium Y_0 du pulsateur de type solaire HD 52265 peuvent être trouvées sur la FIGURE 1.5. Insistons sur le fait que cette analyse repose à la fois sur l'exploitation de données spectroscopiques mais aussi sur celle de données sismiques précises. S'il n'est pas fondamentalement surprenant que des estimations basées sur des modèles physiques distincts diffèrent, on peut néanmoins être étonné par leur distribution et l'amplitude de ces différences. En particulier, lorsque comparés aux incertitudes directement issues des observations, les écarts entre les estimations sont très importants et beaucoup semblent s'exclure mutuellement. Cette idée peut se formuler comme suit :

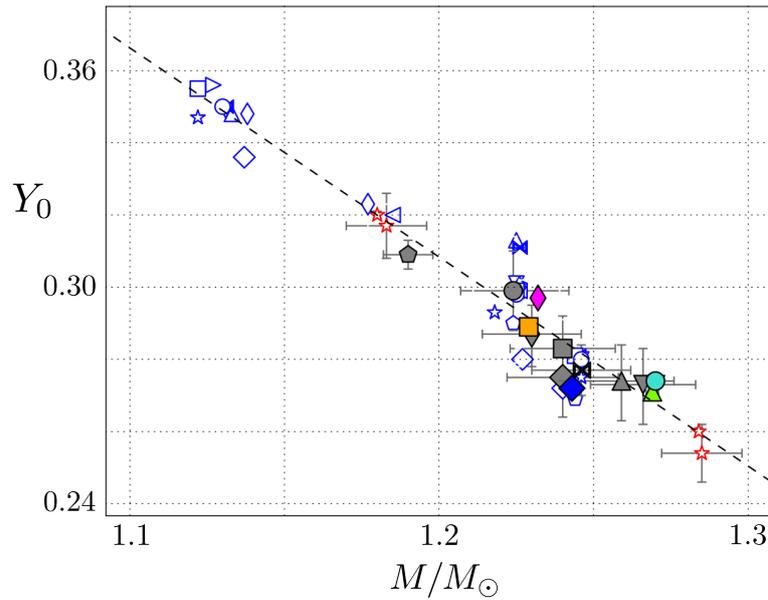


FIGURE 1.5 – *Issu de LEBRETON et GOUPIL, 2014.* Illustration de la dégénérescence $M - Y_0$ pour l'étoile HD 52265. Chaque symbole représente ici une estimation de la masse et de la fraction en l'hélium initiale de l'astre pour un modèle de physique interne différent (voir texte). Celles-ci sont déduites à la fois des observables spectroscopiques (température effective T_{eff} , luminosité L ou gravité de surface $\log g$ et métallicité $[\text{Fe}/\text{H}]$) et sismiques (les fréquences d'oscillations ν_{nl}) à l'aide de données CoRoT (BAGLIN et al., 2002; BAGLIN et al., 2009).

Selon un modèle stellaire \mathcal{M}_1 , il est courant de juger comme *très peu probable* l'estimation la *plus probable* du couple (Y_0, M) faite par un modèle stellaire voisin \mathcal{M}_2 .

Une raison à cela peut être trouvée dans la forme spécifique que prend la distribution de ces estimations. Loin de se répartir de façon isotrope autour d'une valeur moyenne, celles-ci s'étirent sur une direction spécifique du plan (en tirets noirs) sur laquelle il semble difficile de définir une contrainte. C'est ce phénomène que l'on qualifie de dégénérescence, en l'occurrence de *dégénérescence masse-hélium*. Il en résulte une incapacité à définir une région bornée contenant les estimations de chacun des paramètres (cf. schéma FIGURE 1.6, panneau **a**)), occasionnant une large incertitude vis-à-vis de ces derniers. Si j'ai déjà eu l'occasion d'évoquer l'intérêt d'une détermination de la physique des modèles stellaires, une incapacité à estimer de façon fiable un paramètre tel que la masse est un problème en soi. Outre le fait que cette grandeur peut être définie quelque soit le modèle, ce qui lui confère un sens physique propre, elle est aussi déterminante pour nombre d'applications. On peut citer à cet égard : la caractérisation d'un scénario évolutif particulier ou, au contraire, l'étude statistique des propriétés d'un ensemble d'étoiles ainsi que l'estimation de la masse des corps orbitant autour d'une étoile et donc de ses exoplanètes.

Je me permets ici un aparté. À la vue des incertitudes relevant de la propagation directe des incertitudes observationnelles, le problème ici établi ne peut pas être résolu en se contentant d'attendre des données sismiques plus précises (du moins pas directement). Cela se traduirait simplement par la diminution des barres d'incertitudes fournies par chacun des modèles, ce qui en l'état ne ferait que renforcer les exclusions mutuelles des estimations de ces modèles.

Le problème de la dégénérescence masse-hélium n'est ainsi pas un problème de *précision* mais d'*exactitude*.

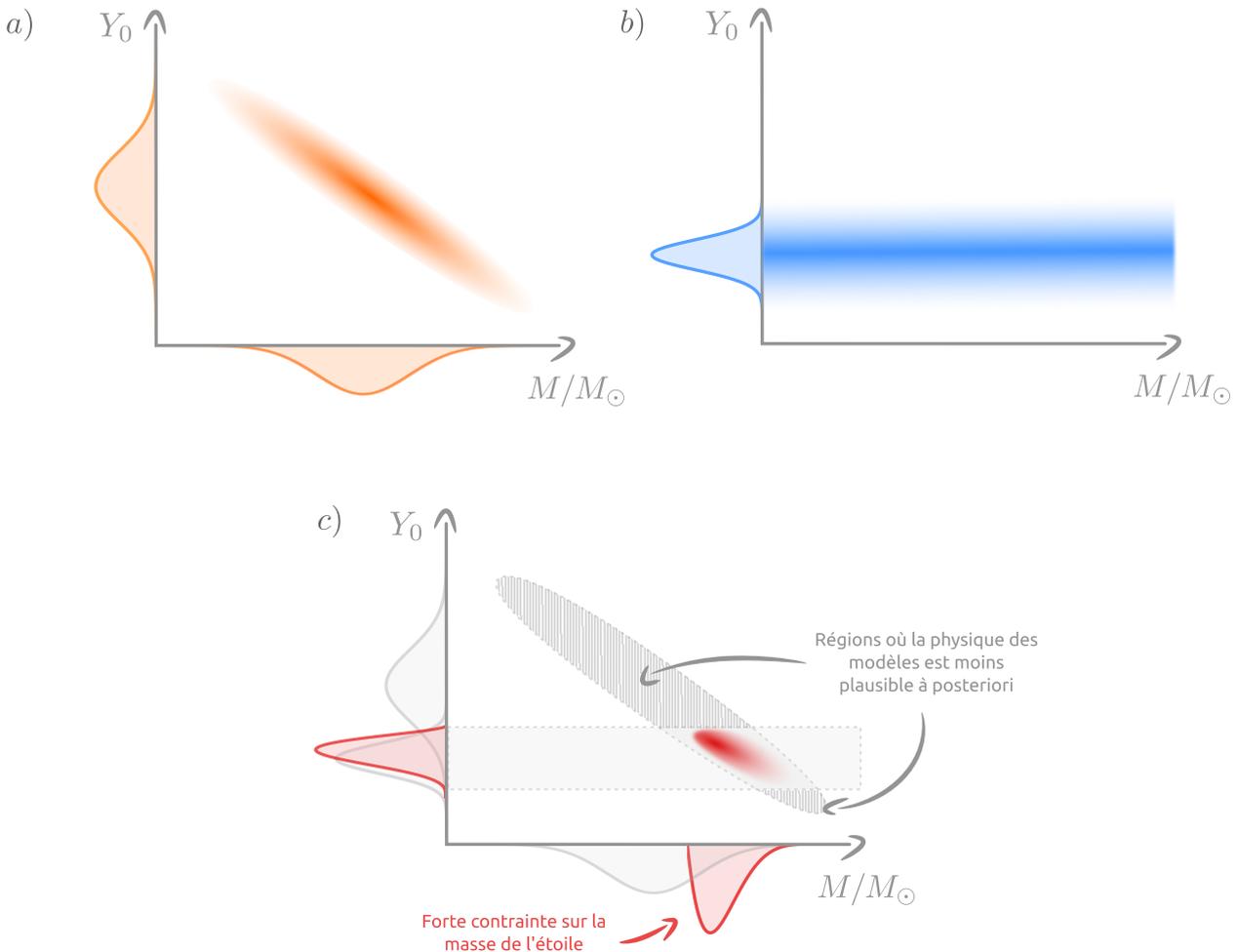


FIGURE 1.6 – Illustration schématique d’une levée de dégénérescence $M - Y_0$ par une estimation indépendante de Y_0 . **a)** Situation de la FIGURE 1.5, les nombreuses possibilités vis-à-vis de la physique interne de l’étoile amène toute une “droite” de l’espace des paramètres à être relativement probable (régions plus foncées). Les deux paramètres, Y_0 et M , sont assez mal contraints (cf. les densités de probabilité marginales sur chacun des axes). **b)** Représentation d’une contrainte sur Y_0 . Celle-ci n’apporte aucune information sur la masse de l’étoile en elle-même. **c)** Ajout de la contrainte indépendante **b)** par rapport à la situation **a)**. Cette dernière permet non seulement de contraindre assez nettement la masse de l’étoile mais également de rendre de nombreuses possibilités de physique interne moins plausibles (celles menant à des couples (M, Y_0) dans les régions hachurées).

Ce sont les biais propres à des modèles plausibles à priori qui causent l’impossibilité de fournir une estimation unique de la masse M et de l’abondance initiale en hélium Y_0 pour une étoile. **Il nous faut ainsi établir une méthode pour rendre moins plausible les modèles les plus biaisés; c’est dans cette optique que s’inscrit mon travail de thèse.**

1.4 Levée de la dégénérescence

Une solution intuitive à ce problème de dégénérescence est illustrée en FIGURE 1.6. S’il est possible d’obtenir une mesure de l’abondance initiale en hélium de l’étoile par une autre méthode³, la combinaison des deux estimations permet de lever la dégénérescence masse-hélium.

3. Notons que le raisonnement se transpose sans problème au cas d’une levée de dégénérescence par une mesure indépendante de la masse de l’étoile. Une telle mesure est notamment permise pour des étoiles binaires.

Bien que cette estimation indépendante ne concerne que l'hélium initial, la masse se trouve fortement contrainte par la direction privilégiée de la dégénérescence. Mais surtout, la plausibilité des modèles menant à des estimations de Y_0 incompatibles avec la nouvelle mesure se trouve réduite, ce qui nous fournit des contraintes sur la physique des modèles.

Ceci exposé, l'obtention d'une contrainte indépendante concernant l'abondance initiale en hélium est loin d'être évidente. En effet, les estimations présentées en FIGURE 1.5 résultent déjà de l'exploitation de données spectroscopiques en plus des fréquences d'oscillation prises individuellement, soit de l'ensemble des données à disposition. Toutefois, il convient de souligner que ces données sont utilisées de façon globale, dans le sens où l'estimation des paramètres stellaires (pour un modèle donné) résulte d'un compromis. Les "meilleurs" paramètres seront ainsi ceux qui minimisent l'écart à la reproduction de toutes les observables à la fois. Pour essayer de trouver une estimation indépendante⁴, il est alors possible de se limiter à un sous-ensemble des données qui serait particulièrement sensible au paramètre à déterminer.

Notons que dans notre cas aucune observable n'est directement sensible à l'abondance initiale en hélium de l'étoile. Il s'agit d'une grandeur utile pour paramétrer des modèles d'évolution stellaire mais qui ne possède pas de réalité physique à l'instant t de l'observation, contrairement à l'abondance en hélium⁵ à cet instant $Y(t)$. Ainsi, il semble naturel de chercher à lever cette dégénérescence indirectement, à l'aide d'une signature sensible à l'abondance en hélium dans les données. Pour les étoiles les plus chaudes en surface ($T_{\text{eff}} > 10\,000$ Kelvins), cette signature peut être directement trouvée dans le spectre de l'étoile. En effet, les raies d'absorptions de l'hélium sont alors visibles dans l'ultraviolet et un modèle d'atmosphère permet d'inférer son abondance (TRAVING, 1955; HUNGER, 1955; WEIDEMANN, 1955; ALLER, 1961). Cette approche n'est cependant pas possible pour les pulsateurs de type solaire, dont les températures de surface restent sous les 8000 Kelvins.

1.5 La région d'ionisation et le glitch sismique

Pour ces étoiles, une alternative consiste à s'intéresser à leur *région d'ionisation*. Au fur et à mesure que l'on s'enfonce dans la structure, la température augmentant, les éléments présents dans l'étoile ont de plus en plus de chance d'accéder à un état excitation. L'atome se voit fournir de l'énergie par son environnement et libère alors un électron en échange. Ce processus se répète en descendant dans l'étoile jusqu'à ce que l'élément n'ait plus d'électron à rendre; on dira alors qu'il est complètement ionisé. Tous les éléments, que ce soit l'hydrogène (de fraction en masse X), l'hélium (Y) ou les métaux (Z) sont sujets au processus d'ionisation. Les températures auxquelles il survient varient d'une espèce et d'un état d'excitation à l'autre. On parle malgré tout de région d'ionisation, car ce phénomène prend place dans une zone relativement restreinte (et généralement proche de la surface) de l'étoile, au-delà de laquelle la grande majorité des atomes sont complètement ionisés.

L'hélium étant le principal constituant des étoiles avec l'hydrogène (un ordre de grandeur typique pour Y est $1/4$), son impact sur la structure dans la région d'ionisation est particulièrement prononcé. Il peut être intuité de la façon suivante : dans la région d'ionisation de l'hélium, une partie non négligeable de l'énergie à disposition est allouée à son excitation. Ainsi, dans cette zone, certaines grandeurs physiques comme la température T (ou la vitesse de propagation des ondes c qui lui est directement reliée) ont tendance à augmenter "moins vite qu'elles ne le devraient" en s'enfonçant dans la structure. À l'inverse, la pression totale dans l'étoile augmente rapidement, la pression des électrons libérés par l'ionisation s'ajoutant à celle des atomes déjà

4. Il conviendrait ici d'explicitier ce que l'on entend par la notion "d'indépendance". Bien entendu, celle-ci n'est pas liée à une indépendance dans les données exploitées puisque l'estimation indépendante repose sur les mêmes observations. Néanmoins, cette estimation peut être le résultat d'un objectif différent (autre fonction à minimiser par exemple) basé sur un modèle différent et paramétré différemment. On pourrait résumer cela de la façon suivante : même information à disposition, procédure d'exploitation différente.

5. Puisque l'abondance en hélium d'une étoile n'est pas homogènement répartie à l'intérieur de la structure, on entendra souvent par là l'abondance d'hélium *en surface*.

présente. Une description plus formelle de ces phénomènes sera donnée dans le paragraphe 2.3.7. On peut néanmoins déjà conjecturer que ces conséquences seront d'autant plus prononcées que l'abondance d'hélium sera élevée.

Pour ces raisons, les fréquences d'oscillation de la structure sont impactées par la région d'ionisation et par l'abondance en hélium dans cette zone. Cette signature reste fine, ce qui explique qu'il soit nécessaire de mettre au point une procédure spécifique pour l'exploiter. Le fait qu'elle ait pour origine une région lui fournit cependant un aspect particulier, celui d'une oscillation (qui sera mise en évidence dans le paragraphe 2.5.6) au sein même des fréquences propres de l'étoile. On appelle une telle signature un *glitch sismique* et, dans ce cas particulier où elle prend son origine dans la région d'ionisation, on parlera du glitch de l'ionisation⁶.

1.6 Objectif et contenu du manuscrit

Le présent manuscrit vise ainsi à exploiter le glitch dû à l'ionisation pour déterminer l'abondance en hélium et donc, indirectement, nous renseigner sur la masse et la physique interne des pulsateurs de type solaire. On pourra alors remarquer que le titre de cette thèse n'est pas tout à fait conforme à l'objectif juste énoncé. En effet, il y est fait mention de *propriétés de la région d'ionisation* et non pas de l'abondance d'hélium. En réalité, si l'objectif initial de ma thèse résidait dans l'estimation de l'abondance en hélium, et ce par le cheminement que j'ai présenté dans les paragraphes précédents, il m'est apparu en cours de route que ce but était inutilement restrictif. Comme on le verra dans la partie principale du manuscrit, le glitch de l'ionisation est en mesure de nous renseigner sur d'autres grandeurs en plus de cette abondance, grandeurs que j'ai regroupées sous la notion de propriétés de la région d'ionisation. À l'instar de l'abondance en hélium et de ce qui est illustré en FIGURE 1.6, ces propriétés peuvent être utilisées comme autant de contraintes supplémentaires sur la masse et les modèles de l'étoile.

Je pourrais maintenant poursuivre par une explication moins évasive de ce qu'est le glitch de l'ionisation (ce sera fait dans le chapitre 3) avant d'aborder en détail mon travail de thèse (chapitres 4 & 5). J'ai tenu, avant cela, à donner un contexte aux notions qui seront abordées dans ces parties. Ainsi, avant d'introduire les notions de sismologie centrales à cette thèse (sections 2.4 et 2.5), on essaiera de donner une compréhension intuitive des principes fondamentaux de la physique stellaire dans la section 2.3. Ces notions permettront en outre d'aborder plus en détail en quoi consiste la modélisation stellaire dont il a été question dans cette introduction. Ce passage sur la physique stellaire sera précédé de deux sections théoriques portant respectivement sur la théorie cinétique (section 2.1) et les équations de l'hydrodynamique (section 2.2). Ces deux sections peuvent être considérées comme plus abstraites et moins essentielles pour appréhender le cœur du texte, il sera donc possible d'en faire abstraction en première lecture (comme indiqué par l'astérisque * qui précède le titre des sections). Leur but sera de dériver et de donner une compréhension des équations utilisées dans ce manuscrit, en plus de pouvoir expliciter les hypothèses sous lesquelles s'appliquent ces relations si souvent utilisées. Finalement, ce chapitre 2 se veut comme un zoom progressif, partant de notions très générales (les équations de Hamilton) pour en arriver à l'objet central de l'étude : le glitch sismique.

Le chapitre 3 donnera plus de détails sur ce qu'est le glitch sismique, ce qu'il permet et comment il a été utilisé pour sonder la structure stellaire. On en profitera pour présenter les travaux qui ont amené cette thèse, leurs apports et limites, et ainsi définir une problématique à cette étude. Suivront les chapitres 4 et 5 qui présenteront mon travail de thèse à proprement parler, de la modélisation de la région d'ionisation à celle du glitch de l'ionisation. Chacune de ces parties sera suivie des publications académiques qu'elles ont entraînées. Enfin, la dernière partie (chapitre 6, chapitre 7 pour la traduction anglaise) de ce manuscrit sera naturellement dédiée à ma conclusion et mes perspectives concernant l'approche présentée.

6. Récemment, la littérature semble avoir plutôt adoptée la dénomination de "glitch de l'hélium" (voir par exemple VERMA et al., 2017; VERMA et al., 2019; VERMA et SILVA AGUIRRE, 2019; FARNIR et al., 2019; FARNIR et al., 2020) pour désigner cette signature sismique. De par l'indissociabilité des signatures de régions de l'hydrogène et de l'hélium que l'on abordera dans le chapitre 5, on lui préférera ici l'appellation globale de "glitch de l'ionisation".

Concernant la forme, on relèvera que le texte de ce manuscrit s'accompagne de nombreuses équations. Néanmoins, les développements qui apparaîtront se veulent utiles et l'accent sera souvent mis sur l'interprétation de ces équations. Lorsque la situation s'y prête, cette interprétation se retrouvera appuyée par des visualisations schématiques des phénomènes en jeu. Les détails purement techniques n'apportant que peu de compréhension sur la physique sous-jacente seront quant à eux repoussés en annexe.

1.7 Purpose and content of the manuscript [en]

The present manuscript aims at exploiting the ionisation glitch to determine the helium abundance and hence, to indirectly inform us on the mass and internal physics of solar-like pulsators. It may then be noted that the thesis title is not entirely in line with the objective just stated since it mentions *physical properties of the ionisation region* and not the helium abundance. Actually, although the initial goal of my thesis was to estimate the helium abundance by the approach I presented in the previous paragraphs, it became clear along the way that this objective was unnecessarily restrictive. As will become evident in the main part of the manuscript, the ionisation glitch is able to inform us about other quantities in addition to this abundance, quantities that I have grouped as properties of the ionisation region. Much like the helium abundance and as illustrated in FIGURE 1.6, these properties can be used as additional constraints on the mass and stellar model.

I could now continue with a more detailed description of the ionisation glitch (this will be done in chapter 3) before discussing my thesis work (chapters 5). Before doing so, I wanted to provide a background to the notions that will be discussed in these sections. Thus, before introducing notions of seismology essential to this thesis (sections 2.4 and 2.5), we will try to give an intuitive understanding of the fundamental principles of stellar physics in section 2.3. These notions will allow us to discuss in more detail stellar modelling, which we have mentioned in this introduction. This part on stellar physics will be preceded by two theoretical sections respectively dealing with the kinetic theory (section 2.1) and the equations of hydrodynamics (section 2.2). These sections can be considered as more abstract and less essential to apprehend the core of the text, it will thus be possible to omit them during a first reading (as indicated by the asterisk * which precedes the sections' titles). Their aim will be to derive and give further insights into the equations used in this manuscript, in addition to clarifying the assumptions under which these much-used relations apply. As a result, this chapter can be seen as a progressive zoom, starting from very general notions (Hamilton's equations) to arrive at the main topic of the study: the seismic glitch.

Chapter 3 will give more details on what the seismic glitch is, what it enables and how it has been used to probe the stellar structure. We will use this opportunity to present the studies that led to this thesis, their contributions and limitations, thus defining what is at stake for the present work. Chapters 4 and 5 will then present my thesis work, from the modelling of the ionisation region to that of the ionisation glitch. Each of these parts will be followed by the academic publication they led to. Finally, the last part (chapter 6, chapter 7 for english translation) of this manuscript will naturally be dedicated to my conclusion and prospects concerning the approach introduced in this work.

Regarding the form, it will be noted that the text of this manuscript contains numerous equations. Nevertheless, the developments that appear are intended to be useful and the emphasis will, in many cases, be placed on the interpretation of these equations. Whenever relevant, this interpretation will be supported by schematic representations of the relevant phenomena. Purely technical details that provide little understanding regarding the underlying physics will be left to the appendices.

Chapitre 2

Théorie

Contenu du chapitre

2.1	* Préambule	16
2.2	* De la théorie cinétique aux équations dynamiques	27
2.3	Des équations dynamiques à la structure stellaire	43
2.4	Les perturbations stationnaires et l'astérosismologie	63
2.5	La théorie des perturbations	82

2.1 * Préambule

Contenu de la section

2.1.1	Le fluide	16
2.1.2	La densité de probabilité à N particules	17
2.1.3	Grandeur physique	18
2.1.4	Le développement en <i>cluster</i>	18
2.1.5	La densité de probabilité à 1 particule	18
2.1.6	Forme explicite de l'équation de Liouville	20
2.1.7	La hiérarchie BBGKY	20
2.1.8	L'Équation de Boltzmann	21
2.1.9	Invariant collisionnel	21
2.1.10	Le théorème H	22
2.1.11	La distribution localement maxwellienne	22
2.1.12	La limite maxwellienne	24
2.1.13	Développement de Chapman-Enskog	24
2.1.14	L'équilibre thermodynamique local	25
2.1.15	Bilan : les différentes échelles	26

On tâchera ici d'introduire un formalisme dans lequel on puisera largement pour écrire cette partie théorique : la théorie cinétique. On cherchera également à donner un sens aux divers échelles ainsi qu'aux notions d'équilibre local et global qui apparaîtront lors de l'étude de la structure stellaire et de ses oscillations. Ce préambule sera également l'occasion de rendre explicite les hypothèses sous-jacentes aux sections qui suivront et ainsi de faire état des limites de chaque cadre.

2.1.1 Le fluide

Dans la suite, on fera régulièrement allusion à un milieu, le fluide, dans lequel on décrira l'évolution de grandeurs physiques. On va donc commencer par donner un sens à ce que l'on entend au travers de cette notion de fluide. Intuitivement, un fluide regroupe un très grand nombre de particules, $N \gg 1$, lesquelles suivent des trajectoires s'influençant mutuellement au cours du temps. Ainsi, à tout instant t , une particule (i) dispose d'une position $\mathbf{q}_i(t)$ et d'une impulsion $\mathbf{p}_i(t)$ caractérisant son évolution. Formellement, la connaissance de l'ensemble des $\mathbf{q}_i(t)$ en tout temps suffit à décrire chacune des trajectoires et donc l'évolution du fluide dans son intégralité. En pratique, cependant, les variations de positions ne peuvent être dissociées des impulsions conformément aux équations de Hamilton :

$$\dot{\mathbf{q}}_i = \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i}, \quad \dot{\mathbf{p}}_i = -\frac{\partial \mathcal{H}}{\partial \mathbf{q}_i}, \quad (2.1.1)$$

qui lient dérivées temporelles des coordonnées $\dot{\mathbf{q}}_i(t)$ et $\dot{\mathbf{p}}_i(t)$ et gradient du Hamiltonien du système à N particules $\mathcal{H}(\mathbf{p}, \mathbf{q})$ ¹. L'état du fluide à un instant t est donc entièrement déterminé par les coordonnées $(\mathbf{p}(t), \mathbf{q}(t))$, de même que tous ces états futurs si l'on dispose de l'expression de l'Hamiltonien \mathcal{H} (et donc d'une formulation des lois physiques auxquelles le fluide se soumet). Cette description propose donc de voir le fluide comme évoluant dans un espace de $6N$ dimensions que l'on appelle *l'espace des phases*. Si ce formalisme est tout à fait adapté à la mécanique classique, il requiert quelques modifications pour décrire un fluide. Pour des phénomènes se déroulant à notre échelle, le nombre de dimension de cet espace est de l'ordre de 10^{24} , si bien qu'il

1. On se permettra la référence aux N coordonnées $(\mathbf{p}_i, \mathbf{q}_i)$ à la fois en omettant l'indice i .

devient absurde de vouloir décrire l'état du système par la position et l'impulsion de chaque particule. À l'inverse, un tel nombre de particules rend les descriptions statistiques particulièrement intéressantes.

2.1.2 La densité de probabilité à N particules

Plutôt que de chercher à connaître exactement la position de chaque particule $(\mathbf{p}(t), \mathbf{q}(t))$ dans l'espace des phases, on peut se demander avec quelle probabilité le système se situe dans le volume $d\Gamma = d\mathbf{p}d\mathbf{q}$ ($= \prod_{i=1}^N d\mathbf{p}_i d\mathbf{q}_i$), délimité par (\mathbf{p}, \mathbf{q}) et $(\mathbf{p} + d\mathbf{p}, \mathbf{q} + d\mathbf{q})$ à l'instant t ? En notant cette probabilité $dF_N(t)$, on introduit naturellement une densité de probabilité à N particules (YVON, 1935), $f_N(\mathbf{p}, \mathbf{q}, t)$, telle que :

$$dF_N(t) \equiv f_N(\mathbf{p}, \mathbf{q}, t)d\Gamma \quad \text{et donc} \quad \int_{\Gamma} f_N(\mathbf{p}, \mathbf{q}, t)d\Gamma = 1, \quad (2.1.2)$$

où $\Gamma \equiv \int d\Gamma$ désigne l'espace des phases dans son ensemble.

À ce stade, notons que l'étude de cette densité de probabilité f_N n'est aucunement plus simple que celle de la trajectoire de chaque particule composant le fluide. Déterminer la probabilité $dF_N(t)$ est en réalité aussi contraignant que de savoir si chaque particule i se situe dans un volume $d\mathbf{p}_i d\mathbf{q}_i$ à l'instant t . L'intérêt majeur de cette description réside dans une propriété de l'espace des phases lui-même. Regardons la variation de l'élément de volume $d\Gamma$ au cours d'un instant Δt le long de la trajectoire :

$$d\mathbf{p}d\mathbf{q}(t + \Delta t) = \left| \frac{\partial(\mathbf{p}, \mathbf{q})_{t+\Delta t}}{\partial(\mathbf{p}, \mathbf{q})_t} \right| d\mathbf{p}d\mathbf{q}(t). \quad (2.1.3)$$

En remarquant que les équations de Hamilton nous donnent au premier ordre

$$\begin{cases} \mathbf{p}(t + \Delta t) = \mathbf{p}(t) + \dot{\mathbf{p}}\Delta t + \mathcal{O}((\Delta t)^2) = \mathbf{p}(t) - \Delta t \frac{\partial \mathcal{H}}{\partial \mathbf{q}} + \mathcal{O}((\Delta t)^2) \\ \mathbf{q}(t + \Delta t) = \mathbf{q}(t) + \dot{\mathbf{q}}\Delta t + \mathcal{O}((\Delta t)^2) = \mathbf{q}(t) + \Delta t \frac{\partial \mathcal{H}}{\partial \mathbf{p}} + \mathcal{O}((\Delta t)^2) \end{cases} \quad (2.1.4)$$

le déterminant de la matrice Jacobienne s'écrit :

$$\left| \frac{\partial(\mathbf{p}, \mathbf{q})_{t+\Delta t}}{\partial(\mathbf{p}, \mathbf{q})_t} \right| = \left(1 - \Delta t \frac{\partial^2 \mathcal{H}}{\partial \mathbf{p} \partial \mathbf{q}} \right) \left(1 + \Delta t \frac{\partial^2 \mathcal{H}}{\partial \mathbf{q} \partial \mathbf{p}} \right) + \mathcal{O}((\Delta t)^2) = 1 + \mathcal{O}((\Delta t)^2). \quad (2.1.5)$$

L'équation (2.1.3) nous permet donc d'affirmer que l'élément de volume $d\Gamma$ est conservé (au premier ordre) le long d'une trajectoire de l'espace des phases. Puisqu'à tout instant on a $d\Gamma/dt = 0$, la définition (2.1.2) donne $dF_N/dt = 0$ et donc, pour un volume arbitraire sur la trajectoire :

$$\frac{d}{dt} \int_{\Gamma(t)} f_N(\mathbf{p}, \mathbf{q}, t) d\Gamma = \int_{\Gamma(t)} \frac{df_N}{dt}(\mathbf{p}, \mathbf{q}, t) d\Gamma = 0. \quad (2.1.6)$$

Cette propriété se vérifiant pour tout volume $\Gamma(t)$, elle conduit nécessairement à

$$\frac{df_N}{dt} = 0, \quad (2.1.7)$$

équation à laquelle on réfère par *théorème de Liouville* (LIOUVILLE, 1838). Cette relation nous fournit un outil fondamental pour suivre l'évolution du fluide dans l'espace des phases. En termes des coordonnées dans l'espace des phases, on pourra développer cette équation comme suit

$$\frac{\partial f_N}{\partial t} + \dot{\mathbf{q}} \cdot \frac{\partial f_N}{\partial \mathbf{q}} + \dot{\mathbf{p}} \cdot \frac{\partial f_N}{\partial \mathbf{p}} = 0. \quad (2.1.8)$$

2.1.3 Grandeur physique

Au-delà de la position des différentes particules dans l'espace des phases, on peut vouloir décrire une grandeur physique en une position \mathbf{r} du fluide au cours du temps t . Une telle grandeur, notons la $A(\mathbf{r}, t)$, dépend de la position de chaque particule non pas dans "l'espace physique" (auquel \mathbf{r} réfère) mais dans l'espace des phases. Mettons que la grandeur que l'on cherche à mesurer en \mathbf{r} est la température du fluide. Il est clair, qu'en plus de l'endroit où on la mesure, cette quantité va dépendre de l'impulsion $\mathbf{p}(t)$ des particules à cet endroit (plus précisément de la variance de cette impulsion) et en cet instant. De façon très générale, la grandeur A mesurée en \mathbf{r} peut être définie comme fonction d'un champ $a_{\mathbf{r}}$ microscopique (dans le sens où il varie à l'échelle des particules) défini sur l'espace des phases :

$$A(\mathbf{r}, t) = \langle a_{\mathbf{r}} \rangle \equiv \int_{\Gamma} a_{\mathbf{r}}(\mathbf{p}, \mathbf{q}) f_N(\mathbf{p}, \mathbf{q}, t) d\Gamma. \quad (2.1.9)$$

La valeur de A en \mathbf{r} à l'instant t peut être vue comme résultant d'une espérance, $f_N(\mathbf{p}, \mathbf{q}, t) d\Gamma$ étant la probabilité que le fluide soit dans l'état (\mathbf{p}, \mathbf{q}) en cet instant et $a_{\mathbf{r}}(\mathbf{p}, \mathbf{q})$ étant la valeur du champ en \mathbf{r} pour un fluide dans cet état. En d'autres termes, pour une description du champ a donnée, la connaissance de la densité de probabilité f_N suffit à estimer la valeur moyenne de ce champ A en tout point et à tout instant. Ceci étant dit, cette densité de probabilité reste une quantité extrêmement complexe à appréhender et il sera question par la suite de se ramener à l'étude d'objets plus simples.

2.1.4 Le développement en *cluster*

En se rappelant que (\mathbf{p}, \mathbf{q}) désigne la position dans l'espace des phases de toutes les particules à la fois, il peut être intéressant de considérer la décomposition du champ microscopique $a_{\mathbf{r}}$ suivante :

$$a_{\mathbf{r}}(\mathbf{p}, \mathbf{q}) = a_{\mathbf{r}}^{(0)} + \sum_{i=1}^N a_{\mathbf{r}}^{(1)}(\mathbf{p}_i, \mathbf{q}_i) + \sum_{i,j'=1}^N a_{\mathbf{r}}^{(2)}(\mathbf{p}_i, \mathbf{p}_{i'}, \mathbf{q}_i, \mathbf{q}_{i'}) + \dots \quad (2.1.10)$$

Ici $a_{\mathbf{r}}^{(j)}$ désigne la partie du champ total dépendant simultanément de la position de j particules dans l'espace des phases. Ainsi, $a_{\mathbf{r}}^{(0)}$ ne dépend d'aucune particule et peut être vu comme la contribution d'un champ extérieur. Le second terme représente la somme des contributions des particules prises individuellement et le troisième la contribution au champ global de leur interaction 2 à 2. Il paraît raisonnable de considérer que la majeure partie du champ global provient des contributions individuelles, puis au second ordre des interactions, etc... . Ainsi apparaît la hiérarchie : $a_{\mathbf{r}}^{(1)} \gg a_{\mathbf{r}}^{(2)} \gg a_{\mathbf{r}}^{(3)} \gg \dots$ et on pourra souvent considérer que

$$a_{\mathbf{r}}(\mathbf{p}, \mathbf{q}) \simeq a_{\mathbf{r}}^{(0)} + \sum_{i=1}^N a_{\mathbf{r}}^{(1)}(\mathbf{p}_i, \mathbf{q}_i) \quad (2.1.11)$$

au premier ordre.

2.1.5 La densité de probabilité à 1 particule

Cette approximation (qui n'est pas systématiquement valable) élimine une part de la complexité du problème. Sous cette hypothèse, la valeur moyenne du champ $A(\mathbf{r}, t)$ peut être écrite comme :

$$\begin{aligned} A(\mathbf{r}, t) &= \int_{\Gamma} a_{\mathbf{r}}(\mathbf{p}, \mathbf{q}) f_N(\mathbf{p}, \mathbf{q}, t) d\Gamma \\ &\simeq a_{\mathbf{r}}^{(0)} \underbrace{\int_{\Gamma} f_N(\mathbf{p}, \mathbf{q}, t) d\Gamma}_1 + \int_{\Gamma} \sum_{i=1}^N a_{\mathbf{r}}^{(1)}(\mathbf{p}_i, \mathbf{q}_i) f_N(\mathbf{p}, \mathbf{q}, t) d\Gamma. \end{aligned} \quad (2.1.12)$$

Penchons-nous sur la signification du terme $a_{\mathbf{r}}^{(1)}(\mathbf{p}_i, \mathbf{q}_i)$. En particulier, on peut se demander si le champ $a_{\mathbf{r}}^{(1)}$ dépend réellement de la position $(\mathbf{p}_i, \mathbf{q}_i)$ de la particule i , ou si la particule 1 placée au même endroit $(\mathbf{p}_1 = \mathbf{p}_i, \mathbf{q}_1 = \mathbf{q}_i)$ aurait eu le même effet. Se poser cette question revient à se demander si les particules i et 1 sont identiques. Généralement, il est possible de séparer les particules en un certain nombre d'espèces $N_{\text{sp}} \ll N$ au sein desquelles il est impossible de les différencier. La plupart des équations dérivées par la suite supposeront que toutes les particules sont identiques ($N_{\text{sp}} = 1$). Dans ce cas, on obtient

$$\begin{aligned} A(\mathbf{r}, t) &\simeq a_{\mathbf{r}}^{(0)} + \int_{\Gamma} \sum_{i=1}^N a_{\mathbf{r}}^{(1)}(\mathbf{p}_1, \mathbf{q}_1) f_N(\mathbf{p}, \mathbf{q}, t) d\Gamma \\ &\simeq a_{\mathbf{r}}^{(0)} + \int_{\Gamma_1} a_{\mathbf{r}}^{(1)}(\mathbf{p}_1, \mathbf{q}_1) \left[N \int_{\bar{\Gamma}_1} f_N(\mathbf{p}, \mathbf{q}, t) d\bar{\Gamma}_1 \right] d\Gamma_1 \end{aligned} \quad (2.1.13)$$

où l'on a introduit les notations $\Gamma_j = \prod_{i \leq j}^N d\mathbf{p}_i d\mathbf{q}_i$ et $\bar{\Gamma}_j = \prod_{i > j}^N d\mathbf{p}_i d\mathbf{q}_i$. Il est alors possible de définir la densité de probabilité² à 1 particule, f_1 , de la façon suivante :

$$f_1(\mathbf{p}_1, \mathbf{q}_1, t) \equiv N \int_{\bar{\Gamma}_1} f_N(\mathbf{p}, \mathbf{q}, t) d\bar{\Gamma}_1. \quad (2.1.14)$$

Cette densité est associée à la probabilité $dF_1 = f_1(\mathbf{p}_1, \mathbf{q}_1, t) d\Gamma_1 / N$ qu'une particule (n'importe laquelle) se situe entre $(\mathbf{p}_1, \mathbf{q}_1)$ et $(\mathbf{p}_1 + d\mathbf{p}_1, \mathbf{q}_1 + d\mathbf{q}_1)$, laquelle est bien plus adaptée au problème que ne l'était dF_N . Puisque A peut-être approximée par

$$A(\mathbf{r}, t) \simeq a_{\mathbf{r}}^{(0)} + \int_{\Gamma_1} a_{\mathbf{r}}^{(1)}(\mathbf{p}_1, \mathbf{q}_1) f_1(\mathbf{p}_1, \mathbf{q}_1, t) d\Gamma_1, \quad (2.1.15)$$

la détermination de f_1 suffit à connaître la grandeur A en tout point et à tout instant (au premier ordre). Considérons par exemple la densité de particule (en nombre) en un point \mathbf{r} au temps t : $n(\mathbf{r}, t)$. Son champ microscopique est modélisé par une somme de Dirac centrés sur chaque particule : $\sum_{i=1}^N \delta(\mathbf{r} - \mathbf{q}_i)$ et se décompose donc de la façon suivante (cf. équation (2.1.11)) :

$$a_{\mathbf{r}}^{(0)} = 0 \quad \text{et} \quad a_{\mathbf{r}}^{(1)}(\mathbf{p}_i, \mathbf{q}_i) = \delta(\mathbf{r} - \mathbf{q}_i). \quad (2.1.16)$$

L'équation (2.1.15) fournit alors la relation fondamentale :

$$\begin{aligned} n(\mathbf{r}, t) &= \int_{\Gamma} \delta(\mathbf{r} - \mathbf{q}_1) f(\mathbf{p}_1, \mathbf{q}_1, t) d\Gamma_1 \\ &= \int_{\mathbb{R}^3} f(\mathbf{p}_1, \mathbf{r}, t) d\mathbf{p}_1. \end{aligned} \quad (2.1.17)$$

La densité de particule en nombre $n(\mathbf{r}, t)$ peut donc être vue comme la densité de probabilité à 1 particule marginalisée sur l'espace des impulsions.

Notons qu'il est possible de généraliser la notion de densité de probabilité à j particules

$$f_j(\mathbf{p}_1, \dots, \mathbf{p}_j, \mathbf{q}_1, \dots, \mathbf{q}_j, t) \equiv \frac{N!}{(N-j)!} \int_{\bar{\Gamma}_j} f_N(\mathbf{p}, \mathbf{q}, t) d\bar{\Gamma}_j. \quad (2.1.18)$$

Le champ à grande échelle $A(\mathbf{r}, t)$ peut alors être décomposée de la façon suivante

$$A(\mathbf{r}, t) = a_{\mathbf{r}}^{(0)} + \sum_{j=1}^k \int_{\Gamma_j} a_{\mathbf{r}}^{(j)}(\mathbf{p}_1, \dots, \mathbf{p}_j, \mathbf{q}_1, \dots, \mathbf{q}_j) f_j(\mathbf{p}_1, \dots, \mathbf{p}_j, \mathbf{q}_1, \dots, \mathbf{q}_j, t) d\Gamma_j \quad (2.1.19)$$

2. La dénomination "densité de probabilité" est quelque peu abusive ici puisque $\int_{\Gamma_1} f_1(\mathbf{p}_1, \mathbf{q}_1, t) d\Gamma_1 = N \int_{\Gamma} f_N(\mathbf{p}, \mathbf{q}, t) d\Gamma = N \neq 1$.

à l'ordre k .

2.1.6 Forme explicite de l'équation de Liouville

Commençons par remarquer que, dans le cas de N particules indiscernables, il est souvent possible de décomposer le Hamiltonien $\mathcal{H}(\mathbf{p}, \mathbf{q})$ en deux termes d'énergie cinétique

$$K(\mathbf{p}) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} = \frac{\mathbf{p}^2}{2m} \quad (2.1.20)$$

(avec m la masse d'une particule) et d'énergie potentielle

$$\Phi(\mathbf{q}) = \sum_{i=1}^N \Phi^{\text{ext}}(\mathbf{q}_i) + \sum_{i=1, j < i}^N \Phi^{\text{int}}(\mathbf{q}_i, \mathbf{q}_j) + \dots \quad (2.1.21)$$

Ici Φ^{ext} désigne le potentiel extérieur agissant sur chacune des particules i et Φ^{int} le potentiel interne généré par i agissant sur j . Les pointillés rendent compte des termes suivants : l'action simultanée de deux particules i et j sur une troisième k et ainsi de suite. En se tenant à l'approximation effectuée plus haut, on ne conservera par la suite que la contribution du champ extérieur et des particules prises individuellement. On pourra alors rendre explicite les équations de Hamilton (2.1.1) :

$$\dot{\mathbf{q}}_i = \frac{\mathbf{p}_i}{m} \quad \text{et} \quad \dot{\mathbf{p}}_i = -\frac{\partial \Phi^{\text{ext}}}{\partial \mathbf{q}_i}(\mathbf{q}_i) - \sum_{j \neq i}^N \frac{\partial \Phi^{\text{int}}}{\partial \mathbf{q}_i}(\mathbf{q}_i, \mathbf{q}_j) \quad (2.1.22)$$

en remarquant que $\Phi^{\text{int}}(\mathbf{q}_i, \mathbf{q}_j) = \Phi^{\text{int}}(\mathbf{q}_j, \mathbf{q}_i)$. L'équation de Liouville (2.1.8) peut également être développée

$$\frac{\partial f_N}{\partial t} + \frac{\mathbf{p}}{m} \cdot \frac{\partial f_N}{\partial \mathbf{q}} + \sum_{i=1}^N \mathbf{F}_i \cdot \frac{\partial f_N}{\partial \mathbf{p}_i} = - \sum_{i=1, j \neq i}^N \mathbf{F}_{ij} \cdot \frac{\partial f_N}{\partial \mathbf{p}_i} \quad (2.1.23)$$

en introduisant \mathbf{F}_i la force externe subie par i et \mathbf{F}_{ij} la force exercée par i sur j .

2.1.7 La hiérarchie BBGKY

De façon cruciale, la distribution à 1 particule, f_1 , n'est généralement pas solution de la même équation que f_N . En intégrant par partie la forme explicite de l'équation de Liouville (2.1.23), il est possible de montrer que (BORN et GREEN, 1946 ; KIRKWOOD, 1946) :

$$\frac{\partial f_1}{\partial t} + \frac{\mathbf{p}_1}{m} \cdot \frac{\partial f_1}{\partial \mathbf{q}_1} + \mathbf{F}_1 \cdot \frac{\partial f_1}{\partial \mathbf{p}_1} = - \int \mathbf{F}_{12} \cdot \frac{\partial f_2}{\partial \mathbf{p}_1} d\mathbf{p}_2 d\mathbf{q}_2 \equiv \hat{\mathcal{J}}[f_2], \quad (2.1.24)$$

où $\hat{\mathcal{J}}[f_2]$ désigne un opérateur faisant intervenir les interactions des particules 2 à 2. La détermination de f_1 nécessite ainsi une connaissance préalable de la fonction de distribution à 2 particules, f_2 . Cependant, f_2 est elle-même solution d'une équation faisant intervenir f_3 et plus généralement, il est possible de montrer que f_j est solution d'une équation dépendant de f_{j+1} . Ainsi, aussi intéressante que soit l'obtention de f_1 , elle semble au premier abord impossible sans celle de f_N et donc sans une résolution complète de l'équation de Liouville.

L'intérêt de l'équation (2.1.24) réside avant tout dans l'appel indirect aux $f_{j>2}$. Elle fait apparaître une hiérarchie parmi les distributions, dans le sens où f_{j+1} peut être vu comme un terme correctif pour la détermination de f_j et ainsi de suite. Il est donc possible de tronquer cette série à un certain ordre k , en proposant un modèle du terme de droite ne dépendant que de f_k et non plus de f_{k+1} . L'erreur induite sur les distributions $f_{j < k}$ d'ordre inférieur sera alors minimale.

2.1.8 L'Équation de Boltzmann

L'approximation la plus courante pour résoudre cette équation est celle dite du *chaos moléculaire* (MAXWELL, 1867), qui propose une troncature dès l'ordre $j = 1$. En supposant que le terme d'interaction se réduit à une description des collisions entre molécules prises 2 à 2, il s'agit de considérer que leurs impulsions ne sont pas corrélées avant collision. Le terme de droite fait alors intervenir un opérateur modélisant ces collisions (on le notera \widehat{C}) et qui n'agit que sur f_1 (deux fois pour les deux particules entrant en collision). L'équation résultante est appelée *équation de Boltzmann* :

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m} \cdot \frac{\partial f}{\partial \mathbf{q}} + \mathbf{F} \cdot \frac{\partial f}{\partial \mathbf{p}} = \widehat{C} [f, f]. \quad (2.1.25)$$

On se permettra d'omettre l'indice 1 sur la distribution $f_1 \rightarrow f$ ainsi que sur les coordonnées $(\mathbf{p}_1, \mathbf{q}_1) \rightarrow (\mathbf{p}, \mathbf{q})$, celles-ci étant les seules à intervenir à partir d'ici (\mathbf{F} est alors la force externe subie par une particule quelconque en \mathbf{q}). Sous les hypothèses évoquées plus haut et en supposant les collisions réversibles à l'échelle microscopique, il est possible de donner une forme explicite à cet opérateur (BOLTZMANN, 1872) :

$$\widehat{C} [f, f](\mathbf{p}, \mathbf{q}, t) \equiv \int d\Omega \int \mathcal{B}(|\mathbf{p} - \mathbf{p}_*|, \Omega) [f^c f_*^c - f f_*] d\mathbf{p}_* \quad (2.1.26)$$

avec \mathbf{p}_* l'impulsion de la particule entrant en collision en \mathbf{q} à l'instant t , Ω l'angle solide de la collision, \mathcal{B} le noyau de collision de Boltzmann et où l'on a introduit les notations :

$$\begin{aligned} f &= f(\mathbf{p}, \mathbf{q}, t) \Big|_{\text{avant collision}} & \text{et} & \quad f_* = f(\mathbf{p}_*, \mathbf{q}, t) \Big|_{\text{avant collision}} \\ f^c &= f(\mathbf{p}, \mathbf{q}, t) \Big|_{\text{après collision}} & \text{et} & \quad f_*^c = f(\mathbf{p}_*, \mathbf{q}, t) \Big|_{\text{après collision}} \end{aligned}$$

Il est à noter que l'équation simplifiée (2.1.25) n'admet pas de solution évidente sous sa forme la plus générale. On pourra néanmoins en étudier plusieurs cas limites dans la suite.

2.1.9 Invariant collisionnel

Considérons tout d'abord l'opérateur suivant :

$$\widehat{I}[\varphi](\mathbf{q}, t) \equiv \int \varphi(\mathbf{p}) \widehat{C} [f, f](\mathbf{p}, \mathbf{q}, t) d\mathbf{p}. \quad (2.1.27)$$

qui agit sur un champ microscopique φ quelconque de l'espace des impulsions. Remplaçons \widehat{C} par sa définition (2.1.26), et considérons les trois transformations sous l'intégrale :

$$\begin{aligned} \mathbf{p} \leftrightarrow \mathbf{p}_* \quad \text{et} \quad \mathbf{p}^c \leftrightarrow \mathbf{p}_*^c & \quad [\text{Inversion des 2 particules}] \\ \mathbf{p} \leftrightarrow \mathbf{p}^c \quad \text{et} \quad \mathbf{p}_* \leftrightarrow \mathbf{p}_*^c & \quad [\text{Retournement du temps}] \\ \mathbf{p} \leftrightarrow \mathbf{p}_*^c \quad \text{et} \quad \mathbf{p}^c \leftrightarrow \mathbf{p}_* & \quad [\text{Inversion des 2 particules et retournement du temps}]. \end{aligned}$$

Le noyau de collision \mathcal{B} est seulement fonction de $|\mathbf{p} - \mathbf{p}_*|$ et Ω , et donc laissé complètement invariant par ces transformations. Le terme $f^c f_*^c - f f_*$, quant à lui, change de signe dans les deux derniers cas. Ainsi, la définition (2.1.27) peut être réécrite des trois façons suivantes :

$$\widehat{I}[\varphi](\mathbf{q}, t) \equiv \int +\varphi(\mathbf{p}_*) \widehat{C} [f, f](\mathbf{p}, \mathbf{q}, t) d\mathbf{p} \quad (2.1.28)$$

$$\widehat{I}[\varphi](\mathbf{q}, t) \equiv \int -\varphi(\mathbf{p}^c) \widehat{C} [f, f](\mathbf{p}, \mathbf{q}, t) d\mathbf{p} \quad (2.1.29)$$

$$\widehat{I}[\varphi](\mathbf{q}, t) \equiv \int -\varphi(\mathbf{p}_*^c) \widehat{C} [f, f](\mathbf{p}, \mathbf{q}, t) d\mathbf{p} \quad (2.1.30)$$

On peut combiner toutes ses définitions :

$$\widehat{\mathcal{I}}[\varphi](\mathbf{q}, t) \equiv \frac{1}{4} \int [\varphi(\mathbf{p}) + \varphi(\mathbf{p}_*) - \varphi(\mathbf{p}^c) - \varphi(\mathbf{p}_*^c)] \widehat{\mathcal{C}}[f, f](\mathbf{p}, \mathbf{q}, t) d\mathbf{p}. \quad (2.1.31)$$

Ainsi si la quantité φ est un *invariant collisionnel*, dans le sens où :

$$\forall(\mathbf{p}, \mathbf{p}_*), \quad \varphi(\mathbf{p}) + \varphi(\mathbf{p}_*) = \varphi(\mathbf{p}^c) + \varphi(\mathbf{p}_*^c), \quad (2.1.32)$$

alors nécessairement $\widehat{\mathcal{I}}[\varphi] = 0$. On notera cependant que la réciproque n'est pas systématiquement vérifiée ; f pourrait être telle que $\widehat{\mathcal{C}}[f, f] = 0$.

2.1.10 Le théorème H

Introduisons à présent une fonction H , telle que³ :

$$H(t) \equiv \int f(\mathbf{p}, \mathbf{q}, t) \ln [f(\mathbf{p}, \mathbf{q}, t)] d\mathbf{p}d\mathbf{q}. \quad (2.1.33)$$

Par intégration par partie, et en supposant f solution de l'équation de Boltzmann (2.1.25), on montre que :

$$\frac{dH}{dt} = \int \widehat{\mathcal{I}}[\ln f](\mathbf{q}, t) d\mathbf{q}. \quad (2.1.34)$$

En utilisant la définition (2.1.31), cette dérivée peut être réécrite

$$\frac{dH}{dt} = \int d\Omega \int \mathcal{B}(|\mathbf{p} - \mathbf{p}_*|, \Omega) g(f^c f_*^c, f f_*) d\mathbf{p}_* d\mathbf{p}d\mathbf{q} \quad (2.1.35)$$

avec $g(x, y) \equiv (x - y) \ln(y/x)$, une fonction qui vérifie $\forall(x, y), g(x, y) \leq 0$. Puisque le noyau de collision \mathcal{B} est strictement positif, la fonction H n'évolue que dans un sens, et ce quelque soit l'état du fluide (BOLTZMANN, 1872) :

$$\frac{dH}{dt} \leq 0. \quad (2.1.36)$$

Il est possible de sentir le lien de cette fonction avec l'entropie totale du système S , dont la valeur ne fait qu'augmenter jusqu'à l'équilibre. On peut montrer que ces deux grandeurs sont en réalité liées par une relation affine : $S = -kH + \text{cste}$ (cf. Annexe A.1) avec k la constante de Boltzmann. On insistera cependant sur le fait que ce résultat présuppose que f est solution de l'équation de Boltzmann et ne constitue pas un principe inviolable. En particulier, il est possible que H (et S) admettent des variations de sens contraires sur des échelles qui ne satisfont pas aux hypothèses dont résulte l'équation (2.1.25). Dans le cas où ces hypothèses sont satisfaites cependant, l'inégalité (2.1.36) atteint l'égalité sur des temps très longs, dans la limite où

$$g(f^c f_*^c, f f_*) = 0 \quad \Rightarrow \quad f^c f_*^c = f f_* \quad (2.1.37)$$

On voit alors que $\ln f$ est un invariant collisionnel, puisque la quantité $\ln f + \ln f_*$ est conservée lors de la collision. Il est important de noter que cette propriété d'invariance à elle seule impose $\widehat{\mathcal{C}}[f, f] = 0$, mais $f^c f_*^c = f f_*$ n'implique en rien que f soit solution de l'équation de Boltzmann et donc que $\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m} \cdot \frac{\partial f}{\partial \mathbf{q}} + \mathbf{F} \cdot \frac{\partial f}{\partial \mathbf{p}} = 0$.

2.1.11 La distribution localement maxwellienne

Commençons par chercher une forme générale pour une distribution f vérifiant l'équation (2.1.37). Dans le cas de collisions élastiques, on peut constater les conservations suivantes lors

3. On remarque que l'argument du logarithme est dimensionné puisque $f d\mathbf{p}d\mathbf{q}$ est sans dimension. On peut cependant imaginer f exprimée en unités de Planck dans le logarithme, et donc comprendre $\ln f$ comme un $\ln(h^3 f)$ à chaque fois qu'il est fait mention de cette quantité.

d'une collision :

$$1 + 1 = 1 + 1 \quad (2.1.38)$$

$$\mathbf{p}_*^c + \mathbf{p}^c = \mathbf{p}_* + \mathbf{p} \quad (2.1.39)$$

$$(\mathbf{p}_*^2)^c + (\mathbf{p}^2)^c = \mathbf{p}_*^2 + \mathbf{p}^2. \quad (2.1.40)$$

Ainsi, il est facile de vérifier que la décomposition

$$\ln f(\mathbf{p}, \mathbf{q}, t) = \alpha(\mathbf{q}, t) + \mathbf{p} \cdot \boldsymbol{\beta}(\mathbf{q}, t) + \mathbf{p}^2 \gamma(\mathbf{q}, t) \quad (2.1.41)$$

donne bien un invariant collisionnel. À ce stade, il est ardu de donner un sens intuitif aux fonctions α , $\boldsymbol{\beta}$ et γ que l'on vient d'introduire. On s'efforcera donc de les exprimer en fonction de grandeurs macroscopiques plus courantes. On a obtenu auparavant la relation suivante :

$$n(\mathbf{r}, t) = \int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p}. \quad (2.1.42)$$

De la même façon, il est possible de relier la vitesse moyenne⁴ au premier moment de la distribution : $\langle \mathbf{v} \rangle \equiv \langle \dot{\mathbf{q}} \rangle = \langle \mathbf{p}/m \rangle$. Il vient

$$\begin{aligned} \langle \mathbf{v} \rangle(\mathbf{r}, t) &= \frac{\int_{\mathbb{R}^3} \frac{\mathbf{p}}{m} f(\mathbf{p}, \mathbf{q}, t) d\mathbf{p}}{\int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{q}, t) d\mathbf{p}} \\ &= \frac{1}{\rho} \int_{\mathbb{R}^3} \mathbf{p} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p}. \end{aligned} \quad (2.1.43)$$

où l'on a introduit la densité en masse de particule $\rho(\mathbf{r}, t) \equiv mn(\mathbf{r}, t)$. On se permettra par la suite de désigner une moyenne sur l'espace des impulsions $\langle \mathbf{v} \rangle$ par $\bar{\mathbf{v}}$. En décomposant la vitesse en une valeur moyenne et une perturbation $\mathbf{v} = \bar{\mathbf{v}} + \mathbf{v}'$, on peut développer le second moment de la distribution comme suit :

$$\begin{aligned} \int_{\mathbb{R}^3} \mathbf{p}^2 f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} &= m^2 \int_{\mathbb{R}^3} (\bar{\mathbf{v}} + \mathbf{v}')^2 f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} \\ &= m\rho\bar{\mathbf{v}}^2 + 2m^2\bar{\mathbf{v}} \int_{\mathbb{R}^3} \mathbf{v}' f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} + m^2 \int_{\mathbb{R}^3} (\mathbf{v}')^2 f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} \end{aligned} \quad (2.1.44)$$

On peut remarquer que

$$\langle \mathbf{v}' \rangle = \langle \mathbf{v} - \bar{\mathbf{v}} \rangle = \bar{\mathbf{v}} - \bar{\mathbf{v}} = \mathbf{0} \quad (2.1.45)$$

$$\langle (\mathbf{v}')^2 \rangle = \langle (\mathbf{v} - \bar{\mathbf{v}})^2 \rangle \equiv \text{Var}(\mathbf{v}) \quad (2.1.46)$$

pour exprimer les moments de f :

$$\int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} = n \quad (2.1.47)$$

$$\int_{\mathbb{R}^3} \mathbf{p} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} = \rho\bar{\mathbf{v}} \quad (2.1.48)$$

$$\int_{\mathbb{R}^3} \mathbf{p}^2 f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p} = m\rho\bar{\mathbf{v}}^2 + 3\rho kT \quad (2.1.49)$$

où l'on a introduit la température locale $(3/2)kT(\mathbf{r}, t) \equiv (1/2)m\text{Var}(\mathbf{v})(\mathbf{r}, t)$. En calculant explicitement les intégrales figurant à gauche, la reparamétrisation de f par ses premiers moments aboutit à une *distribution localement maxwellienne* :

$$f(\mathbf{v}, \mathbf{r}, t) = n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left(-\frac{m(\mathbf{v} - \bar{\mathbf{v}})^2}{2kT} \right) \quad (2.1.50)$$

4. On fait référence à une moyenne sur l'espace des impulsions et non pas l'espace physique.

où $n(\mathbf{r}, t)$, $\bar{\mathbf{v}}(\mathbf{r}, t)$ et $T(\mathbf{r}, t)$ sont des grandeurs qui dépendent de la position et du temps⁵.

2.1.12 La limite maxwellienne

Puisque la distribution localement maxwellienne vérifie l'égalité (2.1.37), il peut être naturel de penser que le théorème H nous dit que la distribution $f(\mathbf{v}, \mathbf{r}, t)$ atteint une distribution localement maxwellienne sur des temps très longs (auquel cas $dH/dt = 0$). Si une telle distribution limite $f(\mathbf{v}, \mathbf{r}, \infty)$ (donc indépendante du temps) donnée par l'équation (2.1.50) vérifie bien par définition de la maxwellienne locale $\widehat{C}[f, f] = 0$, elle n'est généralement pas solution de l'équation de Boltzmann, en l'occurrence :

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m} \cdot \frac{\partial f}{\partial \mathbf{v}} \neq \widehat{C}[f, f] = 0. \quad (2.1.51)$$

En l'absence de forces externes au fluide ($\mathbf{F} = \mathbf{0}$), une solution évidente de cette équation (appelée équation de Vlasov lorsque le terme collisionnel s'annule) est cependant donnée par une fonction indépendante de l'espace et du temps $f(\mathbf{v}, \mathbf{r}, t) = f(\mathbf{v})$. Cette solution de l'équation de Boltzmann dans la limite où $dH/dt \rightarrow 0$,

$$f(\mathbf{v}) = n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left(-\frac{m(\mathbf{v} - \bar{\mathbf{v}})^2}{2kT} \right) \quad (2.1.52)$$

avec n , $\bar{\mathbf{v}}$ et T qui désignent cette fois-ci des constantes de l'espace et du temps, est appelée *distribution maxwellienne*⁶. Il est possible de montrer que, hors conditions aux limites exceptionnelles, il n'existe pas de solution stationnaire plus générale à l'équation de Boltzmann en l'absence de force externe (VILLANI, 2002), et donc dépendant de l'espace. Elle peut être vue comme un cas limite à $t \rightarrow \infty$, un équilibre atteint par le fluide lorsque la fonction $H(t)$, qui ne peut que décroître, finit par atteindre son minimum. Les grandeurs physiques macroscopiques définies par la formule (2.1.15) sont alors des constantes, elles ne dépendent plus de l'espace ni du temps. Lorsque le fluide atteint cet état, on parle d'*équilibre thermodynamique*.

Si une force externe $\mathbf{F}(\mathbf{r}) = -\partial\Phi^{\text{ext}}/\partial\mathbf{r}$ agit sur le fluide, on peut chercher une distribution plus générale de la forme $f(\mathbf{v})g(\mathbf{r})$. En injectant cette expression dans l'équation de Boltzmann, la condition résultante s'écrit $\ln g(\mathbf{r}) + \Phi^{\text{ext}}(\mathbf{r})/kT = 0$. La densité de probabilité à 1 particule tend alors vers une distribution en équilibre avec le potentiel extérieur :

$$f(\mathbf{v}, \mathbf{r}) = n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left(-\frac{(1/2)m(\mathbf{v} - \bar{\mathbf{v}})^2 + \Phi^{\text{ext}}(\mathbf{r})}{kT} \right). \quad (2.1.53)$$

On remarquera que cette distribution reste stationnaire et qu'elle est localement maxwellienne en introduisant une densité fonction de la position :

$$n(\mathbf{r}) \equiv (N/V) e^{-\Phi^{\text{ext}}(\mathbf{r})/kT}. \quad (2.1.54)$$

où $V \equiv \int d\mathbf{q}$ désigne le volume de l'espace physique.

2.1.13 Développement de Chapman-Enskog

Si la distribution maxwellienne semble être une limite inéluctable pour f sur des temps suffisamment longs, rien ne justifie jusqu'ici l'emploi d'une distribution localement maxwellienne. Pourtant, on utilisera cette approximation pour dériver de nombreuses équations de conservation

5. On a ici exprimé f sur les vitesses \mathbf{v} en utilisant $f(\mathbf{p}, \mathbf{r}, t)d\mathbf{p} = dn(\mathbf{r}, t) = f(\mathbf{v}, \mathbf{r}, t)d\mathbf{v}$.

6. Il est possible de trouver la distribution maxwellienne sans la densité de particule n en facteur, auquel cas il est question de distribution normalisée puisque $\int f(\mathbf{v})d\mathbf{v} = \int f(\mathbf{p})d\mathbf{p} = n$. Cette distribution normalisée $f(\mathbf{v})/n$ définit bien une densité de probabilité sur l'espace des vitesses, contrairement à la définition (2.1.52) utilisée dans ce manuscrit. À noter que la densité de probabilité sur l'espace des phases s'obtient en remplaçant le facteur n par $n/N \equiv 1/V$ puisque $\int f(\mathbf{p})d\mathbf{p}d\mathbf{q} = N$.

par la suite (cf. paragraphe 2.2.9). Un argument en faveur de son emploi dans un cadre particulier peut être trouvé en redimensionnant le problème physique (VILLANI, 2002). Considérons une distribution f_Λ et un potentiel $\Phi_\Lambda^{\text{ext}}$, variant sur des petites échelles de temps et d'espace. On pourra écrire :

$$f_\Lambda(\mathbf{v}, \Lambda \mathbf{r}, \Lambda t) = f(\mathbf{v}, \mathbf{r}, t) \quad \text{et} \quad \Phi_\Lambda^{\text{ext}}(\Lambda \mathbf{r}) = \Phi^{\text{ext}}(\mathbf{r}). \quad (2.1.55)$$

avec f solution de l'équation de Boltzmann pour le potentiel Φ^{ext} . En remarquant que

$$\frac{\partial f}{\partial t} = \Lambda \frac{\partial f_\Lambda}{\partial t}, \quad \frac{\partial f}{\partial \mathbf{r}} = \Lambda \frac{\partial f_\Lambda}{\partial \mathbf{r}} \quad \text{et} \quad \frac{\partial \Phi^{\text{ext}}}{\partial \mathbf{r}} = \Lambda \frac{\partial \Phi_\Lambda^{\text{ext}}}{\partial \mathbf{r}} \quad (2.1.56)$$

et que $\widehat{C}[f, f] = \widehat{C}[f_\Lambda, f_\Lambda]$ (il ne s'agit pas d'un opérateur différentiel), l'équation que vérifie f_Λ s'écrit :

$$\Lambda \left[\frac{\partial f_\Lambda}{\partial t} + \mathbf{v} \cdot \frac{\partial f_\Lambda}{\partial \mathbf{r}} + \frac{\mathbf{F}_\Lambda}{m} \cdot \frac{\partial f_\Lambda}{\partial \mathbf{v}} \right] = \widehat{C}[f_\Lambda, f_\Lambda] \quad (2.1.57)$$

Cherchons une solution de cette équation par approche perturbative ; on notera :

$$f_\Lambda = f_\Lambda^{(0)} + \Lambda f_\Lambda^{(1)} + \Lambda^2 f_\Lambda^{(2)} + \dots \quad (2.1.58)$$

La définition de \widehat{C} (2.1.26) nous permet d'écrire le développement du terme de droite en puissance de Λ :

$$\begin{aligned} \widehat{C}[f_\Lambda, f_\Lambda] &= \widehat{C}[f_\Lambda^{(0)}, f_\Lambda^{(0)}] + \Lambda \left(\widehat{C}[f_\Lambda^{(1)}, f_\Lambda^{(0)}] + \widehat{C}[f_\Lambda^{(0)}, f_\Lambda^{(1)}] \right) \\ &+ \Lambda^2 \left(\widehat{C}[f_\Lambda^{(2)}, f_\Lambda^{(0)}] + \widehat{C}[f_\Lambda^{(1)}, f_\Lambda^{(1)}] + \widehat{C}[f_\Lambda^{(0)}, f_\Lambda^{(2)}] \right) + \dots \end{aligned} \quad (2.1.59)$$

Pour Λ petit⁷, les équations que satisfont les $f^{(k)}$ prennent donc la forme :

$$0 = \widehat{C}[f_\Lambda^{(0)}, f_\Lambda^{(0)}] \quad (2.1.60)$$

$$\frac{\partial f_\Lambda^{(1)}}{\partial t} + \mathbf{v} \cdot \frac{\partial f_\Lambda^{(1)}}{\partial \mathbf{r}} + \frac{\mathbf{F}_\Lambda}{m} \cdot \frac{\partial f_\Lambda^{(1)}}{\partial \mathbf{v}} = \widehat{C}[f_\Lambda^{(1)}, f_\Lambda^{(0)}] + \widehat{C}[f_\Lambda^{(0)}, f_\Lambda^{(1)}] \quad (2.1.61)$$

$$\frac{\partial f_\Lambda^{(2)}}{\partial t} + \mathbf{v} \cdot \frac{\partial f_\Lambda^{(2)}}{\partial \mathbf{r}} + \frac{\mathbf{F}_\Lambda}{m} \cdot \frac{\partial f_\Lambda^{(2)}}{\partial \mathbf{v}} = \widehat{C}[f_\Lambda^{(2)}, f_\Lambda^{(0)}] + \widehat{C}[f_\Lambda^{(1)}, f_\Lambda^{(1)}] + \widehat{C}[f_\Lambda^{(0)}, f_\Lambda^{(2)}] \quad (2.1.62)$$

⋮

$$\frac{\partial f_\Lambda^{(r)}}{\partial t} + \mathbf{v} \cdot \frac{\partial f_\Lambda^{(r)}}{\partial \mathbf{r}} + \frac{\mathbf{F}_\Lambda}{m} \cdot \frac{\partial f_\Lambda^{(r)}}{\partial \mathbf{v}} = \sum_{r'=0}^r \widehat{C}[f_\Lambda^{(r')}, f_\Lambda^{(r-r')}] \quad (2.1.63)$$

Cette façon de trouver la solution f_Λ de la hiérarchie ainsi définie, dans le sens où $f_\Lambda^{(0)}$ permet de trouver $f_\Lambda^{(1)}$ qui à son tour permet de trouver $f_\Lambda^{(2)}$ etc..., est appelée développement de Chapman-Enskog (CHAPMAN et COWLING, 1970).

2.1.14 L'équilibre thermodynamique local

Ainsi, dans la limite $\Lambda \rightarrow 0$ et donc des petites échelles de temps et d'espace, ce développement donne la distribution f_Λ comme solution de $\widehat{C}[f_\Lambda, f_\Lambda] = 0$. La forme adoptée par cette densité est alors uniquement influencée par les collisions des particules deux à deux, on a vu qu'elle est donc localement maxwellienne (cf. paragraphe 2.1.11). On fera par la suite référence

7. Cette hypothèse demanderait évidemment à être éclaircie. Un adimensionnement plus poussé de l'équation de Boltzmann permet de voir que Λ doit être de l'ordre du nombre de Knudsen Kn , lequel exprime le rapport entre le libre parcours moyen d'une particule du fluide et une longueur caractéristique du problème macroscopique (BIRD, 1994).

à ce cadre particulier comme celui de l'équilibre *thermodynamique local*. Plus généralement, on voit que f_Λ peut être approximée par une maxwellienne locale à l'ordre 0, au-delà duquel on doit commencer à inclure des corrections de l'ordre de Λ . Il est à noter que la première correction portée à $f_\Lambda^{(0)}$ (donnée par l'équation (2.1.50)), $f_\Lambda^{(1)}$ peut être exprimée analytiquement (cf. équation (2.2.45), CHAPMAN et COWLING, 1970). Les ordres supérieurs restent cependant, encore aujourd'hui, hors de portée (VILLANI, 2002).

2.1.15 Bilan : les différentes échelles

En conclusion, ce préambule fait émerger trois échelles au problème, chacune correspondant à un formalisme et des hypothèses propres. Ces échelles sont représentées schématiquement sur la FIGURE 2.1 :

1. *L'échelle microscopique.*

Ici, le fluide peut être comme un agrégat de N particules qui suivent des trajectoires dans un espace des phases caractérisé par les positions et impulsions de chaque composant $(\mathbf{q}_i, \mathbf{p}_i)$. Typiquement, cette description est adaptée sur des échelles de l'ordre du libre parcours moyen dans le fluide. Les grandeurs sont définies à des échelles très locales comme fonctions de ces coordonnées et peuvent très largement dévier de leur valeur moyenne. La distribution des particules dans l'espace des phases est notamment libre d'adopter des forme non spécifiées, n'étant pas soumise à un équilibre collisionnel local ni au transport à grande échelle. Ainsi, une description complète du fluide nécessiterait la connaissance de la position et de l'impulsion de chaque particule le composant.

Ex : À cette échelle, la densité est donnée par le champ discret $n_{\mathbf{r}}(\mathbf{p}, \mathbf{q}, t) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{q}_i(t))$.

2. *L'équilibre thermodynamique local.*

Dans ce cadre intermédiaire, le fluide est dominé par interactions à faibles échelles (les collisions) et n'est que peu influencé par les phénomènes tendant à uniformiser les grandeurs à plus large échelle. Ce cadre est celui de l'hydrodynamique, lequel repose sur une caractérisation statistique du fluide qui n'était pas possible à l'échelle microscopique. Les grandeurs physiques peuvent être décrites par des fonctions de l'espace \mathbf{r} et du temps t dérivant de moyennes sur l'espace des impulsions. La répartition des particules dans cet espace peut être approchée par une distribution localement maxwellienne, dans le sens où les moments comme la moyenne et la variance qui caractérisent la distribution ne sont définis que localement (cf. FIGURE 2.1). Une description exhaustive du fluide à cette échelle ne nécessite plus une connaissance exacte des trajectoires de chaque particule mais seulement de l'évolution des moments (vitesse moyenne $\bar{\mathbf{v}}(\mathbf{r}, t)$, température $T(\mathbf{r}, t)$, ...) de la fonction de distribution dans l'espace au cours du temps.

Ex : À cette échelle, la densité peut-être décrite par le champ physique $n(\mathbf{r}, t) = \int_{\mathbb{R}^3} f(\mathbf{p}, \mathbf{r}, t) d\mathbf{p}$.

3. *L'équilibre thermodynamique.*

On parle d'équilibre thermodynamique lorsque la fonction $H(t)$ (ou l'entropie $S(t)$) a atteint son minimum (respectivement maximum). Il s'agit du cadre de la thermodynamique classique. Les grandeurs physiques sont stationnaires et leur variations spatiales permettent uniquement de satisfaire un équilibre avec un potentiel extérieur. Les caractéristiques de la distribution sont définies et identiques en tous points, on parle de distribution maxwellienne. Une connaissance parfaite d'un fluide dans cet état ne requiert que la valeur de quelques grandeurs (les variables d'états T , V , N par exemple) obtenues par des moyennes sur l'espace physique.

Ex : À cette échelle, la densité est simplement fournie par une fonction des variables d'état $n(\mathbf{r}) \equiv (N/V) e^{-\Phi^{\text{ext}}(\mathbf{r})/kT}$.

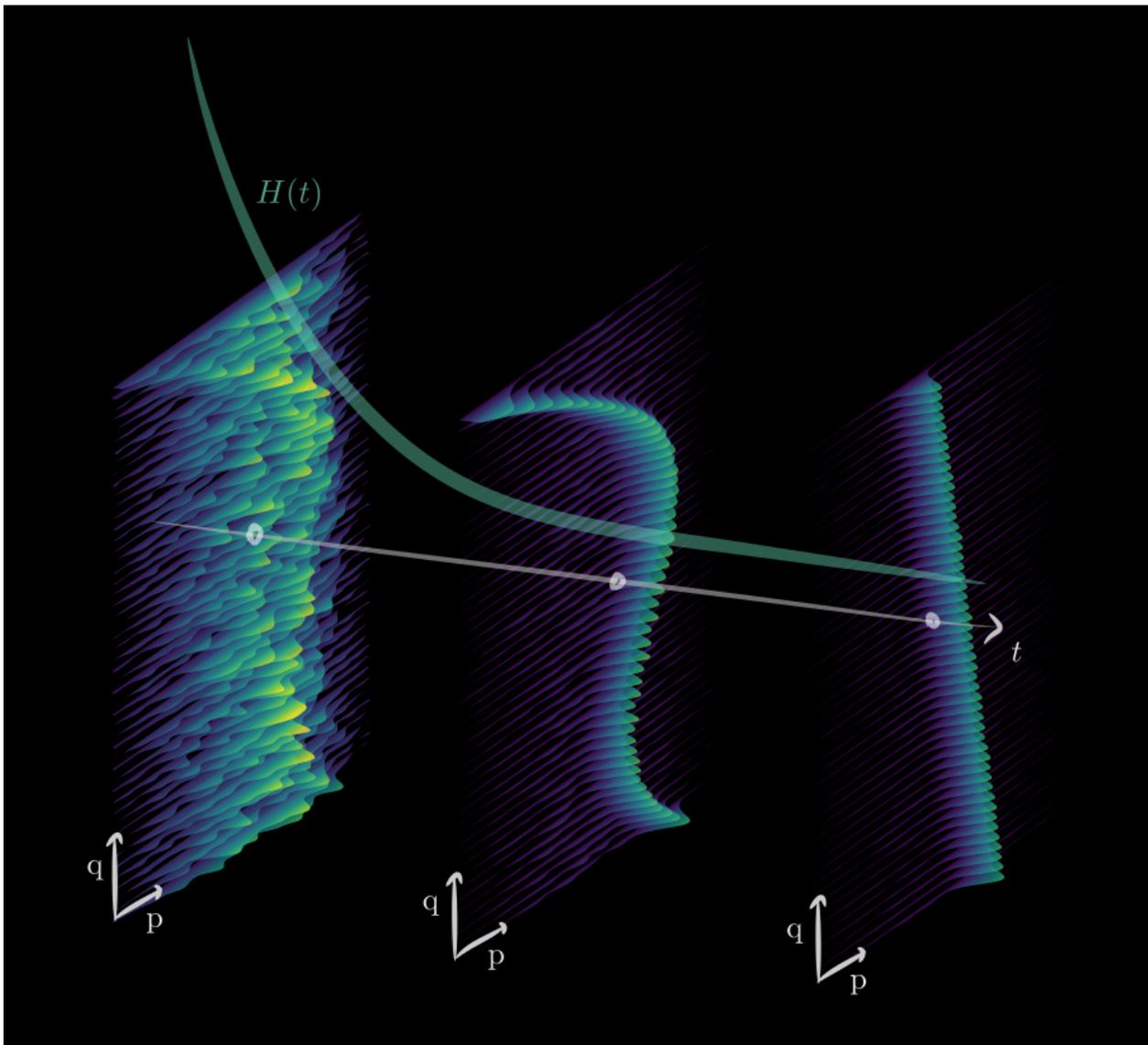


FIGURE 2.1 – Représentation schématique de la densité de probabilité à 1 particule $f(\mathbf{p}, \mathbf{q}, t)$ dans l'espace des phases à 3 instants $t_1 \ll t_2 \ll t_3$.

2.2 * De la théorie cinétique aux équations dynamiques

Contenu de la section

2.2.1	Évolution d'une grandeur moyenne	28
2.2.2	Évolution d'un invariant collisionnel	29
2.2.3	L'équation de continuité	29
2.2.4	L'évolution de la quantité de mouvement	29
2.2.5	Formes conservative et convective, formalismes eulérien et lagrangien . . .	31
2.2.6	La conservation de l'énergie	32
2.2.7	Le transfert radiatif	32
2.2.8	Fermeture des équations dynamiques	35
2.2.9	Un exemple de fermeture : l'équilibre thermodynamique local	36
2.2.10	La fermeture de Navier-Stokes	41

On s'est penché, dans la section 2.1, sur les équations fondamentales qui régissent l'évolution de la fonction de distribution dans l'espace des phases $f(\mathbf{v}, \mathbf{r}, t)$. Si cet objet peut paraître abstrait, son étude nous fournit une bonne compréhension des notions intuitives d'échelle et d'équilibre ainsi que de nombreux outils pour décrire l'évolution des grandeurs physiques. Ici, on va s'intéresser aux relations qui gouvernent ces quantités à l'échelle macroscopique : les équations de l'hydrodynamique.

2.2.1 Évolution d'une grandeur moyenne

Considérons un champ moyen $\bar{\varphi}(\mathbf{r}, t)$ défini localement dans l'espace et dans le temps et décrit microscopiquement par un champ sur l'espace des vitesses $\varphi(\mathbf{v})$ dont la forme ne dépend pas la position \mathbf{r} . À cette échelle, cela équivaut à dire que :

$$\bar{\varphi}(\mathbf{r}, t) \equiv \langle \varphi \rangle(\mathbf{r}, t) \equiv \frac{\int \varphi(\mathbf{v}) f(\mathbf{v}, \mathbf{r}, t) d\mathbf{v}}{\int f(\mathbf{v}, \mathbf{r}, t) d\mathbf{v}} = \frac{1}{n(\mathbf{r}, t)} \int \varphi(\mathbf{v}) f(\mathbf{v}, \mathbf{r}, t) d\mathbf{v} \quad (2.2.1)$$

On supposera systématiquement que $f(\mathbf{v}, \mathbf{r}, t)$ est solution de l'équation de Boltzmann (2.1.25) par la suite. Ainsi, on peut réécrire la définition (2.1.27) de la façon suivante :

$$\begin{aligned} \hat{\Gamma}[\varphi](\mathbf{r}, t) &\equiv \int \varphi(\mathbf{v}) \hat{\mathcal{C}}[f, f](\mathbf{v}, \mathbf{r}, t) d\mathbf{v} = \int \varphi \left[\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m} \cdot \frac{\partial f}{\partial \mathbf{v}} \right] d\mathbf{v} \\ &= \frac{\partial}{\partial t} \int \varphi f d\mathbf{v} + \frac{\partial}{\partial \mathbf{r}} \cdot \int \mathbf{v} \varphi f d\mathbf{v} + \frac{\mathbf{F}}{m} \cdot \int \varphi \frac{\partial f}{\partial \mathbf{v}} d\mathbf{v} \end{aligned} \quad (2.2.2)$$

où l'on a progressivement omis les dépendances pour en alléger l'expression. Puisque la première intégrale fait intervenir le champ moyen $\bar{\varphi}$ défini juste au dessus, l'opérateur $\hat{\Gamma}[\varphi]$ revêt une importance centrale en liant les variations temporelles et spatiales de cette grandeur. On peut décomposer une nouvelle fois la vitesse \mathbf{v} en une moyenne à large échelle, $\bar{\mathbf{v}}$ la vitesse d'ensemble (ou vitesse de convection), et sa composante dans le référentiel comouvant à cette échelle, \mathbf{v}' la vitesse de diffusion. Ces vitesses ont les propriétés suivantes :

$$\bar{\mathbf{v}} \equiv \frac{1}{n} \int \mathbf{v} f d\mathbf{v} \quad \text{et donc} \quad \langle \mathbf{v}' \rangle = \frac{1}{n} \int (\mathbf{v} - \bar{\mathbf{v}}) f d\mathbf{v} = \bar{\mathbf{v}} - \bar{\mathbf{v}} = \mathbf{0}. \quad (2.2.3)$$

La deuxième intégrale de l'équation (2.2.2) se divise donc en deux termes :

$$\int \mathbf{v} \varphi f d\mathbf{v} = n \bar{\varphi} \bar{\mathbf{v}} + n \langle \varphi \mathbf{v}' \rangle \quad (2.2.4)$$

et une intégration par partie permet de réécrire la troisième :

$$\int \varphi \frac{\partial f}{\partial \mathbf{v}} d\mathbf{v} = \int \frac{\partial \varphi f}{\partial \mathbf{v}} d\mathbf{v} - n \left\langle \frac{\partial \varphi}{\partial \mathbf{v}} \right\rangle \quad (2.2.5)$$

Puisque f définit une distribution sur l'espace des vitesses, il est nécessaire qu'elle tende vers 0 à l'infini. On exigera en plus qu'elle le fasse suffisamment vite pour que $\varphi f \xrightarrow{\infty} 0$ pour toute fonction φ "raisonnable"⁸. Le premier terme de l'équation (2.2.5) disparaît donc et l'opérateur $\hat{\Gamma}[\varphi]$ prend la forme :

$$\hat{\Gamma}[\varphi] = \frac{\partial}{\partial t} (n \bar{\varphi}) + \frac{\partial}{\partial \mathbf{r}} \cdot (n \bar{\varphi} \bar{\mathbf{v}} + n \langle \varphi \mathbf{v}' \rangle) - n \frac{\mathbf{F}}{m} \cdot \left\langle \frac{\partial \varphi}{\partial \mathbf{v}} \right\rangle. \quad (2.2.6)$$

8. Cette propriété est notamment nécessaire pour que les moments des vitesses soient définis.

En multipliant cette équation par m , et en introduisant la notation plus courante $\nabla \equiv \partial/\partial \mathbf{r}$, il vient une relation décrivant les variations du champ moyen $\bar{\varphi}$:

$$\frac{\partial}{\partial t}(\rho \bar{\varphi}) + \nabla \cdot (\rho \bar{\varphi} \bar{\mathbf{v}} + \rho \langle \varphi \mathbf{v}' \rangle) = n \mathbf{F} \cdot \left\langle \frac{\partial \varphi}{\partial \mathbf{v}} \right\rangle + m \hat{\Gamma}[\varphi]. \quad (2.2.7)$$

Il peut être difficile de saisir au premier abord à quel point l'équation (2.2.7) (écrite ici sous forme conservative) est fondamentale. Les termes de gauche décrivent l'évolution d'un champ moyen arbitraire, causée d'une part de la dépendance explicite au temps du champ et d'autre part son transport à grande (convection) et petite échelles (diffusion) par les champs de vitesse $\bar{\mathbf{v}}$ et \mathbf{v}' . Cette variation a pour source l'action des forces externes \mathbf{F} (et celle-ci est d'autant plus grande que le champ microscopique dépend de la vitesse) à laquelle s'ajoute les interactions des particules 2 à 2 par collision. La FIGURE 2.2 fournit une représentation schématique de ces effets.

2.2.2 Évolution d'un invariant collisionnel

Si φ se trouve être un invariant collisionnel et donc si

$$\forall(\mathbf{v}, \mathbf{v}_*), \quad \varphi(\mathbf{v}) + \varphi(\mathbf{v}_*) = \varphi(\mathbf{v}^c) + \varphi(\mathbf{v}_*^c), \quad (2.2.8)$$

on a alors montré dans le paragraphe 2.1.9 que $\hat{\Gamma}[\varphi] = 0$. L'équation (2.2.7) ne contient alors plus de source (ou de perte) collisionnelle et s'écrit

$$\frac{\partial}{\partial t}(\rho \bar{\varphi}) + \nabla \cdot (\rho \bar{\varphi} \bar{\mathbf{v}} + \rho \langle \mathbf{v}' \varphi \rangle) = n \mathbf{F} \cdot \left\langle \frac{\partial \varphi}{\partial \mathbf{v}} \right\rangle. \quad (2.2.9)$$

On va voir dans les paragraphes qui suivent que cette équation, appliquée aux invariants collisionnels déjà mentionnés dans la section 2.1, permet de retrouver les équations hydrodynamiques dans toute leur généralité.

2.2.3 L'équation de continuité

Une quantité trivialement invariante par collision est $\varphi(\mathbf{v}) = 1$. Dans ce cas, l'équation (2.2.9) se simplifie considérablement

$$\varphi(\mathbf{v}) = 1 \quad \Rightarrow \quad \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) = 0 \quad (2.2.10)$$

puisque à la fois $\langle \mathbf{v}' \rangle = \mathbf{0}$ et $\partial \varphi / \partial \mathbf{v} = \mathbf{0}$. Cette équation n'est autre que l'équation de continuité, laquelle lie les variations de la masse à son transport à grande échelle⁹. L'absence de terme à droite nous dit qu'il n'y a aucune source locale de masse dans le fluide.

2.2.4 L'évolution de la quantité de mouvement

Si l'on applique maintenant l'équation (2.2.9) à chacune des composantes de la vitesse, on obtient l'équation suivante :

$$\varphi(\mathbf{v}) = \mathbf{v} \quad \Rightarrow \quad \frac{\partial \rho \bar{\mathbf{v}}}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} + \rho \langle \mathbf{v}' \otimes \mathbf{v}' \rangle) = n \mathbf{F} \quad (2.2.11)$$

où \otimes désigne le produit tensoriel. On obtient alors une équation qui décrit l'évolution de la quantité de mouvement au cours du temps. Le flux de cette quantité est composé de deux tenseurs : le premier décrit son transport par le champ de vitesse à grande échelle et le second la contribution

⁹. Notons que la propriété $\langle \mathbf{v}' \varphi \rangle = \mathbf{0}$ nous dit que le transport de masse par diffusion n'a en moyenne aucune influence.

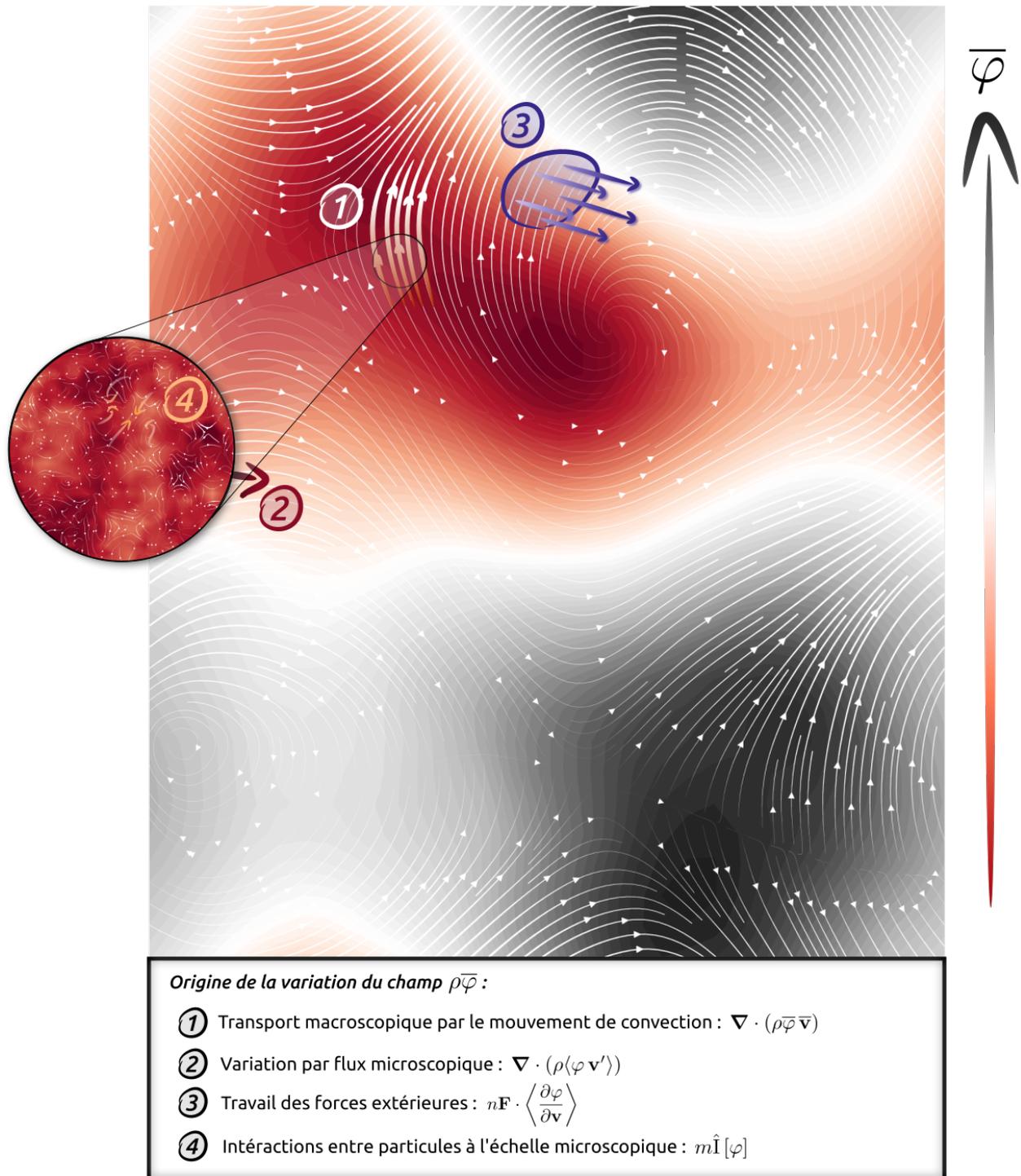


FIGURE 2.2 – Schéma représentant les origines de l'évolution temporelle du champ moyen $\bar{\varphi}$. Les variations de couleurs témoignent des fluctuations (spatiales) d'un champ scalaire quelconque $\bar{\varphi}$ tandis que les lignes de courant illustrent la direction et l'amplitude du champ de vitesse $\bar{\mathbf{v}}$ (respectivement φ et \mathbf{v} pour la loupe).

de la diffusion en moyenne à l'échelle microscopique. On appelle ce second terme tenseur des pressions (LIFSHITZ et PITAEVSKII, 1981) et le note P :

$$\mathbf{P} \equiv \rho\langle\mathbf{v}' \otimes \mathbf{v}'\rangle. \quad (2.2.12)$$

Enfin, le terme de droite permet de voir que les forces externes constituent la seule source

de quantité de mouvement. Notons que, lorsque la contribution des forces externes se limite à la force de gravité $\mathbf{F} = m\mathbf{g}$ (en négligeant notamment les forces électromagnétiques à grandes échelles ainsi que toute force inertielle dans le cas d'un référentiel en mouvement), on trouve alors :

$$\frac{\partial \rho \bar{\mathbf{v}}}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} + \mathbf{P}) = \rho \mathbf{g}. \quad (2.2.13)$$

2.2.5 Formes conservative et convective, formalismes eulérien et lagrangien

Le passage d'une forme conservative à une forme convective (à comprendre comme décrivant l'évolution d'un champ le long du mouvement de convection) se fait à l'aide de l'équation de continuité :

$$\frac{\partial}{\partial t}(\rho \bar{\varphi}) + \nabla \cdot (\rho \bar{\varphi} \bar{\mathbf{v}}) = \rho \left(\frac{\partial \bar{\varphi}}{\partial t} + \bar{\mathbf{v}} \cdot \nabla \bar{\varphi} \right) + \underbrace{\bar{\varphi} \left(\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) \right)}_0 \equiv \rho \frac{D\bar{\varphi}}{Dt}. \quad (2.2.14)$$

La dérivée particulière D/Dt alors introduite peut être vue comme une dérivée temporelle pour le champ $\bar{\varphi}$ dans un formalisme qui décrit une trajectoire de coordonnées $(\mathbf{r}(\mathbf{r}_0, t), t)$ suivant le mouvement de convection, avec $\mathbf{r}(\mathbf{r}_0, t)$ vérifiant :

$$\mathbf{r}(\mathbf{r}_0, t) = \mathbf{r}_0 + \int_0^t \bar{\mathbf{v}}(\mathbf{r}(\mathbf{r}_0, t'), t') dt'. \quad (2.2.15)$$

Le choix de la trajectoire $\mathbf{r}(\mathbf{r}_0, t)$ fait apparaître un formalisme où les grandeurs peuvent s'exprimer comme des fonctions d'une position initiale \mathbf{r}_0 et du temps t que l'on appelle *formalisme lagrangien*. Notons que ce choix de trajectoire est entièrement spécifié par le choix de sa position initiale $\mathbf{r}_0(\mathbf{r}, t)$ ¹⁰ dans le formalisme "classique" (fonction de \mathbf{r} et de t), appelé par opposition *formalisme eulérien*. La correspondance entre les deux représentations est simplement donnée par :

$$\underbrace{\bar{\varphi}(\mathbf{r}, t)}_{\text{Formalisme eulérien}} = \bar{\varphi}(\mathbf{r}(\mathbf{r}_0, t), t) = \underbrace{\bar{\varphi}(\mathbf{r}_0, t)}_{\text{Formalisme lagrangien}} \quad (2.2.16)$$

ce qui donne bien la relation :

$$\left(\frac{\partial \bar{\varphi}}{\partial t} \right)_{\mathbf{r}_0} = \left(\frac{\partial \bar{\varphi}}{\partial t} \right)_{\mathbf{r}} + \left(\frac{\partial \mathbf{r}}{\partial t} \right)_{\mathbf{r}_0} \cdot \left(\frac{\partial \bar{\varphi}}{\partial \mathbf{r}} \right)_{\mathbf{r}} = \frac{\partial \bar{\varphi}}{\partial t} + (\bar{\mathbf{v}} \cdot \nabla) \bar{\varphi} \equiv \frac{D\bar{\varphi}}{Dt} \quad (2.2.17)$$

en vertu de la définition (2.2.15) pour la trajectoire $\mathbf{r}(\mathbf{r}_0, t)$. En prenant les exemples de l'équation de continuité et de la conservation de la quantité de mouvement, les formes convectives s'écrivent simplement :

$$\rho \frac{D}{Dt} \left(\frac{1}{\rho} \right) = \nabla \cdot \bar{\mathbf{v}} \quad (2.2.18)$$

$$\rho \frac{D\bar{\mathbf{v}}}{Dt} = -\nabla \cdot \mathbf{P} + \rho \mathbf{g}. \quad (2.2.19)$$

10. La relation (2.2.15) peut être inversée sous hypothèses de continuité et d'unicité (ARIS, 1990).

2.2.6 La conservation de l'énergie

Appliquée à l'invariant collisionnel $\varphi(\mathbf{v}) = (1/2)\mathbf{v}^2$, l'équation (2.2.9) conduit à une relation un peu plus complexe :

$$\begin{aligned} \varphi(\mathbf{v}) = \frac{1}{2}\mathbf{v}^2 \quad \Rightarrow \quad \frac{\partial}{\partial t} \left(\frac{1}{2}\rho\bar{\mathbf{v}}^2 + \frac{1}{2}\rho\langle\mathbf{v}'\cdot\mathbf{v}'\rangle \right) \\ + \nabla \cdot \left[\left(\frac{1}{2}\rho\bar{\mathbf{v}}^2 + \frac{1}{2}\rho\langle\mathbf{v}'\cdot\mathbf{v}'\rangle \right) \bar{\mathbf{v}} + \rho\langle\mathbf{v}'\otimes\mathbf{v}'\rangle \cdot \bar{\mathbf{v}} + \frac{1}{2}\rho\langle(\mathbf{v}'\cdot\mathbf{v}')\mathbf{v}'\rangle \right] = n\mathbf{F} \cdot \bar{\mathbf{v}}. \end{aligned} \quad (2.2.20)$$

Pour en alléger l'écriture, on pourra introduire l'énergie interne massique ε et le flux de chaleur conductif \mathbf{Q} :

$$\varepsilon \equiv \frac{1}{2}\langle\mathbf{v}'\cdot\mathbf{v}'\rangle \quad (2.2.21)$$

$$\mathbf{Q} \equiv \frac{1}{2}\rho\langle(\mathbf{v}'\cdot\mathbf{v}')\mathbf{v}'\rangle. \quad (2.2.22)$$

La première réfère ainsi à une énergie cinétique "désorganisée" par rapport au champ de vitesse à grande échelle et le second désigne le transport de cette énergie par diffusion. L'équation alors obtenue décrit l'évolution de l'énergie totale $(1/2)\rho\bar{\mathbf{v}}^2 + \rho\varepsilon$ au cours du temps :

$$\frac{\partial}{\partial t} \left(\frac{1}{2}\rho\bar{\mathbf{v}}^2 + \rho\varepsilon \right) + \nabla \cdot \left[\left(\frac{1}{2}\rho\bar{\mathbf{v}}^2 + \rho\varepsilon \right) \bar{\mathbf{v}} + \mathbf{P} \cdot \bar{\mathbf{v}} + \mathbf{Q} \right] = n\mathbf{F} \cdot \bar{\mathbf{v}} \quad (2.2.23)$$

Cette relation fait intervenir le transport engendré par le flux d'énergie à grande échelle (le flux convectif), le travail de la pression et le flux conductif et donne comme source d'énergie la puissance libérée par les forces externes. Dans le cas où les forces externes ne sont constituées que d'une force de gravitation statique $\mathbf{F}(\mathbf{r}) = -m\nabla\Phi(\mathbf{r})$ (où Φ désigne le potentiel de gravitation), il est possible d'exploiter la propriété

$$n\mathbf{F} \cdot \bar{\mathbf{v}} = -\rho\bar{\mathbf{v}} \cdot \nabla\Phi = -\nabla(\rho\Phi\bar{\mathbf{v}}) + \Phi\nabla \cdot (\rho\bar{\mathbf{v}}) = -\nabla(\rho\Phi\bar{\mathbf{v}}) - \frac{\partial\rho\Phi}{\partial t} \quad (2.2.24)$$

pour obtenir la forme complètement conservative :

$$\frac{\partial}{\partial t} \left(\frac{1}{2}\rho\bar{\mathbf{v}}^2 + \rho\varepsilon + \rho\Phi \right) + \nabla \cdot \left[\left(\frac{1}{2}\rho\bar{\mathbf{v}}^2 + \rho\varepsilon + \rho\Phi \right) \bar{\mathbf{v}} + \bar{\mathbf{v}} \cdot \mathbf{P} + \mathbf{Q} \right] = 0 \quad (2.2.25)$$

où l'énergie gravitationnelle $\rho\Phi$ fait désormais partie du bilan interne du système. En remarquant que :

$$\rho \frac{D}{Dt} \left(\frac{\bar{\mathbf{v}}^2}{2} \right) = \rho\bar{\mathbf{v}} \frac{D\bar{\mathbf{v}}}{Dt} = -\bar{\mathbf{v}} \cdot (\nabla \cdot \mathbf{P}) - \rho\bar{\mathbf{v}} \cdot \nabla\Phi = -\bar{\mathbf{v}} \cdot (\nabla \cdot \mathbf{P}) - \rho \frac{D\Phi}{Dt} \quad (2.2.26)$$

d'après l'équation (2.2.19), on soulignera la forme convective particulièrement compacte

$$\rho \frac{D\varepsilon}{Dt} = -\mathbf{P} : (\nabla \otimes \bar{\mathbf{v}}) - \nabla \cdot \mathbf{Q} \quad (2.2.27)$$

en introduisant le produit doublement contracté " : ".

2.2.7 Le transfert radiatif

Il peut sembler curieux (en particulier dans un manuscrit qui s'apprête à parler de physique stellaire) que l'équation (2.2.23) ne fasse qu'intervenir les flux convectif et conductif sans mentionner le flux radiatif. Dans les faits, cette contribution est bien présente mais notre cadre de travail l'a écartée avant même la dérivation de l'équation de Boltzmann. Dans la section 2.1, on a en effet supposé un fluide composé de particules identiques et toutes les grandeurs intervenant

jusqu'ici résultent des seules interactions de ces particules entre elles. Pour tenir compte de la radiation dans nos équations, il faut faire intervenir un nouveau type de particules : les photons.

De façon remarquable, les photons peuvent être traités dans le cadre de la théorie cinétique (la théorie introduite dans la section 2.1) de manière analogue aux particules du fluide. On parle dans ce cas de gaz de photons. En particulier, il est possible de leur associer une distribution de probabilité dans l'espace et en impulsion pour une énergie $h\nu$ donnée (h désigne ici la constante de Planck et ν la fréquence du photon), f_ν , laquelle est solution d'une équation de Boltzmann. Les conventions utilisées pour exprimer cette équation diffère cependant quelque peu du cas de la particule de fluide, la relation entre vitesse \mathbf{v} et impulsion \mathbf{p} du photon différant de $\mathbf{p}/m = \mathbf{v}$ puisque le photon n'a pas de masse. Ainsi, plutôt que de définir la densité de probabilité de la façon suivante $dN(t) \equiv f(\mathbf{p}, \mathbf{q}, t) d\mathbf{p} d\mathbf{q}$, on utilisera (POMRANING, 1973) :¹¹

$$dN^R(t) \equiv f^R(\nu, \boldsymbol{\Omega}, \mathbf{r}, t) d\boldsymbol{\Omega} d\nu d\mathbf{r} \quad (2.2.28)$$

avec la direction de propagation $\boldsymbol{\Omega}$ (caractérisée par l'angle solide Ω), laquelle fixe l'impulsion du photon pour une fréquence ν donnée $\mathbf{p} = (h\nu/c)\boldsymbol{\Omega}$. Les différentes dépendances peuvent être intuitées de la façon suivante : $d\mathbf{r} = c dt dS$ est l'élément de volume contenant les photons s'appêtant à traverser l'élément de surface dS (normal à $\boldsymbol{\Omega}$) durant l'instant dt puisque leur vitesse c est constante. Le terme $d\boldsymbol{\Omega} d\nu$ peut être vu, quant à lui, comme une description "en coordonnées sphériques" du domaine d'impulsion $d\mathbf{p}$, $d\nu$ caractérisant la norme du vecteur et $d\boldsymbol{\Omega}$ son orientation. L'équation de Boltzmann associée s'écrit (POMRANING, 1973) :

$$\frac{\partial f^R}{\partial t} + c\boldsymbol{\Omega} \cdot \nabla f^R = \rho q - \rho c \kappa_a f^R + \hat{C}^R [f^R, f^R]. \quad (2.2.29)$$

On peut noter quelques différences avec l'équation (2.1.25). Premièrement, aucune force externe ne s'applique au gaz de photons ; ceux-ci n'ont pas de masse. Deuxièmement, leur nombre total $N^R(t) = \int f^R d\boldsymbol{\Omega} d\nu d\mathbf{r}$ n'est pas conservé au cours du temps, contrairement au nombre de particules N du fluide. C'est ce que traduisent les quantités $q(\nu, \mathbf{r}, t)$, le taux d'émission spontané du milieu par unité de masse et $\kappa_a(\nu, \mathbf{r}, t)$, le taux d'absorption du milieu par unité de masse ou *opacité*. L'opérateur \hat{C}^R décrit, quant à lui, les interactions des photons avec la matière par diffusion et peut être défini comme

$$\hat{C}^R [f^R, f^R] \equiv \rho \int \left[\mathcal{K}(\nu_* \rightarrow \nu, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}_*) \kappa_s f_*^R - \mathcal{K}(\nu \rightarrow \nu_*, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}_*) \kappa_s f^R \right] d\nu_* d\boldsymbol{\Omega}_* \quad (2.2.30)$$

en réutilisant les conventions dans l'intégrale :

$$\kappa_s f_*^R \equiv \kappa_s(\nu_*, \mathbf{r}, t) f^R(\nu_*, \boldsymbol{\Omega}_*, \mathbf{r}, t) \quad \text{et} \quad \kappa_s f^R \equiv \kappa_s(\nu, \mathbf{r}, t) f^R(\nu, \boldsymbol{\Omega}, \mathbf{r}, t). \quad (2.2.31)$$

Ici, $\kappa_s(\nu_*, \mathbf{r}, t)$ désigne une opacité liée à la diffusion d'un photon par la matière et $\mathcal{K}(\nu \rightarrow \nu_*, \boldsymbol{\Omega} \cdot \boldsymbol{\Omega}_*)$ peut être vu comme la probabilité que ce photon (de fréquence ν et de direction $\boldsymbol{\Omega}$) obtienne une fréquence ν_* et soit dévié d'un angle $\boldsymbol{\Omega} \cdot \boldsymbol{\Omega}_*$ après interaction avec le milieu.

Tout comme pour le fluide, il est possible de lier les moments de la distribution sur l'espace des impulsions à des quantités physiques. On définira notamment :

$$E^R(\mathbf{r}, t) \equiv \int h\nu f^R(\nu, \boldsymbol{\Omega}, \mathbf{r}, t) d\boldsymbol{\Omega} d\nu \quad (2.2.32)$$

$$\mathbf{Q}^R(\mathbf{r}, t) \equiv \int h\nu c \boldsymbol{\Omega} f^R(\nu, \boldsymbol{\Omega}, \mathbf{r}, t) d\boldsymbol{\Omega} d\nu \quad (2.2.33)$$

$$\mathbf{P}^R(\mathbf{r}, t) \equiv \int h\nu \boldsymbol{\Omega} \otimes \boldsymbol{\Omega} f^R(\nu, \boldsymbol{\Omega}, \mathbf{r}, t) d\boldsymbol{\Omega} d\nu \quad (2.2.34)$$

respectivement les énergie, flux et tenseur des pressions radiatifs. Encore une fois, à l'instar des moments dérivés pour le fluide, ces grandeurs sont liées entre elles par intégration de l'équation

11. On utilisera dans la suite l'exposant R pour caractériser toute grandeur se rapportant à la radiation.

de Boltzmann sur f^R . Par exemple, les intégrations $\int hv$ (2.2.29) $d\Omega dv$ et $\int (hv/c) \Omega$ (2.2.29) $d\Omega dv$ donnent :

$$\frac{\partial E^R}{\partial t} + \nabla \cdot \mathbf{Q}^R = \rho \Sigma - \rho c \bar{\kappa}_a E^R + \int hv \hat{C}^R [f^R, f^R] d\Omega dv \quad (2.2.35)$$

$$\frac{1}{c^2} \frac{\partial \mathbf{Q}^R}{\partial t} + \nabla \cdot \mathbf{P}^R = -\rho \overline{\Omega \kappa_a} E^R + \int (hv/c) \Omega \hat{C}^R [f^R, f^R] d\Omega dv \quad (2.2.36)$$

avec $\Sigma = 4\pi \int hv q(v) dv$ et en introduisant la convention :

$$\bar{\varphi}(\mathbf{r}, t) \equiv \frac{\int hv \varphi(v, \mathbf{r}, t) f^R(v, \Omega, \mathbf{r}, t) d\Omega dv}{\int hv f^R(v, \Omega, \mathbf{r}, t) d\Omega dv} = \frac{1}{E^R} \int hv \varphi f^R(v, \Omega, \mathbf{r}, t) d\Omega dv \quad (2.2.37)$$

pour une grandeur se rapportant à la radiation. On notera que la première équation (2.2.35) porte sur l'énergie et la seconde (2.2.36) sur la quantité de mouvement. À ce stade, il est clair que l'étude du fluide ne peut se indépendamment de celle du gaz de photon : les gains d'énergie et de quantité de mouvement pour fluide constituent des pertes d'énergie et d'impulsion pour les photons. De l'autre côté, les propriétés d'absorption, d'émission et de diffusion du milieu qui régissent l'équation de Boltzmann radiative¹² (2.2.29) sont intimement liées à l'état du fluide, notamment au travers des taux d'émission et opacités du milieu.

Le couplage des équations radiatives et hydrodynamiques (qui définit un pan entier de la physique : l'hydrodynamique radiative) doit cependant se faire avec précaution. En effet, alors que les interactions matière-rayonnement (termes de droite dans l'équation de Boltzmann) se décrivent plus facilement dans un référentiel qui suit le mouvement du fluide, les opérateurs différentiels s'appliquant aux quantités physiques (termes de gauche) sont généralement écrit dans un référentiel inertiel. Ceci peut mener à des inconsistances, en particulier dans les fluides se mouvant à grandes vitesses, avec l'apparition de termes en $\mathcal{O}(\bar{\mathbf{v}}/c)$ suite à la transformation de Lorentz (THOMAS, 1930). On peut montrer formellement que les effets Doppler et d'aberration additionnels ne peuvent être négligés si l'on veut maintenir la cohérence du formalisme (CASTOR, 1972; BUCHLER, 1979). Pour l'instant, on supposera que les termes de droites ont été calculés dans le référentiel inertiel sans chercher à les expliciter (on revient sur cette hypothèse dans le paragraphe 2.2.9). On gardera seulement à l'esprit que les moments de la distribution f^R diffèrent quelque peu de ce que l'on entend intuitivement par "énergie radiative" ou "flux radiatif". De cette façon, les équations de l'hydrodynamique (2.2.19) et (2.2.27) peuvent être corrigées en tenant compte des gains et pertes de quantité de mouvement et d'énergie du gaz de photons :

$$\rho \frac{D\bar{\mathbf{v}}}{Dt} + \nabla \cdot \mathbf{P} + \rho \nabla \Phi = - \left(\rho \Sigma - \rho c \bar{\kappa}_a E^R + \int hv \hat{C}^R [f^R, f^R] d\Omega dv \right) \quad (2.2.38)$$

$$\rho \frac{D\boldsymbol{\varepsilon}}{Dt} + \mathbf{P} : (\nabla \otimes \bar{\mathbf{v}}) + \nabla \cdot \mathbf{Q} = - \left(-\rho \overline{\Omega \kappa_a} E^R + \int (hv/c) \Omega \hat{C}^R [f^R, f^R] d\Omega dv \right) \quad (2.2.39)$$

12. Appelée aussi équation du transfert radiatif.

On notera que l'équation de continuité reste inchangée. En utilisant les relations (2.2.35) et (2.2.36), il est possible d'écrire sous forme conservative¹³ :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) = 0 \quad (2.2.40)$$

$$\frac{\partial}{\partial t} \left(\rho \bar{\mathbf{v}} + \frac{\mathbf{Q}^R}{c^2} \right) + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} + \mathbf{P} + \mathbf{P}^R) = -\rho \nabla \Phi \quad (2.2.41)$$

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + E^R + \rho \Phi \right) + \nabla \cdot \left[\left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi \right) \bar{\mathbf{v}} + \mathbf{P} \cdot \bar{\mathbf{v}} + \mathbf{Q} + \mathbf{Q}^R \right] = 0. \quad (2.2.42)$$

Ces équations décrivent la conservation de la masse ainsi que l'évolution de la quantité de mouvement et de l'énergie dans un fluide soumis aux radiations (dans un cadre non relativiste). Elles sont donc d'un intérêt central dans le cadre de la physique stellaire.

2.2.8 Fermeture des équations dynamiques

Les équations (2.2.40)-(2.2.42) étant au nombre de 3, il semble raisonnable chercher à ce qu'elles ne dépendent que de 3 variables en vue de les résoudre (le choix des conditions aux limites et donc le problème de Cauchy dans son ensemble ne sera pas abordé dans cette section). Il est clair que celles-ci font pourtant intervenir bien plus d'inconnues. Pour éclaircir ce point, on peut remarquer que les résultats obtenus jusqu'ici ont été dérivés pour une distribution f solution de l'équation de Boltzmann (2.1.25) sans faire aucune hypothèse supplémentaire quant à sa forme. On a cependant vu dans la section 2.1, qu'aux échelles de temps et d'espace qui nous intéressent en hydrodynamique, la densité de probabilité pouvait être approchée par un développement

$$f = f^{(0)} + f^{(1)} + f^{(2)} + \dots \quad (2.2.43)$$

respectant la hiérarchie : $f^{(0)} \gg f^{(1)} \gg f^{(2)} \gg \dots$

De façon cruciale, ce développement met en lumière que la distribution f (et donc tous ces moments) peut être écrite en ne faisant intervenir que 3 grandeurs macroscopiques définies localement : la densité $\rho(\mathbf{r}, t)$, le champ de vitesse à grande échelle $\bar{\mathbf{v}}(\mathbf{r}, t)$ et la température $T(\mathbf{r}, t)$. Par exemple, on a vu dans le paragraphe 2.1.11 que $f^{(0)}$ prenait la forme d'une maxwellienne locale :

$$f^{(0)}(\mathbf{v}, \mathbf{r}, t) = n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left(-\frac{m(\mathbf{v} - \bar{\mathbf{v}})^2}{2kT} \right). \quad (2.2.44)$$

Une résolution (plus complexe) de l'équation (2.1.61) permet également de exprimer la première correction $f^{(1)}$ comme fonction de $(\rho, \bar{\mathbf{v}}, T)$ (CHAPMAN et COWLING, 1970) :

$$f^{(1)}(\mathbf{v}, \mathbf{r}, t) = -[\mathbf{A} \cdot \nabla \ln T + \mathbf{B} : (\nabla \otimes \bar{\mathbf{v}})] f^{(0)}(\mathbf{v}, \mathbf{r}, t) \quad (2.2.45)$$

avec \mathbf{A} et \mathbf{B} un vecteur et un tenseur respectivement solutions des équations :

$$\frac{1}{n} \hat{\mathbf{I}}[\mathbf{A}] = \left[\frac{m}{2kT} (\mathbf{v} - \bar{\mathbf{v}})^2 - \frac{5}{2} \right] (\mathbf{v} - \bar{\mathbf{v}}) f^{(0)} \quad (2.2.46)$$

$$\frac{1}{n} \hat{\mathbf{I}}[\mathbf{B}] = \left(\frac{m}{kT} \right) [(\mathbf{v} - \bar{\mathbf{v}}) \otimes (\mathbf{v} - \bar{\mathbf{v}}) - (\mathbf{v} - \bar{\mathbf{v}})^2 \mathbb{1}] f^{(0)} \quad (2.2.47)$$

où $\mathbb{1}$ désigne le tenseur unitaire.

On voit donc que les équations (2.2.40)-(2.2.42) forment bien un système d'inconnues $(\rho, \bar{\mathbf{v}}, T)$ et que les expressions des différents moments qui y figure dépendront de l'ordre de troncature choisi dans le développement (2.2.43). Plus précisément, ces moments admettent eux-même un

13. On pourra noter la correspondance avec les équations (17) et (18) de POMRANING, 1982 avec ici la prise en compte du champ de gravité.

développement du type :

$$\bar{\varphi} = \sum_{r=0}^{\infty} \bar{\varphi}^{(r)} \quad \text{avec} \quad \bar{\varphi}^{(r)} \equiv \frac{1}{n} \int \varphi(\mathbf{v}) f^{(r)}(\mathbf{v}, \mathbf{r}, t) d\mathbf{v}. \quad (2.2.48)$$

La troncature du développement de f à l'ordre k impose donc implicitement une relation entre les $f^{(r < k)}$ (et donc entre leurs moments) par la méthode de Chapman-Enskog (cf. paragraphe 2.1.13) :

$$\forall r < k, \quad \mathcal{D}f^{(r)} - \sum_{r'=1}^{r-1} \widehat{\mathcal{C}}[f^{(r')}, f^{(r-r')}] = \widehat{\mathcal{C}}[f^{(r)}, f^{(0)}] + \widehat{\mathcal{C}}[f^{(0)}, f^{(r)}] \quad (2.2.49)$$

$$\mathcal{D}f^{(k)} - \sum_{r=1}^{k-1} \widehat{\mathcal{C}}[f^{(r)}, f^{(k-r)}] = 0 \quad (2.2.50)$$

avec

$$\mathcal{D}f^{(r)} \equiv \frac{\partial f^{(r)}}{\partial t} + \mathbf{v} \cdot \frac{\partial f^{(r)}}{\partial \mathbf{r}} + \frac{\mathbf{F}}{m} \cdot \frac{\partial f^{(r)}}{\partial \mathbf{v}} \quad (2.2.51)$$

puisque $\widehat{\mathcal{C}}[f^{(k)}, f^{(0)}] = \widehat{\mathcal{C}}[f^{(0)}, f^{(k)}] = 0$ dans ce cas. On appelle cette relation équation de fermeture.

2.2.9 Un exemple de fermeture : l'équilibre thermodynamique local

Regardons d'abord ce que l'on obtient dans le cas le plus simple, la troncature à l'ordre 0. Le cadre correspondant a été défini comme celui de l'équilibre thermodynamique local dans la section 2.1. Les différents moment de la distribution sont alors ceux de la maxwellienne locale $f^{(0)}$:

$$\varepsilon^{(0)} = \frac{3}{2} \mathfrak{R}T \quad (2.2.52)$$

$$\mathbf{P}^{(0)} = \rho \mathfrak{R}T \mathbb{1} \quad (2.2.53)$$

$$\mathbf{Q}^{(0)} = \mathbf{0} \quad (2.2.54)$$

avec $\mathfrak{R} \equiv k/m$. Lorsque la radiation peut être négligée, les équations de l'hydrodynamique se réduisent à :

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \bar{\mathbf{v}} \quad (2.2.55)$$

$$\rho \frac{D\bar{\mathbf{v}}}{Dt} = -\nabla(\rho \mathfrak{R}T) - \rho \nabla \Phi \quad (2.2.56)$$

$$\frac{DT}{Dt} = -\frac{2}{3} T \nabla \cdot \bar{\mathbf{v}}. \quad (2.2.57)$$

Notons que la comparaison des équations (2.2.55) et (2.2.57) donnent simplement $d\rho/\rho = (3/2)dT/T$ le long du mouvement. Dans ce cas particulier, la densité peut être éliminée des équations et l'on écrira simplement :

$$\frac{D\bar{\mathbf{v}}}{Dt} = -\nabla \left(\frac{5}{2} \mathfrak{R}T - \Phi \right) \quad (2.2.58)$$

$$\frac{DT}{Dt} = -\frac{2}{3} T \nabla \cdot \bar{\mathbf{v}}. \quad (2.2.59)$$

On reconnaît les équations de l'hydrodynamique idéale¹⁴, accompagnée d'une équation d'état (2.2.53) correspondant au gaz parfait (il s'agit de la fermeture à l'ordre 0). On remarquera la

14. Aussi appelé système d'Euler.

conservation d'une grandeur le long du mouvement pour ce cas spécifique :

$$0 = \mathfrak{R} \left(\frac{3}{2} d \ln T - d \ln \rho \right) = \frac{3\mathfrak{R}}{2\varepsilon^{(0)}} d\varepsilon^{(0)} + \rho \mathfrak{R} d \left(\frac{1}{\rho} \right) = \frac{d\varepsilon^{(0)}}{T} + \frac{p^{(0)}}{T} d \left(\frac{1}{\rho} \right) = ds^{(0)}. \quad (2.2.60)$$

avec $p^{(0)} \equiv \text{Tr}(\mathbf{P}^{(0)})/3 = \rho \mathfrak{R} T$, la composante isotrope du tenseur des pressions. Cette quantité, qui vérifie $d\varepsilon^{(0)} = T ds^{(0)} - p^{(0)} d(1/\rho)$, n'est autre que l'entropie (par unité de masse) à l'équilibre thermodynamique local. On pourra donc qualifier le mouvement du fluide d'isentropique, les seules variations de température sont dues à un flux convectif advecté par le champ de vitesse (cf. équation 2.2.59), lequel vérifie lui-même un équilibre avec les disparités d'énergie du système (cf. équation 2.2.58).

Pour inclure l'impact de la radiation sur le fluide, il nous faudrait trouver pour f^{R} un développement analogue à f et en déduire l'expression de $f^{\text{R},(0)}$. L'équation de Boltzmann vérifiée par f^{R} (2.2.29) est cependant bien plus complexe et il n'est généralement pas possible, sans simplification drastique du problème, de donner une forme explicite de $f^{\text{R},(0)}$. Un tel cas se présente lorsque, en plus de supposer l'équilibre thermodynamique global entre le gaz de photon et le fluide, on impose la stationnarité et l'isotropie de la solution. La solution obtenue peut alors être reliée à la fameuse fonction de Planck $B(\nu)$ (PLANCK, 1967)¹⁵ :

$$h\nu c f^{\text{R},(0)}(\nu, \mathbf{\Omega}, \mathbf{r}, t) = h\nu c f^{\text{R},(0)}(\nu) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/kT} - 1} \equiv B(\nu). \quad (2.2.61)$$

Dans ce cas simple, les différents moments sont obtenus par intégration des relations (2.2.32)-(2.2.34) :

$$E^{\text{R},(0)} = \frac{4\sigma}{c} T^4 \quad (2.2.62)$$

$$\mathbf{Q}^{\text{R},(0)} = 0 \quad (2.2.63)$$

$$\mathbf{P}^{\text{R},(0)} = \frac{4\sigma}{3c} T^4 \mathbb{1} = \frac{E^{\text{R},(0)}}{3} \mathbb{1} \quad (2.2.64)$$

en introduisant la constante de Stefan-Boltzmann

$$\sigma \equiv \frac{2\pi^5 k^4}{15c^2 h^3} \simeq 5,67 \times 10^{-8} \text{ J s}^{-1} \text{ m}^{-2} \text{ K}^{-4}. \quad (2.2.65)$$

De manière générale, l'équilibre thermodynamique global reste un peu trop simpliste pour décrire correctement les interactions entre matière et rayonnement. Pour une application en physique stellaire, il semble en effet peu raisonnable de supposer la distribution comme uniforme et isotrope, ou même stationnaire si l'on veut décrire l'évolution sur les temps longs. Des simplifications plus pertinentes peuvent être déduites d'hypothèses moins fortes, typiquement en supposant un équilibre thermodynamique local en le fluide et le gaz de photons. En effet, une grande partie de la complexité des équations (2.2.35) et (2.2.36) proviennent du taux d'émission et opacités, lesquelles n'admettent généralement pas d'expressions simples. À l'équilibre thermodynamique local, on a vu cependant que toutes les grandeurs physiques s'expriment en fonction de 3 quantités : ρ , $\bar{\mathbf{v}}$ et T . Il est alors possible de chercher des relations simples entre ces grandeurs. Supposons pour commencer que $q(\nu, \mathbf{r}, t)$ et $\kappa_a(\nu, \mathbf{r}, t)$ sont proportionnels, on écrira :

$$q = \alpha \kappa_a \quad (2.2.66)$$

15. Date de la traduction anglaise, la note originale est de 1900.

Prenons pour hypothèse simplificatrice $\kappa_s = 0$ (l'argument tient même sans cette hypothèse, cf. POMRANING, 1982). À l'équilibre thermodynamique *global*, l'équation (2.2.29) s'écrit alors :

$$0 = \rho\alpha\kappa_a - \frac{\rho\kappa_a}{h\nu}B \quad \Rightarrow \quad \alpha = \frac{B}{h\nu} \quad (2.2.67)$$

puisque l'on a vu plus haut que $h\nu c f^{R,(0)}(\nu) = B(\nu)$ avec B la fonction de Planck. Ainsi, pour maintenir la cohérence du formalisme dans cette limite, on pourra éliminer q des équations au profit de $\kappa_a B/h\nu$. Une seconde simplification est obtenue en supposant une diffusion isotrope et cohérente des photons par la matière :

$$\mathcal{K}(\nu \rightarrow \nu_*, \mathbf{\Omega} \cdot \mathbf{\Omega}_*) = \frac{1}{4\pi} \delta(\nu - \nu_*) \quad (2.2.68)$$

avec δ la fonction de Dirac. Le terme de diffusion se simplifie donc en

$$\widehat{C}^R [f^R, f^R] = \frac{\rho c \kappa_s}{4\pi} \int f^R(\nu, \mathbf{\Omega}_*, \mathbf{r}, t) d\mathbf{\Omega}_* - \rho c \kappa_s f^R. \quad (2.2.69)$$

Une propriété très intéressante en résulte : $\int \widehat{C}^R [f^R, f^R] d\mathbf{\Omega} = 0$. Cette dernière peut être exploitée en intégrant l'équation de Boltzmann sur les angles solides (et la multipliant par $h\nu$) :

$$\frac{\partial E_v^R}{\partial t} + \nabla \cdot \mathbf{Q}_v^R = \rho c \kappa_a \left(\frac{4\pi B}{c} - E_v^R \right). \quad (2.2.70)$$

en introduisant X_v , la densité spectrale du moment X . Pour exploiter cette équation, on peut chercher à dériver une seconde relation liant \mathbf{Q}_v^R à E_v^R . On peut commencer par réécrire l'équation (2.2.36) en termes de densité spectrale. Il suffit pour cela de considérer $\int h\nu \mathbf{\Omega} (2.2.29) d\mathbf{\Omega}$, ce qui donne avec les simplifications évoquées :

$$\frac{1}{c} \frac{\partial \mathbf{Q}_v^R}{\partial t} + c \nabla \cdot \mathbf{P}_v^R = -\rho(\kappa_a + \kappa_s) \mathbf{Q}_v^R. \quad (2.2.71)$$

Il reste alors à exprimer \mathbf{P}_v^R en fonction de E_v^R pour clore le système. On peut pour cela supposer que f^R est à symétrie azimutale et admet donc le développement suivant sur les harmoniques sphériques Y_ℓ^m :

$$f^R(\nu, \mathbf{\Omega}, \mathbf{r}, t) = \sum_{\ell=0}^{\infty} a_{\nu,\ell}(\mathbf{r}, t) Y_\ell^0(\mathbf{\Omega}) \quad (2.2.72)$$

avec

$$a_{\nu,\ell}(\mathbf{r}, t) \equiv \int Y_\ell^{0*}(\mathbf{\Omega}') f^R(\nu, \mathbf{\Omega}', \mathbf{r}, t) d\mathbf{\Omega}'. \quad (2.2.73)$$

Si l'on a vu que l'hypothèse d'une distribution isotrope était un peu trop simpliste, on peut cependant supposer que f^R est "quasi-isotrope". Cette hypothèse, appelée approximation de Eddington (EDDINGTON, 1926), peut s'écrire plus formellement comme $a_{\nu,\ell} \simeq 0$ si $\ell \geq 2$. On pourra donc tronquer le développement (2.2.72) et en remarquant que

$$a_{\nu,0} \equiv \int \frac{1}{\sqrt{4\pi}} f^R(\nu, \mathbf{\Omega}', \mathbf{r}, t) d\mathbf{\Omega}' = \frac{E_v^R}{h\nu\sqrt{4\pi}} \quad (2.2.74)$$

$$a_{\nu,1} \equiv \int \frac{3}{\sqrt{4\pi}} (\mathbf{\Omega} \cdot \mathbf{\Omega}') f^R(\nu, \mathbf{\Omega}', \mathbf{r}, t) d\mathbf{\Omega}' = \frac{3(\mathbf{\Omega} \cdot \mathbf{Q}_v^R)}{h\nu c \sqrt{4\pi}}, \quad (2.2.75)$$

cette décomposition se réécrit :

$$f^R \simeq \frac{1}{4\pi h\nu c} \left(c E_v^R + 3 \mathbf{\Omega} \cdot \mathbf{Q}_v^R \right). \quad (2.2.76)$$

La définition (2.2.34) nous permet d'exprimer la densité spectrale du tenseur des pressions radiatives :

$$\mathbf{P}_\nu^R = \frac{E_\nu^R}{4\pi} \int \boldsymbol{\Omega} \otimes \boldsymbol{\Omega} d\boldsymbol{\Omega} + \frac{3\mathbf{Q}_\nu^R}{4\pi c} \cdot \int \boldsymbol{\Omega} \otimes \boldsymbol{\Omega} \otimes \boldsymbol{\Omega} d\boldsymbol{\Omega} = \frac{E_\nu^R}{3} \mathbb{1}. \quad (2.2.77)$$

Ce tenseur reste donc isotrope quand bien même f^R ne possède pas cette symétrie. L'équation (2.2.71) se réécrit donc

$$\frac{1}{c} \frac{\partial \mathbf{Q}_\nu^R}{\partial t} + \frac{c}{3} \boldsymbol{\nabla} E_\nu^R = -\rho(\kappa_a + \kappa_s) \mathbf{Q}_\nu^R. \quad (2.2.78)$$

et clôt bien la relation (2.2.70) entre \mathbf{Q}_ν^R et E_ν^R . Puisque par définition \mathbf{Q}_ν^R est de l'ordre de cE_ν^R , il est généralement possible de négliger le premier terme de cette équation dans la limite non-relativiste en remarquant que $\partial/\partial t \ll c\boldsymbol{\nabla}$. Dans ce cas, le flux d'énergie radiative dérive directement de la densité d'énergie :

$$\mathbf{Q}_\nu^R = -\frac{c}{3\rho\kappa} \boldsymbol{\nabla} E_\nu^R \quad (2.2.79)$$

avec $\kappa = \kappa_a + \kappa_s$ l'opacité totale. Une écriture du type $\mathbf{Q}_\nu^R \propto \boldsymbol{\nabla} E_\nu^R$ rappelle une équation de diffusion, d'où le nom donné à la relation très largement utilisée en physique (2.2.79) de limite diffusive. En pratique, on utilise généralement en physique stellaire une hypothèse encore plus forte qui dérive de la théorie de l'équilibre diffusif (POMRANING, 1973). Cette approximation revient à considérer, en plus de l'équation (2.2.79), la relation (2.2.70) dans sa limite à l'équilibre thermodynamique :

$$E_\nu^R = \frac{4\pi B}{c}. \quad (2.2.80)$$

Ce cadre peut être senti comme pertinent dans un régime dominé par les interactions matière-rayonnement et en ne considérant donc que les termes de droite¹⁶. En intégrant sur le spectre complet, on peut retrouver l'expression des moments intervenant dans les équation de l'hydrodynamique radiative :

$$E^{R,(0)}(\mathbf{r}, t) = \frac{4\sigma}{c} T^4(\mathbf{r}, t) \quad (2.2.81)$$

$$\mathbf{Q}^{R,(0)}(\mathbf{r}, t) = -\int \frac{4\pi}{3\rho\kappa} \boldsymbol{\nabla} B(\nu, \mathbf{r}, t) d\nu \quad (2.2.82)$$

$$\mathbf{P}^{R,(0)}(\mathbf{r}, t) = \frac{4\sigma}{3c} T^4(\mathbf{r}, t) \mathbb{1}. \quad (2.2.83)$$

Ces derniers sont identiques à ceux obtenus à l'équilibre thermodynamique global, à l'exception du flux radiatif¹⁷. Puisque la dépendance spatiale de $B(\nu, \mathbf{r}, t)$ provient seulement de la température (cf. définition (2.2.61)), il est possible de réexprimer ce flux de la façon suivante :

$$\mathbf{Q}^{R,(0)} = -\frac{4\pi}{3\rho} \int \frac{1}{\kappa} \frac{\partial B}{\partial T} d\nu \boldsymbol{\nabla} T = -\frac{4\pi}{3\rho} \frac{\int \frac{1}{\kappa} \frac{\partial B}{\partial T} d\nu}{\int \frac{\partial B}{\partial T} d\nu} \int \frac{\partial B}{\partial T} d\nu \boldsymbol{\nabla} T. \quad (2.2.84)$$

Enfin, puisque $\int \partial B/\partial T d\nu = 4\sigma T^3/\pi$, on trouve généralement le flux radiatif sous la forme (cf. par exemple KIPPENHAHN, WEIGERT et WEISS, 2012) :

$$\mathbf{Q}^{R,(0)} = -\frac{16\sigma}{3\rho\kappa_R} T^3 \boldsymbol{\nabla} T. \quad (2.2.85)$$

16. Notons qu'il s'agit ici d'une approximation utile qui donne de bons résultats en pratique mais difficile à justifier. En effet, si l'on ne conservait que les termes de droite, y compris dans l'équation (2.2.78), on retrouverait $\mathbf{Q}_\nu^R = \mathbf{0}$ et donc l'équilibre thermodynamique global au sens strict.

17. Ils diffèrent aussi dans leur dépendance à la position et au temps puisque la température $T(\mathbf{r}, t)$ est définie localement.

où l'on a introduit au passage la moyenne de Rosseland (ROSSELAND, 1924 ; EDDINGTON, 1926) :

$$\frac{1}{\kappa_R(\mathbf{r}, t)} \equiv \frac{\int \frac{1}{\kappa(\nu, \mathbf{r}, t)} \frac{\partial B}{\partial T} d\nu}{\int \frac{\partial B}{\partial T} d\nu}. \quad (2.2.86)$$

Finalement, en prenant en compte les équations (2.2.52)-(2.2.54) et (2.2.81)-(2.2.83), on voit que l'on a réussi à exprimer toutes les grandeurs intervenant dans les équations dynamiques à l'équilibre thermodynamique en fonction de ρ , $\bar{\mathbf{v}}$ et T ¹⁸. Cette manière de fermer les équations de l'hydrodynamique ne constitue évidemment pas la seule façon de procéder. On pourra notamment mentionner les fermetures entropique ou gaussienne, respectivement proposées par MINERBO, 1978 et LEVERMORE, 1996.

À ce stade, il est tentant de remplacer les expressions des moments obtenus dans les relations (2.2.40)-(2.2.42) pour en dériver les équations de l'hydrodynamique radiative à l'équilibre thermodynamique. Ce serait cependant une erreur. Si les moments qui apparaissent dans ces équations ont été définis dans un référentiel inertiel, l'expression obtenue pour le flux (2.2.85) repose implicitement sur une dérivation où le fluide est immobile. En effet, les hypothèses de diffusion cohérente et de quasi-isotropie de la fonction de distribution n'ont en réalité que peu de sens dans un référentiel où le milieu lui-même est en mouvement. Le flux doit alors être corrigé, et cette correction correspond au passage du cadre eulérien (\mathcal{E}) au cadre lagrangien (\mathcal{L}). Il est possible de montrer que (POMRANING, 1982) :

$$\mathbf{Q}_{\mathcal{E}}^{\mathbf{R},(0)} = \mathbf{Q}_{\mathcal{L}}^{\mathbf{R},(0)} + \frac{4E_{\mathcal{L}}^{\mathbf{R},(0)}}{3} \bar{\mathbf{v}} = \mathbf{Q}_{\mathcal{L}}^{\mathbf{R},(0)} + E_{\mathcal{L}}^{\mathbf{R},(0)} \bar{\mathbf{v}} + \mathbf{P}_{\mathcal{L}}^{\mathbf{R},(0)} \cdot \bar{\mathbf{v}} \quad (2.2.87)$$

avec $X_{\mathcal{E}}$ la quantité exprimée dans le cadre eulérien apparaissant dans (2.2.41)-(2.2.42) et les $X_{\mathcal{L}}$ dans le cadre lagrangien données par les équations (2.2.81)-(2.2.83). Avec le flux radiatif exprimé dans le référentiel suivant le fluide, les équations

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) = 0 \quad (2.2.88)$$

$$\frac{\partial}{\partial t} (\rho \bar{\mathbf{v}}) + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} + \mathbf{P} + \mathbf{P}^{\mathbf{R}}) = -\rho \nabla \Phi \quad (2.2.89)$$

$$\begin{aligned} & \frac{\partial}{\partial t} \left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + E^{\mathbf{R}} + \rho \Phi \right) \\ & + \nabla \cdot \left[\left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + E^{\mathbf{R}} + \rho \Phi \right) \bar{\mathbf{v}} + (\mathbf{P} + \mathbf{P}^{\mathbf{R}}) \cdot \bar{\mathbf{v}} + \mathbf{Q} + \mathbf{Q}^{\mathbf{R}} \right] = 0 \end{aligned} \quad (2.2.90)$$

remplacent les relations (2.2.40)-(2.2.42) (on a retiré le terme relativiste en $1/c^2$ dans l'équation sur la quantité de mouvement). De façon intéressante, on reconnaît alors les équations de l'hydrodynamique auxquelles on a appliqué la transformation $\mathbf{P} \rightarrow \mathbf{P} + \mathbf{P}^{\mathbf{R}}$, $\mathbf{Q} \rightarrow \mathbf{Q} + \mathbf{Q}^{\mathbf{R}}$ et $\varepsilon \rightarrow \varepsilon + \varepsilon^{\mathbf{R}}$ en définissant $\rho \varepsilon^{\mathbf{R}} \equiv E^{\mathbf{R}}$. En particulier, on voit qu'il est possible de donner un sens aux notions de pression, de flux et d'énergie totales en sommant les grandeurs du fluide et du gaz de photons. Cette propriété, parfois invoquée pour dériver les équations (2.2.88)-(2.2.90) comme s'il s'agissait d'une trivialité, n'a en réalité rien d'évident. En effet, les grandeurs associées à la matière et au rayonnement ne sont même pas calculées dans le même référentiel et ces relations ne sont plus vérifiées si la vitesse entre ces deux référentiels devient trop élevée (cadre relativiste).

18. Il nous manque en réalité une écriture explicite pour κ et qui permettrait de calculer κ_R en fonction de ces variables. Il est de coutume d'introduire une forme fonctionnelle simple pour la moyenne de Rosseland $\kappa_R = \kappa_0 \rho T^{-7/2}$ appelée opacité de Kramers (SCHWARZSCHILD, 1958).

À l'équilibre thermodynamique local on pourra écrire :

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \bar{\mathbf{v}} \quad (2.2.91)$$

$$\rho \frac{D\bar{\mathbf{v}}}{Dt} = -\nabla p_{\text{tot}} - \rho \nabla \Phi \quad (2.2.92)$$

$$\rho \frac{D\varepsilon_{\text{tot}}}{Dt} = -p_{\text{tot}} (\nabla \cdot \bar{\mathbf{v}}) + \nabla \cdot \left(\frac{16\sigma}{3\rho\kappa_R} T^3 \nabla T \right) \quad (2.2.93)$$

avec

$$\varepsilon_{\text{tot}}(\rho, T) = \frac{3}{2} \mathfrak{R}T + \frac{4\sigma T^4}{\rho c} \quad (2.2.94)$$

$$p_{\text{tot}}(\rho, T) = \rho \mathfrak{R}T + \frac{4\sigma T^4}{3c} \quad (2.2.95)$$

les énergie massique et pression totales. Si l'on cherche à écrire cette dernière équation sur la température T , on peut voir apparaître la forme suivante :

$$\frac{DT}{Dt} = -\frac{2+2\beta}{3+9\beta} T \nabla \cdot \bar{\mathbf{v}} + \frac{4\beta}{1+3\beta} \frac{c}{\rho\kappa_R} \Delta T + \frac{16\beta}{3+9\beta} \dots \quad (2.2.96)$$

avec $\beta \equiv \varepsilon^R/\varepsilon_{\text{tot}}$ et les pointillés indiquant la présence de termes plus complexes faisant notamment intervenir les variations d'opacité. On constate ainsi la présence d'un terme de diffusion, d'autant plus fort que le rayonnement domine ($\beta \rightarrow 1$) et d'autant plus faible que la moyenne de Rosseland est élevée (l'opacité agit contre la diffusion de la température). On remarque que l'on retrouve bien l'équation (2.2.59) dans la limite où le rayonnement devient négligeable ($\beta \rightarrow 0$).

2.2.10 La fermeture de Navier-Stokes

On finira par mentionner succinctement l'impact que peut avoir une fermeture des équations dynamiques à l'ordre supérieur, cas qui correspond à un léger écart à l'équilibre thermodynamique local. En l'absence de radiation, on a vu qu'une troncature à l'ordre 1 (plutôt qu'à l'ordre 0, laquelle à l'hydrodynamique idéale) du développement (2.2.43) donne pour la distribution :

$$f(\mathbf{v}, \mathbf{r}, t) \simeq [1 - \mathbf{A} \cdot \nabla \ln T - \mathbf{B} : (\nabla \otimes \bar{\mathbf{v}})] n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left(-\frac{m(\mathbf{v} - \bar{\mathbf{v}})^2}{2kT} \right) \quad (2.2.97)$$

avec \mathbf{A} et \mathbf{B} solutions des équations (2.2.46) et (2.2.47). Après de longs développements, ce choix de f conduit aux expressions des moments suivantes :

$$\varepsilon = \frac{3}{2} \mathfrak{R}T \quad (2.2.98)$$

$$\mathbf{P} = \rho \mathfrak{R}T \mathbf{1} - 2\mu \left[\frac{1}{2} \left((\nabla \otimes \bar{\mathbf{v}}) + (\nabla \otimes \bar{\mathbf{v}})^T \right) - \frac{1}{3} (\nabla \cdot \bar{\mathbf{v}}) \mathbf{1} \right] = p \mathbf{1} - \Sigma \quad (2.2.99)$$

$$\mathbf{Q} = -\lambda \nabla T \quad (2.2.100)$$

avec

$$\mu(\mathbf{r}, t) = \frac{m^2 kT}{10\rho^2} \int \mathbf{B} : \hat{\mathbf{I}}[\mathbf{B}] d\mathbf{v} \quad (2.2.101)$$

$$\lambda(\mathbf{r}, t) = \frac{2mk^2 T}{3\rho^2} \int \mathbf{A} \cdot \hat{\mathbf{I}}[\mathbf{A}] d\mathbf{v} \quad (2.2.102)$$

les coefficients de viscosité et de conductivité thermique qui ne font intervenir que la densité ρ et la température T . Ici Σ désigne un tenseur anisotrope qui représente les contraintes visqueuses que subit le fluide. On constate également l'apparition d'un flux de chaleur non nul par conduction. L'écriture des équations de l'hydrodynamique dans ce cadre aboutit au système de Navier-Stokes :

$$\frac{D\rho}{Dt} = -\rho \nabla \cdot \bar{\mathbf{v}} \quad (2.2.103)$$

$$\rho \frac{D\bar{\mathbf{v}}}{Dt} = -\nabla p + \nabla \cdot \Sigma - \rho \nabla \Phi \quad (2.2.104)$$

$$\rho \frac{D\varepsilon}{Dt} = -p (\nabla \cdot \bar{\mathbf{v}}) + \Sigma : (\nabla \otimes \bar{\mathbf{v}}) + \nabla \cdot (\lambda \nabla T). \quad (2.2.105)$$

On voit que ces équations permettent de prendre très facilement en compte la radiation (dans le cadre de l'équilibre diffusif) en procédant aux substitutions $p \rightarrow p_{\text{tot}}$, $\varepsilon \rightarrow \varepsilon_{\text{tot}}$ (cf. Eqs. (2.2.94) et (2.2.95)) et $\lambda \rightarrow \lambda_{\text{tot}}$, en introduisant $\lambda_{\text{tot}} = \lambda + \lambda^{\text{R}}$ avec

$$\lambda^{\text{R}}(\mathbf{r}, t) \equiv \frac{16\sigma}{3\rho\kappa_{\text{R}}} T^3 \quad (2.2.106)$$

conformément à l'expression (2.2.85). En définissant l'entropie hydrodynamique par $Tds \equiv d\varepsilon + pd(1/\rho)$, on peut aussi écrire l'équation (2.2.105) sous la forme :

$$\rho T \frac{Ds}{Dt} = \Sigma : (\nabla \otimes \bar{\mathbf{v}}) + \nabla \cdot (\lambda \nabla T) \quad (2.2.107)$$

en utilisant la relation (2.2.18). Qualitativement, on peut voir que l'augmentation de l'entropie dans un fluide hors équilibre thermodynamique trouve principalement deux origines : la première est l'action des contraintes visqueuses qui tend à amoindrir les fluctuations du champ de vitesse et la seconde est le flux de chaleur qui tend à uniformiser la température du fluide par dissipation. À l'équilibre, ces deux termes disparaissent et $Ds/Dt = 0$, l'entropie a atteint son maximum.

2.3 Des équations dynamiques à la structure stellaire

Contenu de la section

2.3.1	Formes stationnaire et statique des équations fluides	43
2.3.2	Un modèle simple de structure à l'équilibre statique	44
2.3.3	Les étoiles	45
2.3.4	Composition et diffusion	46
2.3.5	Stabilité du système d'équations	49
2.3.6	Les régions stellaires	53
2.3.7	L'ionisation	56
2.3.8	La modélisation stellaire	59
2.3.9	Conditions aux limites	60

Cette section a pour objet l'étude des structures à l'équilibre vis-à-vis des équations obtenues dans la section 2.2. On commencera par se pencher sur les formes stationnaires et statiques des équations, avant d'aborder les structures qui maintiennent dynamiquement leur équilibre. On s'intéressera ensuite aux phénomènes qui font les spécificités des objets centraux de cette étude : les étoiles.

2.3.1 Formes stationnaire et statique des équations fluides

Dans la section précédente, on a parcouru de nombreux systèmes d'équations permettant la description d'un fluide soumis à la gravitation et au rayonnement thermique en présence d'hypothèses plus ou moins fortes. Lorsque ce fluide est proche de l'équilibre thermodynamique local, on pourra retenir le système d'équation suivant (sous forme conservative) :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) = 0 \quad (2.3.1)$$

$$\frac{\partial \rho \bar{\mathbf{v}}}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} + p \mathbb{1} - \Sigma) = -\rho \nabla \Phi \quad (2.3.2)$$

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi \right) + \nabla \cdot \left[\left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi + p \right) \bar{\mathbf{v}} - \Sigma \cdot \bar{\mathbf{v}} + \mathbf{Q} \right] = 0. \quad (2.3.3)$$

où les termes p, ε et \mathbf{Q} peuvent être aisément modifiés pour tenir compte de la radiation ou non. On va chercher à partir d'ici la structure la plus simple possible satisfaisant à ces équations. Tout d'abord, les structures qui vont nous intéresser à partir d'ici seront des structures fluides auto-gravitantes. Le potentiel de gravitation peut dès lors être déduit de la distribution de masse à l'intérieur du fluide par l'équation de Poisson :

$$\Delta \Phi = 4\pi G \rho \quad (2.3.4)$$

avec $G \simeq 6,674 \times 10^{-11} \text{ m}^3 \text{ kg}^{-1} \text{ s}^{-2}$ la constante universelle de gravitation. Dans l'optique de trouver une solution simple aux équations (2.3.1)-(2.3.3), on peut s'intéresser à un fluide ayant évolué sur un temps long devant un temps caractéristique et donc à l'équilibre thermodynamique. Au regard des constantes introduites jusqu'à présent pour décrire le fluide, il n'existe en réalité qu'un seul moyen de les agencer de façon à former un temps caractéristique :

$$t_{\text{dyn}} \sim \sqrt{\frac{V}{GmN}} \quad (2.3.5)$$

où V désigne le volume total du fluide, m la masse d'une particule, N le nombre de particules (et donc mN la masse totale du fluide) et G la constante de gravitation. Ce temps est un temps dynamique et correspond approximativement au temps de chute libre à l'intérieur de la structure sous l'action du champ de gravité. Pour un temps t grand devant t_{dyn} on pourra donc commencer par chercher la structure d'une solution stationnaire ($\partial/\partial t = 0$) de ces équations. En annulant les premiers termes de chaque équation, on voit notamment que les champs

$$\rho \bar{\mathbf{v}} \quad \text{et} \quad \left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi + p \right) \bar{\mathbf{v}} - \Sigma \cdot \bar{\mathbf{v}} + \mathbf{Q},$$

respectivement associé à la quantité de mouvement et flux d'énergie, sont à flux conservatif et dérivent donc d'un potentiel vecteur. Cette propriété contraint grandement leur géométrie et donc celle d'une solution stationnaire à ces équations.

L'étude d'un tel système reste néanmoins complexe, et l'on pourra d'abord chercher une solution qui soit en plus statique ($\bar{\mathbf{v}} = \mathbf{0}$). Dans ce cas, l'équation de continuité est automatiquement satisfaite et l'on obtient le système¹⁹ :

$$\nabla p = -\rho \nabla \Phi \quad (2.3.6)$$

$$\nabla \cdot \mathbf{Q} = 0. \quad (2.3.7)$$

La première équation n'est autre que l'équilibre hydrostatique entre les forces de pression et de gravitation, tandis que la seconde contraint la géométrie du flux de chaleur. En intégrant cette équation sur l'ensemble du domaine, une application du théorème de Stokes donne :

$$\int_V \nabla \cdot \mathbf{Q} \, d\mathbf{r} = \oint_{\partial V} \mathbf{Q} \cdot d\mathbf{S} = 0. \quad (2.3.8)$$

On voit alors qu'il n'y a aucun échange de chaleur avec l'extérieur du domaine.

2.3.2 Un modèle simple de structure à l'équilibre statique

À l'aide de l'équation de Poisson (2.3.4), on peut alors observer que la première équation (2.3.6) est la définition implicite d'une relation entre pression et densité :

$$\nabla p = \rho \nabla \int_V \frac{G\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}'. \quad (2.3.9)$$

Cette relation permet de déterminer le gradient de pression en tout point pour une distribution de masse donnée. En plus de l'équation (2.3.9), on voit qu'il suffit d'une seconde relation entre pression et densité pour imposer complètement la structure de la solution. En s'inspirant de la relation dérivée pour l'hydrodynamique idéale $\rho \propto T^{3/2}$, on peut chercher à imposer une relation du type $p \propto \rho^\gamma$ avec γ une constante. Le cas idéal donnerait alors $\gamma = 5/3$ (puisque $p \propto \rho T$) et une consultation de l'équation (2.2.96) donne $\gamma \simeq 4/3$ dans les milieux dominés par la radiation ($\beta \rightarrow 1$). Une relation de ce type est appelée relation polytropique et la structure qui en résulte polytrope. En substituant ainsi la pression dans l'équation (2.3.9), la distribution de masse doit alors satisfaire à la contrainte :

$$\rho^{\gamma-1} = \int_V \frac{K\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \text{cnst}. \quad (2.3.10)$$

avec $K = G(\gamma - 1)\rho^\gamma / \gamma p = \text{cnst}$. De par sa simplicité, le polytrope a de nombreuses applications en physique et a permis d'effectuer les premières approximations de structures stellaires (EDDINGTON, 1926; CHANDRASEKHAR, 1939).

19. Σ s'annule en vertu de la définition (2.2.99)

2.3.3 Les étoiles

On peut maintenant se demander si une structure régie par l'équation (2.3.10) peut exister de façon réaliste et donc constituer un modèle d'étoile statique. En plus de la stratification polytropique (2.3.10), l'équation (2.3.7) doit être respectée. Si l'on y injecte $\mathbf{Q} = -\lambda \nabla T$, hors coïncidence miraculeuse entre les variations de conductivité et de température, il est généralement nécessaire que $\nabla T = \mathbf{0}$. La température d'une telle structure doit alors être uniforme, ce qui correspond au cas limite $\gamma \rightarrow 1$. Même dans un tel cas, pour que l'équation soit vérifiée sur le bord du domaine, il faut alors que cette température corresponde à celle de l'extérieur T_{ext} . Cette situation correspond à un équilibre thermodynamique global avec l'extérieur et l'on pourra vérifier que la structure dans son ensemble est décrite par un faible nombre de grandeurs comme sa température T_{ext} ou sa densité moyenne mN/V (lesquelles fixent les valeurs des constantes dans l'équation (2.3.10)) qui jouent le rôle de variables d'état. Ce cas de figure semble évidemment irréaliste et en pratique l'équation (2.3.7) n'est jamais respectée, l'étoile n'est pas à l'équilibre thermodynamique global et sa température en surface T_{surf} diffère de $T_{\text{ext}} \simeq 0\text{K}$. Le flux de chaleur \mathbf{Q} n'est donc pas nul et notre structure perd de l'énergie au cours du temps : elle rayonne²⁰. Il semble donc que la structure ne puisse être décrite par des équations stationnaires même après un temps d'évolution grand devant t_{dyn} et qu'il faille à minima considérer les équations (2.3.1)-(2.3.3) en l'absence de champ de vitesse à grande échelle :

$$\frac{\partial \rho}{\partial t} = 0 \quad (2.3.11)$$

$$\nabla p = -\rho \nabla \Phi \quad (2.3.12)$$

$$\frac{\partial}{\partial t} (\rho \varepsilon + \rho \Phi) + \nabla \cdot \mathbf{Q} = 0. \quad (2.3.13)$$

On pourra vérifier, en prenant l'exemple de la structure (2.3.10), que ∇T est orienté vers l'intérieur du domaine pour $\gamma > 1$ ($T \propto \rho^{\gamma-1}$) et donc que $\nabla \cdot \mathbf{Q} > 0$. Ce flux représente donc bien une perte d'énergie locale : $\rho \varepsilon + \rho \Phi$ décroît au fil du temps. Ayant introduit la température en surface de l'étoile, il est désormais possible de trouver un nouveau temps caractéristique pour l'évolution de la structure. Une intégration de la dernière équation (2.3.13) sur l'ensemble du domaine donne :

$$\int_V \left(\frac{\partial}{\partial t} (\rho \varepsilon + \rho \Phi) + \nabla \cdot \mathbf{Q} \right) d\mathbf{r} = \frac{\partial}{\partial t} \int_V \rho \varepsilon + \rho \Phi d\mathbf{r} + \oint_{\partial V} \mathbf{Q} \cdot d\mathbf{S} = 0 \quad (2.3.14)$$

Le premier terme représente la variation temporelle de l'ensemble de l'énergie contenue dans le domaine, laquelle s'équilibre avec l'ensemble du flux rayonné en surface à chaque instant. En supposant l'énergie interne totale du même ordre de grandeur que l'énergie gravitationnelle contenue dans la structure, on pourra écrire en terme d'échelle :

$$\int_V \rho \varepsilon + \rho \Phi d\mathbf{r} \sim \frac{G(mN)^2}{V^{1/3}} \quad \text{et} \quad \oint_{\partial V} \mathbf{Q} \cdot d\mathbf{S} \sim V^{2/3} \sigma T_{\text{surf}}^4 \quad (2.3.15)$$

Au regard de l'équation (2.3.13), le temps caractéristique correspondant est donné par le rapport de ces deux quantités :

$$t_{\text{KH}} \sim \frac{G(mN)^2}{V \sigma T_{\text{surf}}^4}. \quad (2.3.16)$$

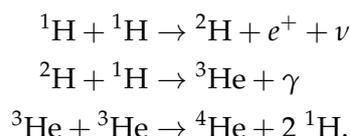
Cette échelle temporelle, appelée temps de Kelvin-Helmholtz (**KELVIN, 1862**), donne une idée de la durée sur laquelle l'étoile peut perdre de l'énergie pour maintenir une température T_{surf} à sa

²⁰. Techniquement, à la vue de ce qui a été introduit jusqu'à présent, la structure perd aussi de l'énergie par conduction puisque $\lambda_{\text{tot}} = \lambda + \lambda^{\text{R}}$. On pourra cependant considérer que la conductivité thermique à l'extérieur du domaine est très faible.

surface (on remarquera que ce temps diminue très vite lorsque T_{surf} augmente). On peut facilement voir que ce temps est généralement très grand devant le temps dynamique t_{dyn} et donc, qu'après un temps d'évolution grand devant t_{KH} , l'étoile devrait se trouver dans un équilibre à la fois dynamique et thermodynamique. Après avoir estimé l'ordre de grandeur de cette échelle pour le soleil, Lord Kelvin écrivait (KELVIN, 1862) :

It seems, therefore, on the whole most probable that the sun has not illuminated the earth for 100,000,000 years, and almost certain that he has not done so for 500,000,000 years. As for the future, we may say, with equal certainty, that inhabitants of the earth can not continue to enjoy the light and heat essential to their life for many million years longer unless sources now unknown to us are prepared in the great storehouse of creation.

Au regard de nos connaissances actuelles, il est clair que notre description manque d'un ingrédient essentiel, le soleil étant en réalité bien plus âgé que ça. Cette source additionnelle de chaleur provient de la fusion nucléaire qui a lieu au cœur de l'étoile et qui est complètement omise par la description cinétique du fluide que l'on a menée jusqu'ici. Comme on va le voir par la suite, ce nouveau facteur introduit une grande complexité si l'on souhaite décrire la structure dans son intégralité. De façon concise, on peut voir la production d'énergie par fusion nucléaire comme la réaction d'un ensemble de particules de masse m_i formant des particules de masse m_j , libérant au passage une énergie de masse $E^N = (\sum_i m_i - \sum_j m_j)c^2$. À titre d'exemple, la chaîne de réaction la plus simple (pp1), permet une libération d'énergie au travers d'une conversion d'hydrogène (^1H) en hélium (^4He) :



Grossièrement, on peut voir que le bilan de la réaction est la production d'un noyau d'hélium de masse $m_{\text{He}} = 4.0026 m_u$ à partir de 4 noyaux d'hydrogène de masse $4m_{\text{H}} = 4.0316 m_u$, et représente donc une perte de masse d'environ $0.029 m_u \simeq 0.007 \times 4m_{\text{H}}$ pour le système (KIPPENHAHN, WEIGERT et WEISS, 2012). Pour introduire une nouvelle échelle de temps, on peut chercher la durée sur laquelle une étoile, initialement composée à 100% d'hydrogène, peut maintenir une température T_{surf} en produisant de l'hélium par réaction nucléaire pp1. L'énergie potentielle dont elle dispose pour ce processus est de $0.007 \times m_{\text{H}} N c^2$ et le flux rayonné en surface est de $V^{2/3} \sigma T_{\text{surf}}^4$ (cf. 2.3.15). Cette échelle est donc donnée par

$$t_{\text{nuc}} \sim 0.007 \times \frac{m_{\text{H}} N c^2}{V^{2/3} \sigma T_{\text{surf}}^4} \quad (2.3.17)$$

qui est de l'ordre de $10^3 t_{\text{KH}}$ et de $10^{15} t_{\text{dyn}}$ pour le soleil. Une telle source d'énergie permet donc de concevoir l'existence de structures n'ayant toujours pas atteint l'équilibre thermodynamique après plusieurs milliards d'années.

2.3.4 Composition et diffusion

On voit immédiatement que la fusion nucléaire induit une représentation qui tranche largement avec notre description du fluide jusqu'ici, à savoir un ensemble de N particules identiques influencées par le rayonnement. En effet, pour décrire la structure convenablement (même après un temps très long), il semble indispensable d'inclure différents types de particules dans la description du fluide et de connaître l'abondance de chacun de ces éléments dans l'étoile. Plus explicitement, on pourra supposer une mixture de N_{sp} espèces différentes, indexées par i , de masse m_i et chacune comptabilisant N_i particules (on a ainsi $N = \sum_i N_i$). Ce simple changement induit de nombreux effets sur les équations étudiées jusqu'ici, le premier étant la modification de l'équation de Boltzmann. Il est désormais nécessaire d'introduire N_{sp} équations de Boltzmann

différentes, chacune s'appliquant à $f_i \equiv f(\mathbf{v}_i, \mathbf{r}, t)$:

$$\frac{\partial f_i}{\partial t} + \mathbf{v}_i \cdot \frac{\partial f_i}{\partial \mathbf{r}} + \frac{\mathbf{F}_i}{m_i} \cdot \frac{\partial f}{\partial \mathbf{v}_i} = \sum_{j=1}^{N_{\text{sp}}} \widehat{\mathcal{C}} [f_i, f_j] \quad (2.3.18)$$

et qui rendent compte à la fois des collisions homogènes ($j = i$) et hétérogènes ($j \neq i$) entre les particules. Il est alors possible de définir des moyennes par espèce à partir des distributions f_i :

$$\bar{\varphi}_i \equiv \langle \varphi_i \rangle \equiv \frac{\int \varphi(\mathbf{v}_i) f(\mathbf{v}_i, \mathbf{r}, t) d\mathbf{v}_i}{\int f(\mathbf{v}_i, \mathbf{r}, t) d\mathbf{v}_i} = \frac{1}{n_i} \int \varphi(\mathbf{v}_i) f_i d\mathbf{v}_i. \quad (2.3.19)$$

Les quantités moyennes pour l'ensemble du système peuvent être obtenues par des moyennes de masses pour les variables intensives (comme la vitesse) :

$$\bar{\varphi} \equiv \sum_i X_i \bar{\varphi}_i \quad (2.3.20)$$

avec $X_i \equiv \rho_i / \rho = m_i n_i / \sum_j m_j n_j$ l'abondance de l'espèce i , ou en sommant directement pour une variable extensive $\bar{\psi}_i = \rho_i \bar{\varphi}_i$ puisque :

$$\sum_i \bar{\psi}_i = \sum_i \rho_i \bar{\varphi}_i = \rho \sum_i X_i \bar{\varphi}_i = \rho \bar{\varphi} = \bar{\psi}. \quad (2.3.21)$$

En définissant la vitesse d'ensemble du système $\bar{\mathbf{v}} \equiv \sum_i X_i \bar{\mathbf{v}}_i$, une des premières conséquences de l'introduction de plusieurs espèces sur les équations est que la vitesse de diffusion $\mathbf{v}'_i \equiv \mathbf{v}_i - \bar{\mathbf{v}}$ n'est pas nulle en moyenne. En effet :

$$\langle \mathbf{v}'_i \rangle \equiv \frac{1}{n_i} \int (\mathbf{v}_i - \bar{\mathbf{v}}) f_i d\mathbf{v}_i = \bar{\mathbf{v}}_i - \bar{\mathbf{v}} \neq \mathbf{0}. \quad (2.3.22)$$

Cette propriété introduit un nouveau phénomène physique dans les équations : la diffusion atomique. Pour illustrer ce point, on peut reprendre l'équation (2.2.9) pour chacune des fonctions $\varphi(\mathbf{v}_i) = X_i$. On obtient l'équivalent de N_{sp} équations de continuité pour chaque espèce

$$\frac{\partial \rho_i}{\partial t} + \nabla \cdot (\rho_i \bar{\mathbf{v}} + \mathbf{J}_i) = 0 \quad (2.3.23)$$

où l'on a défini pour la première fois un moment d'ordre 1 :

$$\mathbf{J}_i \equiv \rho_i \langle \mathbf{v}'_i \rangle = \rho_i (\bar{\mathbf{v}}_i - \bar{\mathbf{v}}) \quad (2.3.24)$$

le flux de masse de l'élément i . Puisque $\rho_i = \rho X_i$, ces équations peuvent être mises sous forme convective et ainsi décrire l'évolution en abondance de chaque élément

$$\rho \frac{DX_i}{Dt} = -\nabla \cdot \mathbf{J}_i. \quad (2.3.25)$$

Ainsi, de façon naturelle et pour des raisons cinétiques, les abondances des divers éléments composant le fluide varient au cours du temps de par leurs disparités de masse. Notons que ce phénomène n'intervient pas dans les équations idéales, même en présence de plusieurs espèces. Dans le cadre d'une fermeture à l'ordre 1 cependant (cf. paragraphe 2.2.10), à l'instar de la pression ou du flux de chaleur, la théorie cinétique nous offre la possibilité d'explicitier les flux de masse \mathbf{J}_i (par généralisation des équations de CHAPMAN et COWLING, 1970, à plus de deux espèces) :

$$\mathbf{J}_i = -\frac{\rho}{m^2} \sum_{j \neq i} m_i m_j \mathcal{D}_{ij} [\nabla x_i + (x_i - X_i) \nabla \ln p] - \rho \mathcal{D}_i^T \nabla \ln T \quad (2.3.26)$$

avec $x_i \equiv n_i/n$ l'abondance volumique de l'élément i et $m \equiv \rho/n = \sum_i m_i x_i$ la masse moyenne d'une particule du fluide. Les \mathcal{D}_{ij} ($= \mathcal{D}_{ji}$) sont au nombre de $N_{\text{sp}}(N_{\text{sp}} - 1)/2$ et représentent les coefficients de diffusion binaires entre les espèces i et j . Les \mathcal{D}_i^T sont au nombre de N_{sp} et constituent quant à eux la contribution thermique de la diffusion. On distingue 3 origines au processus de diffusion. La première résulte de la surabondance (ou sous-abondance) d'une espèce donnée qui tend à s'homogénéiser en créant au passage un flux de matière : on parle d'osmose (terme en ∇x_i). La seconde se situe au niveau du gradient de pression, on parle de "sédimentation gravitationnelle" (*gravitational settling* [en]) en physique stellaire, les variations de pression s'équilibrant avec la gravité. Les éléments les plus lourds sont amenés à suivre le gradient de pression (et donc le champ de gravité) tandis que les plus légers sont ramenés vers la surface. Cela peut être mis en évidence en remarquant que $-m_i(x_i - X_i)\nabla \ln p = (m_i - m)X_i(\rho/p)\mathbf{g}$; le sens du courant diffusion est donc orienté selon le signe de $m_i - m$. À noter que, dans les régions où la pression radiative se fait forte, un terme en ∇p^{R} apparaît. Ce dernier s'oppose à la gravité (puisque $\nabla p = \rho\mathbf{g} - \nabla p^{\text{R}}$), ce qui lui vaut le nom de "lévitation radiative" (*radiative buoyancy* [en]). Le dernier terme de l'expression (2.3.26) est un peu moins intuitif et peut être interprété comme une réponse variable des éléments lorsque soumis à un gradient de température (on parle d'effet Soret ou de thermophorèse). Ce dernier processus tend à s'exercer dans le même sens que la sédimentation gravitationnelle en ramenant les éléments les plus lourds au centre de l'étoile (THOUL, BAHCALL et LOEB, 1994).

Tout comme les conductivité λ , coefficient de viscosité μ ou opacité de Rosseland κ_R , les coefficients de diffusions binaires \mathcal{D}_{ij} et thermiques \mathcal{D}_i^T peuvent être vues comme des propriétés du milieu. À noter que, depuis l'introduction d'un mélange d'élément dans les équations, ces grandeurs ne s'expriment généralement plus comme des fonctions de ρ et T seulement²¹. Formellement, le système thermodynamique peut être intégralement contraint si l'on connaît les triplets densité, vitesse, température de chaque espèce $(\rho_i, \bar{\mathbf{v}}_i, T_i)$, (ou les $(n_i, \bar{\mathbf{v}}_i, T_i)$ puisque $n_i = \rho_i/m_i$) qui fixent les caractéristiques des fonctions de distribution f_i . Hors cas très spécifique, il est rare de considérer que les températures diffèrent d'une espèce à l'autre et l'on suppose $\forall i, T_i = T$. En plus de cela, puisque $\mathbf{J}_i = \rho_i(\bar{\mathbf{v}}_i - \bar{\mathbf{v}})$, l'équation (2.3.26) permet d'exprimer les $\bar{\mathbf{v}}_i$ sur les variables restantes, ce qui laisse à déterminer les (ρ_i, T) en plus de la vitesse $\bar{\mathbf{v}}$. À noter qu'en physique stellaire, on exprime traditionnellement les grandeurs comme des fonctions de (ρ, T, X_i) en utilisant le lien $\rho X_i = \rho_i$ (plus de détails sur les différentes représentations sont à trouver dans l'annexe A.4).

Au regard des équations (2.3.23) et (2.3.26), on voit qu'il est possible de définir une échelle de temps caractéristique pour le phénomène de diffusion :

$$t_{\text{diff}} \sim \frac{V^{2/3}}{\mathcal{D}} \quad (2.3.27)$$

avec \mathcal{D} un ordre de grandeur typique pour les \mathcal{D}_{ij} et \mathcal{D}_i^T . Une estimation de cette valeur pour une étoile donne typiquement $t_{\text{diff}} \sim t_{\text{nuc}}$, faisant de la fusion nucléaire et la diffusion deux phénomènes impactant les mêmes échelles de temps pour ce qui est des variations d'abondance. L'absence de phénomène dominant à plus court terme nous oblige donc à les prendre en compte simultanément dans l'équation (2.3.23).

Si l'on cherche maintenant à inclure les effets de la fusion nucléaire sur les variations d'abondance, il nous faut introduire les taux de conversion (locaux) par fusion r_{ij} de l'espèce i vers j . On pourra alors introduire une source de masse locale $m_i \sum_j r_{ji}$ par fusion nucléaire ainsi qu'un

21. La dépendance en $\bar{\mathbf{v}}$ des moments de la fonction de distribution n'apparaît pas dans les coefficients, que l'on préfère indépendants du champ de vitesse si possible. Un tel cas figure dans l'équation (2.2.99), lors du choix de définition du coefficient de viscosité μ par $\Sigma(\rho, T, \bar{\mathbf{v}}) = 2\mu(\rho, T)F(\bar{\mathbf{v}})$.

puits de masse $-m_i \sum_j r_{ij}$ dans l'équation (2.3.23) :

$$\frac{\partial \rho_i}{\partial t} + \nabla \cdot (\rho_i \bar{\mathbf{v}} + \mathbf{J}_i) = m_i \sum_{j \neq i} (r_{ji} - r_{ij}). \quad (2.3.28)$$

Ces deux termes représentent respectivement la création de masse qui résulte de toute réaction impliquant l'élément i comme produit de fusion et la destruction de masse par toute réaction l'impliquant comme réactif. Cette équation s'écrit sous forme convective sur les abondances de la façon suivante :

$$\rho \frac{DX_i}{Dt} = -\nabla \cdot \mathbf{J}_i + m_i \sum_{j \neq i} (r_{ji} - r_{ij}). \quad (2.3.29)$$

Enfin, en plus de son impact sur la composition, la fusion libère de l'énergie dans le milieu, ce qui constitue une source pour l'équation (2.3.13) que l'on notera $\rho \dot{\epsilon}_{\text{nuc}}$.

2.3.5 Stabilité du système d'équations

En tenant compte de la fusion nucléaire et de la diffusion, il apparaît que les équations qui décrivent l'évolution de notre structure prennent la forme suivante :

$$\frac{\partial}{\partial t} (\rho X_i) + \nabla \cdot \mathbf{J}_i = m_i \sum_{j \neq i} (r_{ji} - r_{ij}) \quad (2.3.30)$$

$$\nabla p = -\rho \nabla \Phi \quad (2.3.31)$$

$$\frac{\partial}{\partial t} (\rho \epsilon + \rho \Phi) + \nabla \cdot \mathbf{Q}^R = \rho \dot{\epsilon}_{\text{nuc}}. \quad (2.3.32)$$

en considérant que le flux d'énergie est dominé par la radiation plutôt que par la conduction.

À la vue des points abordés jusqu'ici, ce système semble a priori pertinent pour décrire la structure d'une étoile de façon réaliste. On remarquera cependant que l'on a fait une hypothèse très forte dès le début de la section en vue de simplifier le problème : supposer que la structure ne possède pas de champ de vitesse à grande échelle ($\bar{\mathbf{v}} = \mathbf{0}$). On va maintenant chercher à trouver sous quelles hypothèses cette simplification se révèle adaptée.

De façon intuitive, on peut voir qu'une telle situation se présente dans le cas où la stratification de l'étoile tend à faire disparaître les déplacements sur des échelles macroscopiques. Considérons par exemple une particule de fluide située en \mathbf{r} que l'on déplacerait légèrement²² jusqu'en $\mathbf{r} + \boldsymbol{\xi}$. Si ce déplacement s'effectue sur une échelle de temps dynamique ($\boldsymbol{\xi}$ peut être pensé comme un $\bar{\mathbf{v}} \delta t$ avec $\delta t \sim t_{\text{dyn}}$), cette particule de fluide n'a a priori aucune raison de se trouver en équilibre thermodynamique avec l'extérieur. Cet équilibre nécessiterait un ajustement thermique par conduction sur une durée $\delta t \sim t_{\text{cond}}$ avec

$$t_{\text{cond}} \sim \frac{\delta \epsilon}{\delta |\mathbf{Q}^R|} \sim \frac{Nk}{\lambda V} |\boldsymbol{\xi}| \quad (2.3.33)$$

qui peut largement dominer t_{dyn} pour un déplacement sur une échelle macroscopique. Pour la même raison, on peut envisager que la composition de cette particule diffère de celle de son entourage puisque $t_{\text{diff}} \sim t_{\text{nuc}} \gg t_{\text{dyn}}$. Cette différence de propriétés avec l'extérieur fait notamment que la particule ne subit pas la même force (en l'occurrence, la gravité) que son entourage. S'il était à l'équilibre à sa position initiale, l'élément de fluide ressent de façon relative

22. On entend par là une condition du type $|\boldsymbol{\xi}| \ll V^{1/3}$.

une nouvelle force en $\mathbf{r} + \boldsymbol{\xi}$, \mathbf{F}' ²³. Ainsi, si cette force tend à s'exercer dans la même direction que le déplacement, c'est à dire $\boldsymbol{\xi} \cdot \mathbf{F}' \geq 0$ ²⁴, ce dernier va avoir tendance à se perpétuer. Ce cas est donc propice aux mouvements de grande échelle puisque chaque déplacement est encouragé par la stratification du milieu. Lorsque ce cas se présente, on parle d'instabilité dynamique. À l'opposé, si $\boldsymbol{\xi} \cdot \mathbf{F}' < 0$, on voit que ces déplacements ont tendance à s'amortir : on parle de stratification stable. Ce phénomène est illustré en FIGURE 2.3.

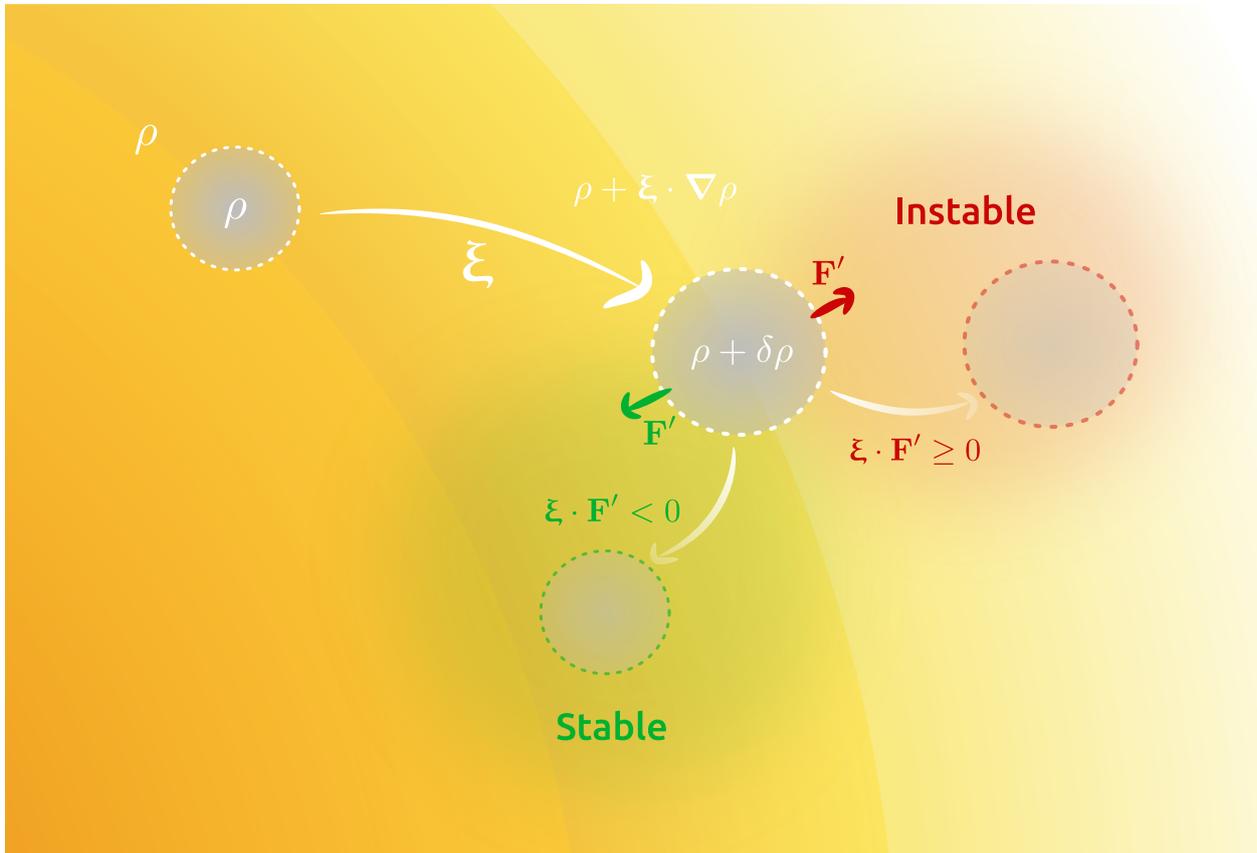


FIGURE 2.3 – Illustration du critère de stabilité sur la stratification stellaire.

On peut chercher formellement une condition à la caractérisation de cette stratification. Au cours du déplacement, des changements vont s'opérer dans la pression, la densité ou la gravité, et ce aussi bien à l'intérieur de la particule de fluide qu'à l'extérieur. Les variations externes s'expriment facilement, elles sont uniquement le résultat du changement de position et peuvent être approximées par $\boldsymbol{\xi} \cdot \nabla p$, $\boldsymbol{\xi} \cdot \nabla\rho$ et $\boldsymbol{\xi} \cdot \nabla\Phi$, au premier ordre. Notons qu'il s'agit d'une propriété générale concernant le potentiel de gravité, qu'on soit à l'intérieur ou à l'extérieur de la particule, et donc $\delta\Phi = \boldsymbol{\xi} \cdot \nabla\Phi$. De la même façon, il est absolument nécessaire que la pression soit continue à l'interface de l'élément déplacé à son arrivée puisque les forces de surface ne se compenseraient pas le cas échéant. Les pressions étant identiques au départ, on a également $\delta p = \boldsymbol{\xi} \cdot \nabla p$. On peut alors écrire la nouvelle force relative comme la différence des forces de

23. On pourrait s'interroger sur ce choix de notation. Comme on le verra par la suite, la notation " δ " désigne la variation d'une grandeur suivie au cours de son mouvement. Elle est donc inadaptée pour référer à une différence de force entre la particule et son entourage, ce que l'on souhaite faire ici. De façon formelle, cette variation peut être vue comme un changement à position fixe mais entre deux instants, avant la présence de la particule déplacée et après son arrivée : $\mathbf{F}' = \mathbf{F}(\mathbf{r} + \boldsymbol{\xi}, t + \delta t) - \mathbf{F}(\mathbf{r} + \boldsymbol{\xi}, t)$. On pourra vérifier que ce choix de notation est bien cohérent avec celui proposé en section 2.4 (voir l'équation (2.4.18)).

24. Il s'agit simplement du travail de cette nouvelle force au cours du déplacement

gravité à l'intérieur et à l'extérieur de la particule :

$$\begin{aligned} \mathbf{F}' &= -(\rho + \delta\rho)\nabla(\Phi + \delta\Phi) + (\rho + \xi \cdot \nabla\rho)\nabla(\Phi + \xi \cdot \nabla\Phi) \\ &= -(\delta\rho - \xi \cdot \nabla\rho)\nabla\Phi \end{aligned} \quad (2.3.34)$$

où l'on a négligé ici les termes d'ordre supérieur (on suppose que le déplacement s'effectue sur un temps court). Il ne reste désormais qu'à exprimer de façon explicite la variation de densité dans l'élément déplacé. Cherchons à expliciter cette variation en fonction des variables d'état usuelles. À l'intérieur de la particule de fluide, cette relation n'est pas unique et dépend en réalité de la transformation thermodynamique que subit l'élément au cours de son déplacement. Une telle transformation peut toujours s'écrire comme une contrainte sur les variables d'états ρ, T et $(X_i)_i$:

$$\mathcal{T}(\rho, T, X_i) = \text{cnst.} \quad (2.3.35)$$

La pression est elle-même fonction de ces variables

$$p = \frac{\rho k T}{m} \quad (2.3.36)$$

puisque m s'exprime facilement en fonction des X_i :

$$m(X_i) \equiv \frac{\rho}{n} = \frac{\rho}{\sum_i n_i} = \frac{\rho}{\sum_i \rho_i / m_i} = \left(\sum_i \frac{X_i}{m_i} \right)^{-1}. \quad (2.3.37)$$

Ainsi, d'après le théorème d'inversion locale, la température peut s'écrire lors de la transformation :

$$T = \mathcal{T}^{-1}(\rho, X_i) \quad (2.3.38)$$

et la variation de pression peut donc être décomposée sur les variables ρ et X_i seulement (à \mathcal{T} fixée) :

$$\delta p = \frac{p}{\rho} \left(\frac{\partial \ln p}{\partial \ln \rho} \right)_{\mathcal{T}, X_i} \delta \rho + p \sum_i \left(\frac{\partial \ln p}{\partial X_i} \right)_{\mathcal{T}, \rho, X_{j \neq i}} \delta X_i. \quad (2.3.39)$$

Plus de précisions sur les précautions à prendre lors de l'écriture d'une telle expression sont disponibles en annexe A.6. On peut illustrer ce que donne cette relation pour 2 exemples :

1. Le cas trivial du transport isotherme ($\mathcal{T} = T = \text{cste}$). On a alors

$$\left(\frac{\partial \ln p}{\partial \ln \rho} \right)_{T, X_i} = 1 \quad \text{et} \quad \forall i, \left(\frac{\partial \ln p}{\partial X_i} \right)_{T, \rho, X_{j \neq i}} = \frac{m}{m_i}$$

ce qui permet de retrouver la relation attendue :

$$\delta p = \frac{p}{\rho} \delta \rho - \frac{p}{m} \delta m. \quad (2.3.40)$$

2. Un cas plus intéressant correspond à celui d'un mouvement de convection, dont on avait vu qu'il était isentropique ($\mathcal{T} = s = \text{cste}$). Il n'est pas possible de trouver une expression explicite pour les dérivées partielles sans une expression de l'entropie du fluide. On peut néanmoins introduire le premier exposant adiabatique Γ_1 :

$$\Gamma_1(\mathbf{r}) \equiv \left(\frac{\partial \ln p}{\partial \ln \rho} \right)_{s, X_i} \quad (2.3.41)$$

et une perturbation sans dimension (moins usuelle) liée à une éventuelle variation d'abondance le long du transport $\delta\chi_s$. On pourra ainsi écrire de façon compacte :

$$\delta p = \frac{\Gamma_1 p}{\rho} (\delta\rho + \rho\delta\chi_s). \quad (2.3.42)$$

L'expression de $\delta\chi_s$ peut être obtenu en effectuant explicitement le produit des déterminants des matrices jacobiniennes :

$$\begin{aligned} \delta\chi_s(\mathbf{r}) &\equiv \sum_i \left(\frac{\partial \ln p}{\partial \ln \rho} \right)_{s, X_i, X_{j \neq i}}^{-1} \left(\frac{\partial \ln p}{\partial X_i} \right)_{s, \rho, X_{j \neq i}} \delta X_i \\ &= \sum_i \left| \frac{\partial(\ln p, s, X_i, X_{j \neq i})}{\partial(\ln \rho, s, X_i, X_{j \neq i})} \right|^{-1} \left| \frac{\partial(\ln p, s, \ln \rho, X_{j \neq i})}{\partial(X_i, s, \ln \rho, X_{j \neq i})} \right| \delta X_i, \end{aligned} \quad (2.3.43)$$

où $|A|$ désigne le déterminant de la matrice A et $\partial(y_1, \dots, y_n) / \partial(x_1, \dots, x_p)$ la matrice jacobienne des fonctions y_i pour les variables x_j . Les propriétés du déterminant et des matrices jacobiniennes permettent de réécrire cette expression comme :

$$\delta\chi_s = \sum_i \underbrace{\left| \frac{\partial(\ln \rho, s, X_i, X_{j \neq i})}{\partial(X_i, s, \ln \rho, X_{j \neq i})} \right|}_1 \left| \frac{\partial(\ln p, s, \ln \rho, X_{j \neq i})}{\partial(\ln p, s, X_i, X_{j \neq i})} \right| \delta X_i, \quad (2.3.44)$$

puisque $\{\ln \rho, s, X_i, X_{j \neq i}\}$ est issu d'une permutation paire de l'ensemble $\{X_i, s, \ln \rho, X_{j \neq i}\}$. Finalement, l'expression de la variation $\delta\chi_s$ prend la forme :

$$\delta\chi_s = \sum_i \left(\frac{\partial \ln \rho}{\partial X_i} \right)_{s, \rho, X_{j \neq i}} \delta X_i. \quad (2.3.45)$$

Maintenant, puisque $\delta p = \xi \cdot \nabla p$, la condition de stabilité peut être rendue plus explicite en injectant les expressions (2.3.47) et (2.3.39) dans l'équation (2.3.34) pour une transformation donnée. Dans le cas d'un déplacement isentropique, le travail de \mathbf{F}' est donné par :

$$\xi \cdot \mathbf{F}' = - \left[\frac{\rho}{\Gamma_1 p} \xi \cdot \nabla p - \rho\delta\chi_s - \xi \cdot \nabla \rho \right] (\xi \cdot \nabla \Phi). \quad (2.3.46)$$

Cette expression peut être encore retravaillée en remarquant que le gradient de pression s'exprime de façon unique en fonction du gradient de densité²⁵ :

$$\nabla p = \gamma p \frac{\nabla \rho}{\rho} \quad (2.3.47)$$

puisque à chaque valeur de $p(\mathbf{r})$ est associé une valeur de $\rho(\mathbf{r})$. Cette relation peut-être vue comme une généralisation de la relation polytropique (cf. paragraphe 2.3.2) à des exposants $\gamma(\mathbf{r})$ variables :

$$\gamma \equiv \frac{d \ln p}{d \ln \rho} \quad (2.3.48)$$

Finalement, le travail s'écrit :

$$\xi \cdot \mathbf{F}' = \frac{\gamma p}{\rho^2} \left[\left(\frac{\gamma}{\Gamma_1} - 1 \right) \xi \cdot \nabla \rho - \rho\delta\chi_s \right] (\xi \cdot \nabla \rho) \quad (2.3.49)$$

25. On peut vérifier que les deux gradients sont alignés à l'aide de l'équation (2.3.31). En effet, le rotationnel de cette équation donne $\nabla \times (\nabla p / \rho) = \mathbf{0}$ et donc $\nabla p \times \nabla \rho = \mathbf{0}$: isodensités et isobares sont confondues. Cette propriété se vérifie tant que $\nabla p / \rho$ dérive d'un potentiel, qu'il soit purement gravitationnel ou centrifuge.

et la condition de stabilité prend la forme

$$\xi \cdot \mathbf{F}' < 0 \quad \Rightarrow \quad \left[\left(\frac{\gamma}{\Gamma_1} - 1 \right) \xi \cdot \nabla \rho - \rho \delta \chi_s \right] (\xi \cdot \nabla \rho) < 0. \quad (2.3.50)$$

2.3.6 Les régions stellaires

Le critère dérivé plus haut peut être vu comme une généralisation du critère de stabilité de Ledoux (LEDoux et WALRAVEN, 1958) pour des mouvements de direction arbitraire au sein de la structure. Pour interpréter l'inégalité (2.3.50) intuitivement, on pourra commencer par se placer dans le cas où $\delta \chi_s = 0$. Puisque la quantité $(\xi \cdot \nabla \rho)^2$ est évidemment positive (ou nulle), l'inégalité prend la forme simple $\Gamma_1 > \gamma$, qui n'est autre qu'une réécriture du critère de Schwarzschild (SCHWARZSCHILD, 1958). Ainsi, toute région dont l'exposant effectif γ est inférieur au premier exposant adiabatique Γ_1 verra les déplacements de matière amortis. Ces dernières ne sont donc pas propices à des mouvements convectifs, et l'on pourra considérer que le système d'équation (2.3.30)-(2.3.32) décrit convenablement la structure de ces régions et leur évolution. On voit ainsi que le transport de l'énergie n'est porté que par \mathbf{Q}^R et ne s'effectue donc que par radiation (et éventuellement par conduction pour de très fortes opacités, auquel cas $\mathbf{Q}^R \rightarrow 0$). De telles régions sont dès lors appelées *régions radiatives*.

À l'inverse, si $\Gamma_1 \leq \gamma$, l'hypothèse selon laquelle $\bar{\mathbf{v}} = \mathbf{0}$ ne se justifie pas puisque les mouvements de convection sont possibles (et même encouragés dans le cas d'une inégalité stricte)²⁶. Une telle région porte le nom de *région convective* et doit donc être étudiée d'un point de vue dynamique. À la vue des ordres de grandeurs des temps de fusion et de diffusion comparés au temps de convection, ce dernier phénomène l'emporte largement sur les deux autres pour ce qui est de la répartition des abondances X_i . On pourra considérer une telle région comme totalement homogène en termes de composition, car constamment mélangée par le champ $\bar{\mathbf{v}}$. De façon intéressante, on peut alors s'intéresser aux abondances sur l'ensemble de la région puisque leur valeur moyenne et locale correspondent. À noter que si le phénomène de diffusion peut être totalement négligé, ce n'est pas le cas de la fusion nucléaire dont le bilan sur la région peut être non nul pour un élément i donné. En intégrant l'équation (2.3.28) sur la région convective :

$$\frac{d}{dt} \int_{V^{cv}} \rho X_i dV + \oint_{\partial V^{cv}} \rho X_i \bar{\mathbf{v}} \cdot d\mathbf{S} = \int_{V^{cv}} m_i \sum_{j \neq i} (r_{ji} - r_{ij}) dV \quad (2.3.51)$$

Ici, le premier terme n'est autre que la variation de masse totale de l'élément i dans la région convective puisque :

$$\int_{V^{cv}} \rho X_i dV = m_i \int_{V^{cv}} n_i dV = m_i N_i^{cv} = M_i^{cv}. \quad (2.3.52)$$

Le second terme correspond au flux de masse de l'élément i entrant dans la région et le terme de droite à l'ensemble de la masse créée (ou détruite) par fusion nucléaire (on le notera \dot{M}_i^{nuc}). Maintenant, il est clair que la composition moyenne²⁷ $\bar{X}_i = M_i^{cv} / M^{cv}$ avec $M^{cv} \equiv \int_{V^{cv}} \rho dV$ et l'équation sur la composition moyenne prend la forme :

$$\frac{d}{dt} (M^{cv} \bar{X}_i) + \oint_{\partial V^{cv}} \rho X_i \bar{\mathbf{v}} \cdot d\mathbf{S} = \dot{M}_i^{\text{nuc}}. \quad (2.3.53)$$

On peut sommer toutes ces équations sur i pour obtenir la variation de l'ensemble de la masse de la région :

$$\frac{dM^{cv}}{dt} + \oint_{\partial V^{cv}} \rho \bar{\mathbf{v}} \cdot d\mathbf{S} = \sum_i \dot{M}_i^{\text{nuc}} \simeq 0 \quad (2.3.54)$$

26. Remarquons que, d'après la définition (2.3.41), une condition suffisante pour cela est de se situer dans une région adiabatique ($ds = 0$) et d'abondances homogènes ($dX_i = 0$). Dans un tel cas, l'exposant effectif correspond exactement au premier exposant adiabatique : $\Gamma_1 \equiv \left(\frac{\partial \ln p}{\partial \ln \rho} \right)_{s, X_i} = \frac{\partial \ln p}{\partial \ln \rho} \equiv \gamma$, violant ainsi le critère de stabilité (2.3.50).

27. On gardera à l'esprit qu'il s'agit ici d'une moyenne sur l'espace physique et non pas sur l'espace des vitesses.

car $\sum_i \bar{X}_i = \sum_i X_i = 1$ par définition. Le bilan global de la destruction de masse par fusion peut être pris comme approximativement nul, car négligeable devant les variations individuelles. Typiquement, on avait vu pour la chaîne de fusion pp1 que :

$$\left| \sum_i \dot{M}_i^{\text{nuc}} \right| \sim 0.007 \dot{M}_{\text{He}}^{\text{nuc}}. \quad (2.3.55)$$

La variation des abondances moyennes au cours du temps s'obtient alors de façon explicite puisque $d(M^{\text{cv}} \bar{X}_i)/dt = M^{\text{cv}} d\bar{X}_i/dt + \bar{X}_i dM^{\text{cv}}/dt$:

$$M^{\text{cv}} \frac{d\bar{X}_i}{dt} + \oint_{\partial V^{\text{cv}}} \rho (X_i - \bar{X}_i) \bar{\mathbf{v}} \cdot d\mathbf{S} = \dot{M}_i^{\text{nuc}}. \quad (2.3.56)$$

En plus de la variation d'abondance assez intuitive due aux réactions de fusion dans la région, on voit apparaître un second terme lié au flux de masse en frontière de zone convective. On peut remarquer la chose suivante : de l'autre côté de la frontière, le champ de vitesse est progressivement amorti puisque $\gamma < \Gamma_1$ par définition. On voit donc l'apparition d'une région de transition entre les zones convectives et radiatives appelée *région d'overshoot*. Dans cette zone, le champ de vitesse s'annule progressivement au fur et à mesure que l'on atteint la zone radiative. Cette région participe bien aux variations d'abondances dans les zones radiatives et convectives et l'évaluation de son étendue peut se faire de nombreuses manières (bien qu'elles ne concordent pas toujours, cf. [RENZINI, 1987](#)). En plus de cette vitesse de pénétration, on trouve aussi la vitesse relative par rapport à la frontière elle-même qui n'est généralement pas statique. L'impact sur les abondances est le suivant : si la région convective progresse sur la région radiative ($\bar{\mathbf{v}} \cdot d\mathbf{S} < 0$, la matière passe la frontière de l'extérieur vers l'intérieur), alors la région convective sera enrichie en élément i ($d\bar{X}_i/dt > 0$) seulement si l'abondance à l'extérieur X_i est supérieure à l'abondance moyenne de la région \bar{X}_i . L'effet sera évidemment l'inverse si la région radiative est plus pauvre en élément i que la région convective.

À cette échelle, on peut aussi s'interroger sur l'importance d'autres phénomènes physiques qui tendrait à dissiper ce champ de vitesse comme la viscosité dynamique. On peut également lui associer une échelle de temps caractéristique :

$$t_{\text{visc}} \sim \frac{Nm}{\mu V^{1/3}} \quad (2.3.57)$$

en fonction du coefficient de viscosité μ . Une estimation de cette échelle pour le soleil donne $t_{\text{visc}} \sim 10^{18} t_{\text{dyn}}$, soit plus que l'âge de l'univers ([KIPPENHAHN, WEIGERT et WEISS, 2012](#)). On pourra donc considérer que la viscosité dynamique a un impact négligeable²⁸ dans les milieux stellaires et que $\Sigma = 0$. En ce qui concerne l'énergie, son transport se fait par convection avec le flux $((1/2)\rho \bar{\mathbf{v}}^2 + \rho\varepsilon + \rho\Phi + p) \bar{\mathbf{v}}$ en plus de la radiation (de flux \mathbf{Q}^{R}) si la région convective est proche du centre de l'étoile. Finalement, les équations qui gouvernent ce milieu prennent la forme :

$$M^{\text{cv}} \frac{d\bar{X}_i}{dt} + \oint_{\partial V^{\text{cv}}} \rho (X_i - \bar{X}_i) \bar{\mathbf{v}} \cdot d\mathbf{S} = \dot{M}_i^{\text{nuc}} \quad (2.3.58)$$

$$\frac{\partial}{\partial t} (\rho \bar{\mathbf{v}}) + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} + p\mathbb{1}) = -\rho \nabla \Phi \quad (2.3.59)$$

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho\varepsilon + \rho\Phi \right) + \nabla \cdot \left[\left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho\varepsilon + \rho\Phi + p \right) \bar{\mathbf{v}} + \mathbf{Q}^{\text{R}} \right] = \rho \dot{\varepsilon}_{\text{nuc}}. \quad (2.3.60)$$

Globalement, on voit que les régions radiatives et convectives diffèrent fortement par les phénomènes physiques qui y dominent ainsi que par les échelles de temps sur lesquelles ils

²⁸. C'est beaucoup moins vrai pour la viscosité turbulente (non introduite ici), de l'ordre de $10^5 t_{\text{dyn}}$ ([KIPPENHAHN, WEIGERT et WEISS, 2012](#)).

prennent place. Leur détermination fait donc partie des enjeux fondamentaux de la physique stellaire. Cette tâche est d'autant plus complexe que ces régions ne sont pas statiques comme souligné plus haut. Elles sont libres de se déplacer au cours de la vie d'une étoile, influencées notamment par l'opacité du milieu (qui détermine l'efficacité du transport radiatif) et le taux de production d'énergie par fusion nucléaire (qui influe sur la quantité d'énergie à transporter).

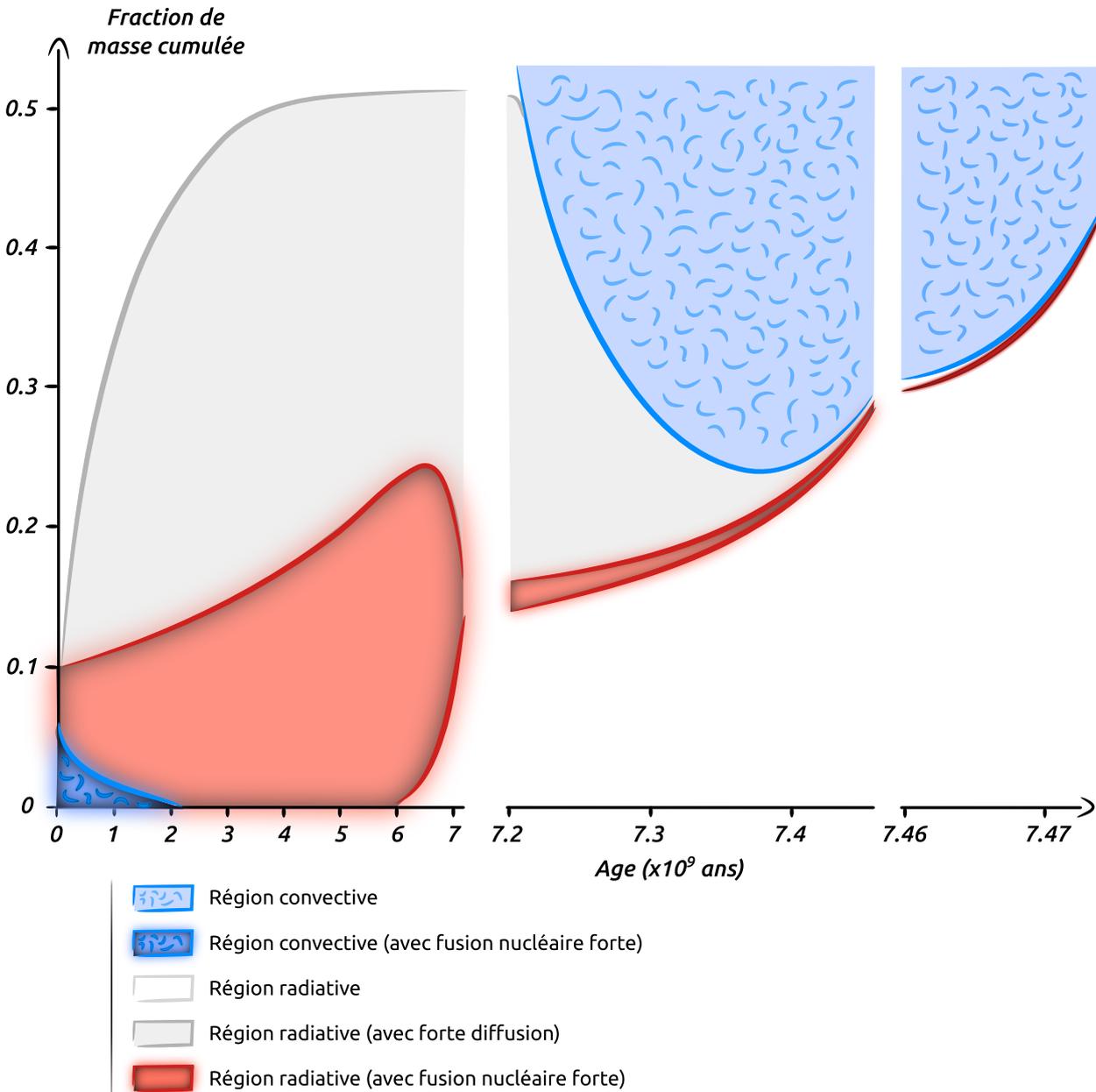


FIGURE 2.4 – Figure adaptée de THOMAS, 1967. Évolution des régions convectives, radiatives et de fusion nucléaire avec l'âge d'une étoile de masse $1.3M_{\odot}$.

Ce phénomène est illustré en FIGURE 2.4 avec le scénario d'évolution d'une étoile légèrement plus massive que le soleil. On voit que la fusion nucléaire commence au cœur de l'étoile, à la fois dans des régions convective et radiative. On y crée de l'hélium à partir d'hydrogène ce qui déséquilibre les abondances relatives de chaque espèce par rapport au reste de l'étoile. Une zone de forte diffusion se crée progressivement : les flux J_H emmènent l'hydrogène vers la surface de l'étoile tandis que J_{He} transporte l'hélium vers le cœur. L'étoile approche la fin de vie lorsqu'il n'y a plus d'hydrogène à "brûler" au centre, la fusion se fait alors en couche au dessus

d'une zone riche en hélium. Au même moment, la zone de diffusion favorise (entre autre) la progression d'une enveloppe convective vers le centre de l'étoile. En effet, on peut montrer dans ce cas que le terme $\rho\delta\chi_s$ s'oppose systématiquement à $\xi \cdot \nabla\rho$ et que leur produit agit comme un terme déstabilisateur dans l'inégalité (2.3.50). On terminera par une dernière remarque relative au déplacement ξ lui-même dans le critère de stabilité. Dans le cas d'une dissipation cinétique négligeable, aucune force ne s'oppose à un déplacement le long d'une isodensité (ou d'une isobare), ce qui s'illustre par le fait que la condition (2.3.50) est toujours insatisfaite si $\xi \cdot \nabla\rho = 0$. On pourrait donc s'attendre à des mouvements perpétuels sur ses surfaces. En pratique, ces courants sont dissipés par l'action de la turbulence sur cette surface (ZAHN, 1992).

2.3.7 L'ionisation

On va maintenant s'intéresser à un phénomène lié à la composition qui sera central dans ce manuscrit. En plus d'une distinction des particules en différents éléments, il est possible de prendre en considération leur état d'ionisation. Ainsi, une espèce i donnée (on les supposera ordonnées par numéro atomique : $Z_i = i$) peut être trouvée dans un état d'ionisation r si r électrons lui ont été retirés. On a donc évidemment $\forall i \in \{1, \dots, N_{\text{sp}}\}$, $0 \leq r \leq i$. En plus de cette partition, les électrons libérés dans le milieu deviennent une espèce à considérer ; on dénotera leurs propriétés avec l'indice e . D'un système composé d' N_{sp} éléments distincts, la prise en compte de l'ionisation nous impose l'étude de $N_{\text{sp}}(N_{\text{sp}} + 3)/2 + 1$ espèces différentes. En considérant que l'ionisation prend place dans une région convective (et l'on va donner des arguments en ce sens par la suite), le système de $N_{\text{sp}} + 2$ équations (2.3.58)-(2.3.60) (l'équation de continuité y figure implicitement) ne suffira pas à décrire l'évolution du système. Il nous faut trouver des équations supplémentaires régissant les interactions entre espèces ionisées.

Commençons par introduire les fractions en espèces ionisées et en électrons de la façon suivante :

$$x_i^r \equiv \frac{n_i^r}{n_i} = x_i \frac{n_i^r}{n} \quad \text{et} \quad x_e \equiv \frac{n_e}{n}. \quad (2.3.61)$$

Cette définition peut sembler un peu curieuse à première vue (il est notamment possible que $x_e > 1$), il semble plus intuitif d'introduire un nombre total de particules $n_{\text{tot}} = n + n_e$ et de définir toutes les fractions de la même façon par n_α/n_{tot} . La convention (2.3.61) nous propose à la place de décrire le système par $N_{\text{sp}}(N_{\text{sp}} + 5)/2 + 1$ fractions (au lieu de $N_{\text{sp}}(N_{\text{sp}} + 3)/2 + 1$) : les x_i, x_i^r, x_e , munies de N_{sp} contraintes : $\forall i \in \{1, \dots, N_{\text{sp}}\}$, $\sum_r x_i^r = 1$. Il y a là un avantage notable : puisque $\bar{X}_i = m_i x_i / m$, les équations (2.3.58) suffisent à contraindre les x_i même avec la prise en compte de l'ionisation. Il nous suffit donc de trouver $N_{\text{sp}}(N_{\text{sp}} + 1)/2 + 1$ équations supplémentaires.

En présence d'ionisation, le bilan d'énergie au cours du temps se trouve être modifié. En effet, le gain où la perte d'un électron modifie le niveau d'excitation propre à chaque atome, ce qui rend l'énergie interne dépendante des fractions d'ionisations x_i^r et x_e . Puisque l'ionisation prend place à des échelles de température et de densité où $\dot{\epsilon}_{\text{nuc}} = \nabla \cdot \mathbf{Q}^R \simeq 0$ et que la masse moyenne ne varie pas le long du mouvement²⁹ $Dm/Dt = 0$, on pourra écrire l'équation sur l'énergie (2.3.60) de la façon suivante :

$$\frac{Dm\epsilon}{Dt} = -p \frac{D}{Dt} \left(\frac{1}{n} \right) + \sum_{ir} \left(\frac{\partial m\epsilon}{\partial x_i^r} \right)_{s,\rho,x_i,x_i^s,x_e} \frac{Dx_i^r}{Dt} + \left(\frac{\partial m\epsilon}{\partial x_e} \right)_{s,\rho,x_i^r} \frac{Dx_e}{Dt}. \quad (2.3.62)$$

Le fait d'écrire cette équation en fonction de l'énergie atomique³⁰ $m\epsilon$ peut sembler superflu, cela nous permet cependant d'introduire rigoureusement les potentiels chimiques μ_α dérivés

29. Cela se comprend en écrivant $Dm/Dt = \sum_i m_i Dx_i/Dt$. Sur des échelles où la diffusion et la fusion sont négligeables, on a $Dx_i/Dt = 0$ pour chacun des éléments.

30. À interpréter dans le sens "par atome"

dans l'annexe A.3 :

$$\left(\frac{\partial m\varepsilon}{\partial x_i^r}\right)_{s,\rho,x_i,x_j^s,x_e} = \mu_i^r \quad \text{et} \quad \left(\frac{\partial m\varepsilon}{\partial x_e}\right)_{s,\rho,x_i,x_i^r} = \mu_e. \quad (2.3.63)$$

En terme d'entropie atomique ms , l'équation (2.3.62) prend alors la forme suivante :

$$T \frac{Dms}{Dt} = \sum_{ir} \mu_i^r \frac{Dx_i^r}{Dt} + \mu_e \frac{Dx_e}{Dt}. \quad (2.3.64)$$

Puisqu'à l'équilibre : $Dms/Dt \rightarrow 0$, les potentiels chimiques sont liés les uns aux autres par la relation :

$$\sum_{ir} \mu_i^r \frac{Dx_i^r}{Dt} + \mu_e \frac{Dx_e}{Dt} = 0. \quad (2.3.65)$$

Maintenant, il est clair que les variations en densité des différents niveaux d'ionisation ne sont pas indépendantes. Puisque la réaction d'ionisation s'écrit à tous les niveaux :



il est nécessaire d'avoir (pour une réaction donnée) :

$$\forall 1 \leq i \leq N_{sp}, \quad 1 \leq r \leq i, \quad \frac{Dx_i^r}{Dt} = \frac{Dx_e}{Dt} = -\frac{Dx_i^{r-1}}{Dt}. \quad (2.3.67)$$

L'identité (2.3.65) peut donc être divisée en $N_{sp}(N_{sp} + 1)/2$ relations à satisfaire sur les potentiels chimiques :

$$\forall 1 \leq i \leq N_{sp}, \quad 1 \leq r \leq i, \quad \mu_i^r + \mu_e - \mu_i^{r-1} = 0. \quad (2.3.68)$$

Pour aller plus loin, il nous faut trouver de façon explicite une expression des potentiels chimiques sur les variables d'état. Une telle dérivation est menée dans les annexes A.1 et A.2 et l'on se contentera ici de citer le résultat final pour une espèce α quelconque :

$$\mu_\alpha(T, n_\alpha) = kT \ln \left(\frac{n_\alpha \lambda_\alpha^3}{\mathcal{Z}_\alpha} \right) \quad (2.3.69)$$

avec λ_α la longueur d'onde thermique de De Broglie de cet espèce :

$$\lambda_\alpha^3(T) \equiv \left(\frac{m_\alpha h^2}{2\pi kT} \right)^{3/2} \quad (2.3.70)$$

et \mathcal{Z}_α sa fonction de partition interne. On peut alors réécrire les équations (2.3.68) comme des liens entre les fractions d'ionisation de niveau $r - 1$ et r :

$$\forall 1 \leq i \leq N_{sp}, \quad 1 \leq r \leq i, \quad \frac{x_i^r}{x_i^{r-1}} = \frac{\mathcal{Z}_i^r \mathcal{Z}_e}{n_e \lambda_e^3 \mathcal{Z}_i^{r-1}}. \quad (2.3.71)$$

Les fonctions de partitions sont liées à la distribution des particules sur les différents d'énergie quantiques. En notant χ_i^r l'énergie requise pour ioniser l'espèce i du niveau $r - 1$ au niveau r , la fonction de partition prend la forme suivante :

$$\mathcal{Z}_i^r(T) = \left(\sum_q g_{iq}^r e^{-\varepsilon_{iq}^r/kT} \right) e^{-\sum_{s \leq r} \chi_i^s/kT} \quad (2.3.72)$$

avec ε_{iq}^r les niveaux d'énergie composant la structure fine et g_{iq}^r leur dégénérescence. Pour traiter l'ionisation, il est courant de négliger la différence en énergie entre les états de la structure fine et de les associer au niveau fondamental $\sum_{s \leq r} \chi_i^s$. Pour ce qui est des électrons, on négligera de la même façon la différence d'énergie entre leurs deux états de spin, ce qui donne $\mathcal{Z}_e = 2$.

Finalement, les équations qui gouvernent la répartition des particules sur les différents niveaux d'excitations s'écrivent :

$$\forall 1 \leq i \leq N_{\text{sp}}, 1 \leq r \leq i, \quad \frac{x_i^r}{x_i^{r-1}} = \frac{2g_i^r}{g_i^{r-1}} \frac{1}{n_e \lambda_e^3} e^{-\chi_i^r/kT} \quad (2.3.73)$$

avec $g_i^r \equiv \sum_q g_{iq}^r$. On reconnaît ici les fameuses équations de Saha (SAHA, 1920; SAHA et FOWLER, 1921) qui donnent le rapport d'occupation de deux niveaux consécutifs en fonction des conditions de densité et de température. Ces équations sont au nombre de $N_{\text{sp}}(N_{\text{sp}} + 1)/2$, il nous reste donc une relation à trouver afin de clore le système. Celle-ci s'obtient en supposant le fluide neutre à l'échelle macroscopique, ce qui fixe le nombre d'électrons par atome. En effet, il est nécessaire pour cela que :

$$n_e = \sum_{ir} r n_i^r \quad \Rightarrow \quad x_e = \sum_{ir} r x_i x_i^r. \quad (2.3.74)$$

On parle alors d'électroneutralité.

Compte tenu des valeurs des potentiels d'ionisation de l'hydrogène et de l'hélium, l'ionisation prend place dans des régions de température assez faibles, ce qui en fait un phénomène presque surfacique pour de nombreuses étoiles. On peut regarder ce qu'il induit pour une grandeur physique comme la pression. En prenant en considération l'électroneutralité :

$$p = \left(n_e + \sum_{ir} n_i^r \right) kT = nkT \times \underbrace{\left(1 + \sum_{ir} r x_i x_i^r \right)}_{\text{facteur causé par l'ionisation}}. \quad (2.3.75)$$

Dans les régions (de faible température) où l'ionisation peut être négligée, ce facteur vaut $1 + \sum_{ir} r x_i x_i^r \simeq 1$. On retrouve bien la relation idéale $p = nkT$. Maintenant, si l'on suppose le fluide complètement ionisé : $1 + \sum_{ir} r x_i x_i^r \simeq 1 + \sum_i i x_i = 1 + \langle i \rangle$ avec $\langle i \rangle$ le numéro atomique moyen du fluide. Les étoiles étant constituées en grande majorité d'hydrogène ($i = 1$), $\langle i \rangle$ est légèrement supérieur à 1 en général. On voit donc que l'ionisation a un impact considérable sur la pression, l'augmentant de plus d'un facteur 2 entre le début et la fin du processus ! Puisque $\langle i \rangle$ est une constante dans une région convective, on remarque cependant que la pression continue de varier comme nkT aussi bien en absence d'ionisation que dans un milieu totalement ionisé. La zone où le fluide est partiellement ionisé, appelée *région d'ionisation* constitue la seule exception. Comme évoqué plus haut, le fait que l'ionisation intervienne à des températures peu élevées fait que cette région est assez peu étendue (le milieu est rapidement complètement ionisé). Dans cette région, en plus des variations dues à la densité et à la température, la pression est localement influencée par les augmentations et diminutions des différents x_i^r . Un témoin de ces fluctuations se trouve être le premier exposant adiabatique Γ_1 , qui pourra être schématiquement décomposé comme³¹ :

$$\Gamma_1 \equiv \left(\frac{\partial \ln p}{\partial \ln \rho} \right)_{s, x_i} = \left(\frac{\partial \ln nkT}{\partial \ln \rho} \right)_{s, x_i} + \left(\frac{\partial \ln(1 + \sum_{ir} r x_i x_i^r)}{\partial \ln \rho} \right)_{s, x_i} \quad (2.3.76)$$

Dans cette expression, le second terme s'annule aussi bien en absence d'ionisation qu'après avoir atteint l'ionisation totale. En dehors de la région d'ionisation, le premier exposant adiabatique peut ainsi être considéré comme constant³². À l'opposé, celui-ci varie fortement à l'intérieur de la région. Le second terme étant négatif, on verra que l'ionisation tend à diminuer la valeur de Γ_1 . Notons au passage que cela abaisse le critère (2.3.50), d'où l'hypothèse que les régions

31. L'équation (2.3.76) n'indique qu' $N_{\text{sp}} + 1$ conservations sur les dérivées partielles, la rendant assez ambiguë puisqu'il y a en réalité $N_{\text{sp}}(N_{\text{sp}} + 3)/2 + 3$ variables d'états. Les $N_{\text{sp}}(N_{\text{sp}} + 1)/2 + 1$ conservations implicites lors des dérivations peuvent être trouvées dans l'annexe A.5.3.

32. La relation $n(\rho, s, x_i) = \rho/m = \rho/\sum_i m_i x_i$ est évidemment fixe et la relation $T(\rho, s, x_i)$ est également inchangée même lorsque le transfert radiatif se renforce (cf. (A.1.18)). Cette dernière ne doit pas être confondue avec la relation $T(\rho, p, x_i)$ qui, elle, dépend fortement du rayonnement (cf. (2.2.96)).

d'ionisations se situent souvent dans des régions convectives; elles favorisent en réalité leur apparition.

On pourra donc retenir que l'ionisation induit des changements de structures abrupts dans l'étoile, et ce sur une étendue limitée. Cette propriété constitue le point de départ des chapitres 3, 4 et 5, ainsi que de mon travail de thèse.

2.3.8 La modélisation stellaire

On dispose à ce stade de deux systèmes d'équations pour représenter les intérieurs stellaires. Dans les régions radiatives de l'étoile, l'évolution de la structure peut être décrite par les équations (2.3.30)-(2.3.32). À l'opposé, les zones convectives sont soumises au système d'équations (2.3.58)-(2.3.60), auquel on pourra ajouter les équations de Saha si l'on se situe dans la région d'ionisation. L'appartenance à un régime ou l'autre peut être établie à l'aide du critère (2.3.50). Si l'on cherche maintenant à résoudre ces équations, on se rend compte que de nombreuses informations restent manquantes. Par exemple : comment déterminer les opacités nécessaires au calcul du transport radiatif? Ou comment évaluer les taux de fusion nucléaire des espèces présentes dans l'étoile?

Répondre à ces questions tient du domaine de la modélisation stellaire. En l'occurrence, ce pan de la discipline est qualifié de modélisation de la *micro-physique*. Il est question d'étudier les interactions fondamentales des particules au niveau microscopique afin d'en dériver les modèles s'appliquant aux équations dynamiques. La détermination des opacités du milieu par exemple, nécessite au préalable la connaissance de l'équation d'état du milieu $p(\rho, T, X_i)$. Si l'on ne s'est limité dans cette section qu'à une description idéale du fluide, il est possible de modéliser les interactions entre particules de façon bien plus exacte en incluant notamment les effets coulombiens et quantiques. L'équation d'état peut ainsi faire intervenir la dégénérescence électronique à haute densité, le potentiel de Debye-Hückel ou encore la pression d'échange (ROGERS et IGLESIAS, 1994; ROGERS, SWENSON et IGLESIAS, 1996). Cette micro-physique est également requise pour décrire l'ionisation dans les milieux denses où les équations de Saha ne s'appliquent plus. Les interactions coulombiennes et quantiques tendent dans ce cas à abaisser les potentiels d'ionisation et requièrent l'introduction de corrections (VEDENOV et LARKIN, 1959; ABE, 1959; LARKIN, 1960). La dérivation précise des taux de fusion nucléaire nécessite, quant à elle, la détermination des sections efficaces de réaction et ainsi la connaissance du S-factor et des énergies de résonance (ANGULO et al., 1999).

Lorsque la micro-physique est modélisée, nos systèmes d'équations peuvent être *techniquement* résolus. Dans les faits, un obstacle manifeste reste les échelles de temps dynamique et d'évolution, lesquelles semblent inconciliables. Là où une description du champ de vitesse nécessiterait une étude du système (2.3.58)-(2.3.60) par pas de temps $\delta t \ll t_{\text{dyn}}$, son évolution requerrait une étude sur un temps $t_{\text{nuc}} \gg \gg t_{\text{dyn}}$! Ainsi, une fois la micro-physique établie, une part au moins aussi importante de la modélisation stellaire consiste à proposer des modèles de *macro-physique*. Puisqu'il est impossible d'inclure le champ de vitesse \bar{v} à l'échelle de l'évolution, il nous faut trouver des recettes (*prescriptions* [en]) pour modéliser son effet. Un premier défi est alors de rendre compte du flux d'énergie $((1/2)\rho \bar{v}^2 + \rho \epsilon + \rho \Phi + p) \bar{v}$ transporté par ce champ de vitesse. On trouve ainsi des modèles de la contribution convective \mathbf{Q}^{cv} à ce flux en fonction des variables d'états : ce sont les modèles MLT (*Mixing Length Theory*, COX et GIULI, 1968; GOUGH et WEISS, 1976) et FST (*Full Spectrum Turbulence*, CANUTO et MAZZITELLI, 1991; CANUTO, GOLDMAN et MAZZITELLI, 1996)). Le premier dérive une expression du flux convectif pour une longueur caractéristique de tourbillon donnée (qui joue alors le rôle de paramètre libre) alors que le second inclut toutes les échelles de la turbulence. En plus de l'aspect énergétique de la convection, il se pose la question de son impact sur la composition, en particulier dans les régions d'overshoot. Divers modèles de la pénétration convective peuvent être obtenus (RENZINI, 1987),

menant à des capacités de mélange variables dans ces régions. On a vu avec le critère (2.3.50) qu'une région avec une structure stable ($\Gamma_1 < \gamma$) pouvait tout de même entretenir des mouvements de convection selon son gradient de composition. Ce phénomène de *semi-convection*³³ fait également partie des éléments participant au mélange et à inclure dans les modèles stellaires (CHARBONNEL et ZAHN, 2007; NOELS et al., 2010). Enfin, si le flux de diffusion peut être explicitement inclus (car s'effectuant sur des temps $t_{\text{diff}} \sim t_{\text{nuc}}$), le modèle choisi pour ses coefficients peut avoir des conséquences considérables sur l'évolution de l'étoile. En particulier, l'inclusion des accélérations radiatives impacte fortement la structure des étoiles plus massives et s'oppose à l'appauvrissement en élément lourds en surface dû à la sédimentation gravitationnelle (DEAL et al., 2018).

Point jusqu'ici éludé, un champ de vitesse à grande échelle ne résulte pas nécessairement de mouvements convectifs. Lorsque ce champ a pour origine la rotation de l'astre, il est possible de montrer qu'une structure statique ne peut satisfaire l'équation du transport de l'énergie (2.3.32). Les isothermes et isopotentiels ne sont généralement pas confondues, ce qui fait qu'il est impossible de compenser le flux radiatif (VON ZEIPPEL, 1924; EDDINGTON, 1925). Une nouvelle composante dynamique doit alors être modélisée, il s'agit de la circulation méridionale (ZAHN, 1992). Ce mouvement à grande échelle se développe donc dans les milieux radiatifs et transporte à la fois les éléments (l'effet est souvent assimilé à de la diffusion, même si le transport avec des mouvements ascendants et descendants en fait un phénomène plus complexe CHABOYER et ZAHN, 1992; MAEDER, 2009) et du moment cinétique. Sa prise en compte est particulièrement complexe, notamment à cause de son interaction avec d'autres phénomènes de transport induits par la rotation différentielle comme les instabilités de Solberg-Høiland (ENDAL et SOFIA, 1976; ENDAL et SOFIA, 1978) ou le mélange par cisaillement (*shear mixing* [en]) (ZAHN, 1992; PRAT et LIGNIÈRES, 2014). En plus de ces effets impactant directement la composition dans l'étoile, les modélisations les plus complètes tiennent aujourd'hui compte de processus agissant sur le champ de vitesse et qui n'impactent qu'indirectement la structure (AMARD et al., 2019). On pourra ainsi citer le transport de moment cinétique par les ondes de gravité internes (ZAHN, TALON et MATIAS, 1997; PINÇON et al., 2017) ou encore le freinage causé par le champ magnétique (lui même couplé à la rotation par effet dynamo, EKSTRÖM et al., 2012).

Évidemment, la prise en compte de tant d'effets en simultané (et de leurs couplages!) introduit au passage de nombreux paramètres libres dans les équations. Non contraints, ils constituent autant d'inconnues dans la physique interne et donc de possibilités de coller aux observations sans apporter de réelles contraintes sur les processus sous-jacents. Cela justifie la recherche de données toujours plus précises en physique stellaire ou de simulations pouvant contraindre certains paramètres de façon indépendante. À cet égard, on verra dans les sections 2.4 et 2.5 que l'astérosismologie fournit des outils puissants pour obtenir des informations sur l'ensemble de la structure interne.

2.3.9 Conditions aux limites

Avant cela, il reste une dernière chose à faire avant d'obtenir des solutions à nos équations : imposer des conditions aux limites. Le problème possédant à la fois un aspect spatial et temporel, il est techniquement nécessaire de fournir des fonctions spatiales au premier instant t_0 pour le résoudre. En pratique, on peut néanmoins supposer que la structure satisfait déjà au moins l'équilibre dynamique, ce qui fait que l'on pourra choisir d'imposer seulement des valeurs en surface (ou au centre) pour ces grandeurs (KIPPENHAHN, WEIGERT et WEISS, 2012). On pourra notamment choisir un rayon R pour l'étoile, une masse M définie comme :

$$M \equiv \int_V \rho dV \quad (2.3.77)$$

33. Aussi appelée convection thermohaline

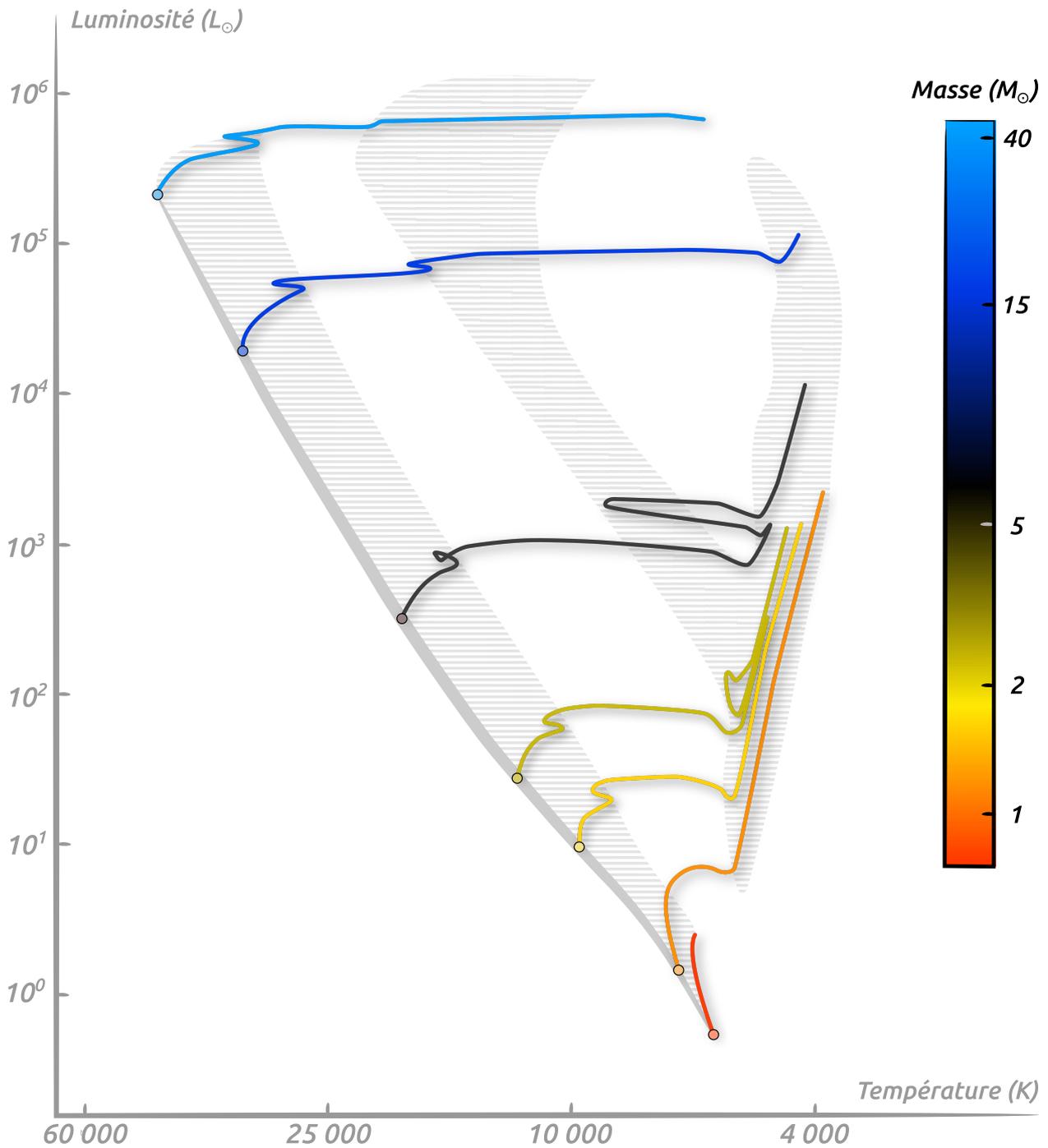


FIGURE 2.5 – Figure adaptée de SCHALLER et al., 1992. Trajectoires évolutives de modèles stellaires entre 0.8 et $40M_{\odot}$ dans un diagramme Hertzsprung–Russell (HR). Tous ces modèles possèdent la même composition initiale $X_i(t_0)$ et illustrent la diversité de structures et d'évolutions possibles en fonction de la masse choisie. Les zones hachurées correspondent aux phases de fusions nucléaires lentes, et où l'étoile se trouve donc à l'équilibre énergétique.

et une luminosité L avec

$$L \equiv \oint_{\partial V} \mathbf{Q}^R \cdot d\mathbf{S}. \quad (2.3.78)$$

Ayant choisi un rayon, on peut imposer de façon équivalente une température effective T_{eff} définie par :

$$4\pi R^2 \sigma T_{\text{eff}}^4 \equiv L. \quad (2.3.79)$$

À noter que le choix du modèle d'atmosphère impose une relation entre la masse et le rayon si l'on veut satisfaire la continuité des pressions en surface. À masse fixée, on dispose donc de deux degrés de liberté pour définir l'étoile représentés par le choix du doublet (R, L) ou (R, T_{eff}) . Il est également possible d'imposer une masse variable au cours de l'évolution $M(t)$ si l'on définit un modèle de perte de masse \dot{M} . Au centre, il est clair que les rayon, masse et luminosité s'annulent et les conditions de densité et de température (la pression en découle par l'équation d'état) résultent indirectement des conditions imposées en surface. Pour les équations qui portent sur les abondances, il est cette fois nécessaire de connaître le mélange initial (*initial mixture* [en]) en tout point³⁴, $X_i(t_0, \mathbf{r})$, avant d'intégrer les équations.

Il est clair que le choix des conditions aux limites pour ces équations influe considérablement sur le type de structure qui en sera solution (cf. FIGURE 2.5). Leur détermination précise est ainsi un enjeu central de la physique stellaire, on qualifiera ici ces grandeurs $(R, M, L, X_i(t_0))$ de paramètres fondamentaux.

34. Elle est généralement supposée homogène pour limiter ce choix de fonction à celui d'une constante : $X_i(t_0)$.

2.4 Les perturbations stationnaires et l'astérosismologie

Contenu de la section

2.4.1	Structure à l'équilibre	63
2.4.2	Perturbations eulériennes et lagrangiennes	65
2.4.3	Linéarisation des équations	66
2.4.4	Perturbations adiabatiques	68
2.4.5	Oscillations stationnaires	69
2.4.6	Décomposition radiale et angulaire des perturbations	71
2.4.7	Conditions aux limites en surface	72
2.4.8	Géométrie particulière du champ de déplacement	73
2.4.9	Séparabilité des solutions	74
2.4.10	L'équation d'onde	76
2.4.11	Analyse JWKB	78
2.4.12	Relation asymptotique des fréquences propres	80

Dans la section précédente, nous avons vu de quelles équations la structure stellaire était solution sur des échelles de temps atteignant celle de l'évolution. Dans un tel cadre, il était nécessaire de proposer des modèles des effets dynamiques, dont la description fine reste incompatible avec des échelles de temps aussi grandes. Dans cette section, on va se pencher sur des solutions particulières des équations sur des échelles de temps dynamiques : les ondes stationnaires. On précise que, si l'on s'attardera sur les propriétés de ces oscillations, la question de leur origine ne sera pas abordée ici (on pourra se référer à [SAMADI, BELKACEM et SONOI, 2015](#) pour une revue approfondie des mécanismes d'excitation).

2.4.1 Structure à l'équilibre

Commençons par considérer le système d'équations dynamiques suivant :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) = 0 \quad (2.4.1)$$

$$\frac{\partial}{\partial t} (\rho \bar{\mathbf{v}}) + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} + p \mathbb{1}) = -\rho \nabla \Phi \quad (2.4.2)$$

$$\frac{\partial}{\partial t} \left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi \right) + \nabla \cdot \left[\left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi + p \right) \bar{\mathbf{v}} + \mathbf{Q}^R \right] = \rho \dot{\varepsilon}_{\text{nuc}}. \quad (2.4.3)$$

où \mathbf{Q}^R désigne la contribution radiative au flux total et où p englobe les contributions cinétiques et radiatives à la pression. On a négligé ici les effets de la diffusion, laquelle agit sur des échelles de temps bien plus grandes et l'on supposera donc que les abondances des différents éléments X_i sont des constantes du temps (elles sont néanmoins libres de varier en fonction de la position dans l'étoile \mathbf{r}). De la même façon, on pourra supposer que, sur cette échelle de temps dynamique et donc dans la suite de la section, le taux de production d'énergie nucléaire volumique est constant : $\rho \dot{\varepsilon}_{\text{nuc}}(t, \mathbf{r}) = \rho \dot{\varepsilon}_{\text{nuc}}(\mathbf{r})$.

Ces équations sont suffisamment générales pour décrire à la fois un mouvement dans une région radiative ou convective, et donc tout à fait adaptées à décrire les perturbations qui y prennent place et qui nous intéresseront par la suite. En plus de potentielles perturbations, on peut chercher à modéliser la structure dans un état d'équilibre, au repos sur l'échelle de temps dynamique. Pour y parvenir, on pourra donc imposer la condition $\partial/\partial t = 0$ au système (2.4.1)-(2.4.3). La première équation ne constitue alors qu'une tautologie et le système se réduit au

couple d'équations considérablement plus simple :

$$\nabla p = -\rho \nabla \Phi \quad (2.4.4)$$

$$\nabla \cdot \mathbf{Q}^R = \rho \dot{\varepsilon}_{\text{nuc}}. \quad (2.4.5)$$

On peut cependant voir une limite à la description de la structure à l'équilibre par ce système : le champ de vitesse $\bar{\mathbf{v}}$ dépend intrinsèquement du temps et possède malgré cela influence considérable sur cette structure à l'équilibre. La solution à ce problème est évoquée dans le paragraphe 2.3.8 : il est possible de modéliser les effets du champ de vitesse en moyenne, en fonction des variables d'état (ρ, T, X_i) . On définira ainsi les modèles :

$$\rho \bar{\mathbf{v}}^{\text{eq}}(\mathbf{r}) \simeq \langle \rho \bar{\mathbf{v}} \rangle_t, \quad (2.4.6)$$

$$p^{\text{eq}}(\mathbf{r}) \simeq \frac{1}{3} \text{Tr} (\langle \rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} \rangle_t), \quad (2.4.7)$$

$$\mathbf{Q}^{\text{eq}}(\mathbf{r}) \simeq \left\langle \left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi + p \right) \bar{\mathbf{v}} \right\rangle_t, \quad (2.4.8)$$

constituant les approximations respectives des flux de masse, de quantité de mouvement³⁵ et d'énergie dus au mouvement en moyenne et inclus dans la structure à l'équilibre (d'où la désignation de ces grandeurs par "eq"). En pratique, p^{eq} modélise la contribution du champ de vitesse turbulent à la pression, là où \mathbf{Q}^{eq} inclut des contributions aussi variées que les flux convectif et turbulent en plus du transport à grande échelle de l'enthalpie, de l'énergie cinétique à grande échelle et turbulente, de l'énergie gravitationnelle et du tenseur de Reynolds (CANUTO, 1997). On peut remarquer que la première équation (2.4.1) s'écrit simplement $\nabla \cdot (\rho \bar{\mathbf{v}}^{\text{eq}}) = 0$ et ne constitue qu'une condition préalable à la conservation de la distribution de masse dans la structure à l'équilibre, typiquement respectée par les mouvements de rotation stationnaire (auquel cas on a $\bar{\mathbf{v}}^{\text{eq}} = \boldsymbol{\Omega} \times \mathbf{r}_\perp$, avec \mathbf{r}_\perp la distance à l'axe de rotation). On mettra de côté cette relation pour nous concentrer sur les deux suivantes par la suite. On redéfinira la vitesse apparaissant dans le système (2.4.1)-(2.4.3) afin qu'elle ne reflète que les mouvements de perturbation (par opposition à ceux inclus dans la structure à l'équilibre) : $\bar{\mathbf{v}} \rightarrow \bar{\mathbf{v}} - \bar{\mathbf{v}}^{\text{eq}}$. À l'inverse, on inclura dans les définitions des pressions et flux les contributions du mouvement en moyenne en procédant aux transformations $p \rightarrow p + p^{\text{eq}}$ et $\mathbf{Q} \rightarrow \mathbf{Q}^R + \mathbf{Q}^{\text{eq}}$ dans ces équations. Les équations de la structure à l'équilibre sont alors celles obtenues lorsque $\bar{\mathbf{v}} = \mathbf{0}$:

$$\nabla p = -\rho \nabla \Phi, \quad (2.4.9)$$

$$\nabla \cdot \mathbf{Q} = \rho \dot{\varepsilon}_{\text{nuc}}, \quad (2.4.10)$$

lesquelles correspondent aux équations (2.4.4) et (2.4.5) avec l'inclusion des pressions et flux à l'équilibre. En y ajoutant de façon explicite l'équation de Poisson :

$$\Delta \Phi = 4\pi G \rho, \quad (2.4.11)$$

le système (2.4.9)-(2.4.11) fait ressortir un problème aux variations temporelles lentes (que l'on supposera même nulles en comparaison de l'échelle dynamique) là où le système (2.4.1)-(2.4.3) expose les fluctuations d'un second problème physique bien plus rapide et complexe. S'il est clair qu'une solution à aux équations (2.4.9)-(2.4.11) ne peut tenir compte des effets engendrés par les variations d'abondances et donc de l'évolution de la structure sur une échelle de temps de l'ordre de $t_{\text{diff}} \sim t_{\text{nuc}}$ (cf. paragraphe 2.3.8), elle permet de définir convenablement ce qu'est une structure à l'équilibre sur un temps $t \sim t_{\text{dyn}}$. Il est important de noter que les définitions des

35. On a supposé son isotropie en moyenne afin de simplifier l'écriture de la pression totale.

quantités considérées apportent des modifications aux équations dynamiques :

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \bar{\mathbf{v}}) = 0 \quad (2.4.12)$$

$$\frac{\partial}{\partial t} (\rho \bar{\mathbf{v}}) + \nabla \cdot (\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}} + \rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}}|_{\text{coupl}} + p \mathbb{1}) = -\rho \nabla \Phi \quad (2.4.13)$$

$$\begin{aligned} \frac{\partial}{\partial t} \left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi + \rho (\bar{\mathbf{v}} \cdot \bar{\mathbf{v}})|_{\text{coupl}} \right) \\ + \nabla \cdot \left[\left(\frac{1}{2} \rho \bar{\mathbf{v}}^2 + \rho \varepsilon + \rho \Phi + p \right) \bar{\mathbf{v}} + \frac{1}{2} \rho (\bar{\mathbf{v}} \cdot \bar{\mathbf{v}})|_{\text{coupl}} \bar{\mathbf{v}} + \mathbf{Q} \right] = \rho \dot{\varepsilon}_{\text{nuc}}. \end{aligned} \quad (2.4.14)$$

où l'on a défini :

$$\rho (\bar{\mathbf{v}} \cdot \bar{\mathbf{v}})|_{\text{coupl}} \equiv 2\rho (\bar{\mathbf{v}} \cdot \bar{\mathbf{v}}^{\text{eq}}), \quad (2.4.15)$$

$$\rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}}|_{\text{coupl}} \equiv \rho \bar{\mathbf{v}}^{\text{eq}} \otimes \bar{\mathbf{v}} + \rho \bar{\mathbf{v}} \otimes \bar{\mathbf{v}}^{\text{eq}}, \quad (2.4.16)$$

$$\rho (\bar{\mathbf{v}} \cdot \bar{\mathbf{v}})|_{\text{coupl}} \bar{\mathbf{v}} \equiv \rho (\bar{\mathbf{v}} \cdot \bar{\mathbf{v}}) \bar{\mathbf{v}}^{\text{eq}} + 2\rho (\bar{\mathbf{v}} \cdot \bar{\mathbf{v}}^{\text{eq}}) (\bar{\mathbf{v}} + \bar{\mathbf{v}}^{\text{eq}}) + \rho (\bar{\mathbf{v}}^{\text{eq}} \cdot \bar{\mathbf{v}}^{\text{eq}}) \bar{\mathbf{v}}. \quad (2.4.17)$$

Les termes supplémentaires qui apparaissent résultent du couplage en le champ de vitesse inclus dans la structure à l'équilibre et celui induit par la perturbation. On négligera par la suite ces termes qui complexifient les développements en utilisant le système (2.4.1)-(2.4.3) pour décrire les mouvements dynamiques. L'hypothèse implicite à cette simplification est la supposition de champs de vitesses faibles à l'équilibre. Notons que cela retire du champ d'étude tous les mouvements cohérents sur la moyenne (2.4.6) et notamment la prise en compte de mouvements de rotation dans la structure à l'équilibre.

2.4.2 Perturbations eulériennes et lagrangiennes

Notre but est d'étudier les solutions des équations dynamiques (2.4.1)-(2.4.3) qui se décomposent schématiquement de la façon suivante : $y = y_{\text{eq}} + y_{\text{dyn}}$ avec $y_{\text{eq}}(\mathbf{r})$ la part de y composant la structure à l'équilibre statique solution de (2.4.9)-(2.4.11) et donc indépendante du temps. L'autre part, $y_{\text{dyn}}(\mathbf{r}, t) = y - y_{\text{eq}}$ correspond alors à la contribution purement dynamique à y , que l'on qualifie de perturbation.

Il convient maintenant de définir un peu plus explicitement ce que l'on entend par perturbation. En fait, il existe au moins deux façons de procéder, correspondant à chacun des formalismes considérés dans le paragraphe 2.2.5. Dans le cadre eulérien, les quantités étant fonctions de (\mathbf{r}, t) , il paraît intuitif de définir la perturbation de toute grandeur comme on l'a fait plus haut :

$$y'(\mathbf{r}, t) \equiv y(\mathbf{r}, t) - y(\mathbf{r}, 0) \quad (2.4.18)$$

où l'on a pris comme origine du temps l'instant où la perturbation est nulle³⁶ : $y(\mathbf{r}, 0) = y_{\text{eq}}(\mathbf{r})$. Notons que, sur cet intervalle de temps t , en plus de la variation perçue à position fixée y' , une particule de fluide subit une variation due à son mouvement conformément à la définition (2.2.15). Ainsi, dans le cadre du formalisme lagrangien, on aura plutôt tendance à définir une perturbation comme suit :

$$\delta y(\mathbf{r}_0, t) \equiv y(\mathbf{r}_0, t) - y(\mathbf{r}_0, 0). \quad (2.4.19)$$

On pourra s'arrêter un temps sur les conventions de notation présente dans cette équation. De manière générale, on notera dans la cadre lagrangien $y(\mathbf{r}_0, t)$ la quantité correspondant à

36. Il n'est pas dit, en pratique, qu'il soit possible de trouver un tel temps. Il paraît en effet probable que des perturbations soient systématiquement présentes dans l'étoile, et donc impossible de désigner un instant spécifique où la structure dans son intégralité est "au repos". Cette origine des temps est donc à comprendre d'un point de vue conceptuel. Comme on le verra par la suite, on s'intéresse ici aux solutions stationnaires. Cette origine des temps peut donc être pensée comme le temps depuis lequel un régime transitoire mène au régime stationnaire observé. Il permet donc de définir clairement ce qu'est la structure à l'équilibre parmi les oscillations : il s'agit de la structure précédant le régime transitoire, aux temps $t < 0$.

$y(\mathbf{r}(\mathbf{r}_0, t), t)$ dans le cadre eulérien. Cette convention permet notamment de faire ressortir les nombreuses symétries entre les deux formalismes. On peut cependant remarquer qu'elle introduit une certaine ambiguïté. En effet, la notation $y(\mathbf{r}_0, t)$ pourrait tout aussi bien correspondre à $y(\mathbf{r}(\mathbf{r}_0, 0), t)$ ou encore $y(\mathbf{r}(\mathbf{r}_0, t), 0)$ dans le cadre eulérien. La première correspond bien en notation eulérienne à $y(\mathbf{r}_0, t)$ et la seconde à $y(\mathbf{r}, 0)$. Notons cependant que ces quantités correspondront bien à des notations *lagrangiennes* distinctes de $y(\mathbf{r}_0, t)$, à savoir $y(\mathbf{r}_1, t)$ (vérifiant $\mathbf{r}(\mathbf{r}_1, t) = \mathbf{r}_0$) et $y(\mathbf{r}, 0)$. Ces grandeurs n'intervenant de toute façon nulle part dans les calculs qui suivent, on pourra adopter la convention présentée en équation (2.4.19) pour en alléger les écritures³⁷.

D'après les définitions (2.4.18) et (2.4.19), il est clair que $\delta y \neq y'$, d'où les deux notations introduites. On utilisera dans la suite des calculs un formalisme ou l'autre à notre convenance, de façon à minimiser les développements évitables. Un premier lien entre ces perturbations peut être établi en remarquant que le déplacement induit sur ce temps t s'écrit :

$$\mathbf{r}(\mathbf{r}_0, t) - \mathbf{r}_0 = \mathbf{r}(\mathbf{r}_0, t) - \mathbf{r}(\mathbf{r}_0, 0) \equiv \delta \mathbf{r}(\mathbf{r}_0, t) \quad (2.4.20)$$

par définition³⁸. Cette quantité intervenant souvent dans les équations, on lui réserve la notation plus compacte $\xi \equiv \delta \mathbf{r}$. En utilisant la relation (2.2.15), on peut dériver une première propriété dans le formalisme lagrangien³⁹ :

$$\xi(\mathbf{r}_0, t) = \mathbf{r}(\mathbf{r}_0, t) - \mathbf{r}_0 = \int_0^t \bar{\mathbf{v}}(\mathbf{r}_0, t') dt' \Rightarrow \left(\frac{\partial \xi}{\partial t} \right)_{\mathbf{r}_0} = \bar{\mathbf{v}}(\mathbf{r}_0, t), \quad (2.4.21)$$

que l'on préférera souvent ré-exprimer dans le cadre eulérien :

$$\frac{D\xi}{Dt} = \bar{\mathbf{v}}(\mathbf{r}, t). \quad (2.4.22)$$

On peut également démontrer les propriétés qui suivent, propres à chaque formalisme :

$$\left(\frac{\partial y}{\partial t} \right)' = \left(\frac{\partial y}{\partial t} \right)_{\mathbf{r}}(t) - \left(\frac{\partial y}{\partial t} \right)_{\mathbf{r}}(0) = \frac{\partial}{\partial t} (y(\mathbf{r}, t) - y(\mathbf{r}, 0)) = \frac{\partial y'}{\partial t} \quad (2.4.23)$$

$$\delta \left(\frac{Dy}{Dt} \right) = \left(\frac{\partial y}{\partial t} \right)_{\mathbf{r}_0}(t) - \left(\frac{\partial y}{\partial t} \right)_{\mathbf{r}_0}(0) = \frac{\partial}{\partial t} (y(\mathbf{r}_0, t) - y(\mathbf{r}_0, 0)) = \left(\frac{\partial \delta y}{\partial t} \right)_{\mathbf{r}_0} = \frac{D\delta y}{Dt} \quad (2.4.24)$$

$$\frac{\partial y(\mathbf{r}, t)}{\partial t} = \left(\frac{\partial y}{\partial t} \right)_{\mathbf{r}}(t) = \frac{\partial}{\partial t} (y(\mathbf{r}, 0) + y'(\mathbf{r}, t)) = \frac{\partial y'}{\partial t} \quad (2.4.25)$$

$$\frac{Dy(\mathbf{r}, t)}{Dt} = \left(\frac{\partial y}{\partial t} \right)_{\mathbf{r}_0}(t) = \frac{\partial}{\partial t} (y(\mathbf{r}_0, 0) + \delta y(\mathbf{r}_0, t)) = \left(\frac{\partial \delta y}{\partial t} \right)_{\mathbf{r}_0} = \frac{D\delta y}{Dt}. \quad (2.4.26)$$

2.4.3 Linéarisation des équations

Pour simplifier les futurs développements, une hypothèse courante est de limiter à l'étude de perturbations de faible amplitude devant la structure à l'équilibre : $y', \delta y \ll y$. Il est alors possible de linéariser toutes les relations obtenues jusqu'ici. On pourra notamment passer d'une perturbation à l'autre en exprimant les coordonnées lagrangiennes dans le formalisme eulérien :

$$\delta y(\mathbf{r}_0, t) = y(\mathbf{r}(\mathbf{r}_0, t), t) - y(\mathbf{r}(\mathbf{r}_0, 0), 0) = y(\mathbf{r}_0 + \xi, t) - y(\mathbf{r}_0, 0) = y'(\mathbf{r}_0, t) + \xi \cdot \nabla y|_{\mathbf{r}_0}. \quad (2.4.27)$$

37. Notons par symétrie qu'il existe cette même ambiguïté autour de la notation $y(\mathbf{r}, t)$ du cadre eulérien sans qu'il soit nécessaire de préciser s'il est question de la quantité $y(\mathbf{r}_0(\mathbf{r}, t), t)$, $y(\mathbf{r}_0(\mathbf{r}, 0), t)$ ou $y(\mathbf{r}_0(\mathbf{r}, t), 0)$ dans le formalisme lagrangien.

38. Notons que l'on a immédiatement $\mathbf{r}' = \mathbf{r} - \mathbf{r} = \mathbf{0}$.

39. L'équation (2.2.15) faisant intervenir le champ de vitesse total, il aurait été nécessaire d'inclure $\bar{\mathbf{v}}^{\text{eq}}$ à l'intégrale ci-contre pour tenir compte d'un mouvement dans la structure à l'équilibre (de rotation par exemple). L'équation (2.4.22) aurait alors donné pour résultat : $D\xi/Dt = \bar{\mathbf{v}} + \bar{\mathbf{v}}^{\text{eq}}$.

Dans cette équation, toutes les quantités sont évalués au même point, on pourra donc l'omettre pour alléger l'écriture : $\delta y = y' + \xi \cdot \nabla y$. On gardera à l'esprit que y n'est pas une fonction du temps ici, il s'agit de sa valeur à l'équilibre statique. En réalité, à l'aide des propriétés (2.4.23)-(2.4.26), il est toujours possible de faire disparaître les $y(\mathbf{r}, t)$ et $y(\mathbf{r}_0, t)$ en faveur des y' et δy . Dans toutes les équations qui suivront, on pourra donc considérer sans ambiguïté une quantité indépendante du temps chaque fois qu'il sera fait mention d'une grandeur y .

Puisque le champ de vitesse s'exprime en fonction d'une grandeur d'ordre 1 (cf. équation (2.4.22)), on écrira de façon générale :

$$\frac{D\delta y}{Dt} \equiv \frac{\partial \delta y}{\partial t} + (\bar{\mathbf{v}} \cdot \nabla) \delta y = \frac{\partial \delta y}{\partial t} \quad (2.4.28)$$

au premier ordre. Appliquée à l'équation (2.4.22) elle-même :

$$\frac{D\xi}{Dt} = \frac{\partial \xi}{\partial t} = \bar{\mathbf{v}}. \quad (2.4.29)$$

On peut maintenant réécrire les équations (2.4.1)-(2.4.3) sous forme linéarisée en remplaçant $\bar{\mathbf{v}}$ par $\partial \xi / \partial t$. L'équation de continuité (2.4.1) s'écrit maintenant :

$$\frac{\partial \rho'}{\partial t} + \nabla \cdot \left(\rho \frac{\partial \xi}{\partial t} \right) = \frac{\partial}{\partial t} (\rho' + \nabla \cdot (\rho \xi)) = 0. \quad (2.4.30)$$

Ayant pris pour convention les perturbations nulles en $t = 0$, cette équation s'intègre sans problème entre 0 et t :

$$\rho' + \nabla \cdot (\rho \xi) = 0. \quad (2.4.31)$$

La propriété (2.4.27) nous donne l'équation sur la perturbation lagrangienne sans plus de difficulté :

$$\delta \rho + \rho \nabla \cdot \xi = 0. \quad (2.4.32)$$

En remarquant que $\bar{\mathbf{v}} \otimes \bar{\mathbf{v}}$ peut être négligé à l'ordre 1 (sous hypothèse de faible couplage, de nombreux termes d'ordre 1 peuvent apparaître dans le cas contraire, cf. équations (2.4.15)-(2.4.17)), l'équation d'Euler (2.4.2) prend la forme :

$$\rho \frac{\partial^2 \xi}{\partial t^2} + \nabla(p + p') = -\rho \nabla \Phi - \rho' \nabla \Phi - \rho \nabla \Phi'. \quad (2.4.33)$$

Dans cette expression, les termes d'ordre 0 s'annulent en vertu de l'équation (2.4.9). On peut également facilement éliminer la variable ρ' avec l'équation (2.4.31) et la variable Φ' avec l'équation de Poisson ($\Delta \Phi' = 4\pi G \rho'$) :

$$\Phi' = - \int_V \frac{G \rho'(\tilde{\mathbf{r}})}{|\mathbf{r} - \tilde{\mathbf{r}}|} d\tilde{\mathbf{r}} = \int_V \frac{G \nabla \cdot (\rho \xi)}{|\mathbf{r} - \tilde{\mathbf{r}}|} d\tilde{\mathbf{r}}, \quad (2.4.34)$$

ce qui mène finalement à

$$\rho \frac{\partial^2 \xi}{\partial t^2} + \nabla p' = \nabla \cdot (\rho \xi) \nabla \Phi - \rho \nabla \int_V \frac{G \nabla \cdot (\rho \xi)}{|\mathbf{r} - \tilde{\mathbf{r}}|} d\tilde{\mathbf{r}}. \quad (2.4.35)$$

Il n'existe pas de lien évident entre les perturbations ξ et p' à cet instant. Pour linéariser la dernière équation (2.4.3), il est en réalité bien plus simple de partir de sa forme convective :

$$T \frac{Ds}{Dt} = \dot{\epsilon}_{\text{nuc}} - \frac{1}{\rho} \nabla \cdot \mathbf{Q}, \quad (2.4.36)$$

Si de nombreux termes apparaissent dans le formalisme eulérien, l'équation qui porte sur la perturbation lagrangienne de l'entropie δs prend une forme compacte à l'aide de la propriété

(2.4.26) :

$$T \frac{D\delta s}{Dt} = \dot{\epsilon}_{\text{nuc}} + \delta \dot{\epsilon}_{\text{nuc}} - \frac{1}{\rho} [\nabla \cdot \mathbf{Q} + (\nabla \cdot \boldsymbol{\xi})(\nabla \cdot \mathbf{Q}) + \nabla \cdot \delta \mathbf{Q}], \quad (2.4.37)$$

où l'on a utilisé l'équation (2.4.32) pour éliminer $\delta\rho/\rho$. Le terme d'ordre 0 s'annule une nouvelle fois avec l'équation statique (2.4.10) :

$$T \frac{D\delta s}{Dt} = \delta \dot{\epsilon}_{\text{nuc}} - (\nabla \cdot \boldsymbol{\xi}) \dot{\epsilon}_{\text{nuc}} - \frac{1}{\rho} \nabla \cdot \delta \mathbf{Q} \quad (2.4.38)$$

Cette équation nous permet de faire une remarque importante. Quand bien même la structure à l'équilibre serait stratifiée de façon isentropique ($\gamma = \Gamma_1$), on voit que la perturbation δs ne s'annule généralement pas. Les deux premiers termes désignent la variation nette⁴⁰ du taux de fusion nucléaire et le dernier la variation du flux au cours du mouvement. On dira dans ce cas que la perturbation dynamique est non-adiabatique⁴¹. Expliciter formellement le contenu des termes de droite (et faire apparaître la relation manquante entre $\boldsymbol{\xi}$ et p') nécessiterait un travail plus poussé (voir UNNO et al., 1989 par exemple) et l'on se limitera ici à l'étude de perturbations strictement adiabatiques.

2.4.4 Perturbations adiabatiques

L'hypothèse d'adiabaticité simplifie considérablement le problème, et tout autant l'allure de la dernière relation recherchée. On retrouve en réalité le cas étudié dans le paragraphe 2.3.5 : celui de la perturbation induite par un déplacement isentropique. La lien entre p' et $\boldsymbol{\xi}$ s'obtient alors combinant la relation (2.3.42) et les propriétés (2.4.27) et (2.4.32) :

$$\begin{aligned} p' &= \delta p - \boldsymbol{\xi} \cdot \nabla p \\ &= \frac{\Gamma_1 p}{\rho} (\delta\rho + \rho \delta\chi_s) - \boldsymbol{\xi} \cdot \nabla p \\ &= -\Gamma_1 p (\nabla \cdot \boldsymbol{\xi} - \delta\chi_s) - \boldsymbol{\xi} \cdot \nabla p. \end{aligned} \quad (2.4.39)$$

Il est assez rare en pratique d'inclure la variation de pression liée aux abondances (terme en $\delta\chi_s$), laquelle demanderait une étude approfondie des équations (2.3.29). L'injection de la relation (2.4.39) dans l'équation du mouvement (2.4.35) mène à :

$$\rho \frac{\partial^2 \boldsymbol{\xi}}{\partial t^2} = \nabla (\Gamma_1 p \nabla \cdot \boldsymbol{\xi} + \boldsymbol{\xi} \cdot \nabla p) + \nabla \cdot (\rho \boldsymbol{\xi}) \nabla \Phi - \rho \nabla \int_V \frac{G \nabla \cdot (\rho \boldsymbol{\xi})}{|\mathbf{r} - \tilde{\mathbf{r}}|} d\tilde{\mathbf{r}}. \quad (2.4.40)$$

L'équation ci-contre ne fait intervenir que le déplacement et lie ses dérivées temporelles (à gauche) à ses dérivées spatiales (à droite). Il est possible de réarranger un peu les termes pour y

40. Au sens où l'on y retranche la variation induite par une dilatation ou contraction au cours du mouvement : $\nabla \cdot \boldsymbol{\xi} = \oint \rho \frac{d(1/\rho)}{dt} dt$.

41. Il est rare de faire la distinction entre adiabaticité et isentropie à cette échelle.

voir plus clair :

$$\begin{aligned}
\frac{\partial^2 \xi}{\partial t^2} &= \frac{1}{\rho} \nabla (\Gamma_1 p \nabla \cdot \xi) - \frac{1}{\rho} \nabla (\rho \xi \cdot \nabla \Phi) + \frac{1}{\rho} \nabla \cdot (\rho \xi) \nabla \Phi - \nabla \int_V \frac{G \nabla \cdot (\rho \xi)}{|\mathbf{r} - \tilde{\mathbf{r}}|} d\tilde{\mathbf{r}} \\
&= \frac{\Gamma_1 p}{\rho} \nabla (\nabla \cdot \xi) + \frac{1}{\rho} \nabla (\Gamma_1 p) \nabla \cdot \xi - \nabla (\xi \cdot \nabla \Phi) \\
&\quad + \frac{1}{\rho} \left[-(\nabla \rho)(\xi \cdot \nabla \Phi) + (\xi \cdot \nabla \rho)(\nabla \Phi) \right] + \nabla \Phi \nabla \cdot \xi - \nabla \int_V \frac{G \nabla \cdot (\rho \xi)}{|\mathbf{r} - \tilde{\mathbf{r}}|} d\tilde{\mathbf{r}} \quad (2.4.41) \\
&= \frac{\Gamma_1 p}{\rho} \nabla (\nabla \cdot \xi) + \left[\nabla \left(\frac{\Gamma_1 p}{\rho} \right) - \left(\frac{\Gamma_1}{\gamma} - 1 \right) \nabla \Phi \right] \nabla \cdot \xi - \nabla (\xi \cdot \nabla \Phi) \\
&\quad - \nabla \int_V \frac{G \nabla \cdot (\rho \xi)}{|\mathbf{r} - \tilde{\mathbf{r}}|} d\tilde{\mathbf{r}}
\end{aligned}$$

où l'on a utilisé le fait qu'isopotentiels et isodensités sont superposées pour éliminer le crochet de la seconde ligne ainsi que l'indice de stratification effective ($\nabla \ln \rho = (1/\gamma) \nabla \ln p$) pour faire disparaître le gradient de densité. Puisque

$$\nabla (\nabla \cdot \xi) = \Delta \xi + \nabla \times (\nabla \times \xi), \quad (2.4.42)$$

le développement (2.4.41) conduit en réalité à une équation de propagation à la vitesse

$$c^2 \equiv \frac{\Gamma_1 p}{\rho} = \left(\frac{\partial p}{\partial \rho} \right)_{s, X_i}. \quad (2.4.43)$$

En plus des termes classiques de l'équation d'onde ($\partial_t^2 \xi - c^2 \Delta \xi = 0$), de nombreux termes antisymétriques par retournement de l'espace figurent à droite de l'équation et trahissent donc des directions de propagation privilégiées. On peut y voir l'action des disparités de vitesse du son ($(\nabla c^2) \nabla \cdot \xi$), le rappel exercé par une stratification non adiabatique du milieu ($(\Gamma_1/\gamma - 1) \mathbf{g} \nabla \cdot \xi$), l'entraînement du mouvement par la gravité ($\nabla (\xi \cdot \mathbf{g})$) et même la rétroaction du champ de gravité causé par le déplacement de masse ($\mathbf{g}' = -\nabla \int_V G \nabla \cdot (\rho \xi) / |\mathbf{r} - \tilde{\mathbf{r}}| d\tilde{\mathbf{r}}$). Une visualisation schématique de ces effets est proposée en FIGURE 2.6.

2.4.5 Oscillations stationnaires

Ayant obtenu une équation d'onde pour la propagation de la perturbation, il est naturel d'en chercher une solution stationnaire⁴². En supposant une pulsation ω pour cette onde, on effectuera la substitution $\xi(\mathbf{r}, t) \rightarrow \xi(\mathbf{r}) e^{i\omega t}$ dans cette équation (la partie réelle de la solution oscille donc périodiquement $\text{Re} \{ \xi(\mathbf{r}, t) \} \propto \cos(\omega t)$). En définissant un opérateur \mathcal{L} comme suit :

$$\mathcal{L} \xi \equiv -c^2 \nabla (\nabla \cdot \xi) - \left[\nabla c^2 + \left(\frac{\Gamma_1}{\gamma} - 1 \right) \mathbf{g} \right] \nabla \cdot \xi - \nabla (\xi \cdot \mathbf{g}) + \nabla \int_V \frac{G \nabla \cdot (\rho \xi)}{|\mathbf{r} - \tilde{\mathbf{r}}|} d\tilde{\mathbf{r}}, \quad (2.4.44)$$

l'équation de propagation (2.4.41) s'écrit simplement comme

$$\mathcal{L} \xi = \omega^2 \xi. \quad (2.4.45)$$

Cette écriture du problème réduit donc l'étude des oscillations stationnaires dans la structure statique à celle des valeurs et vecteurs propres de l'opérateur \mathcal{L} . Du spectre de cet opérateur résulte les fréquences d'oscillations ω^2 des ondes stationnaires tandis que l'espace propre associé fournit les modes d'oscillation $\xi(\mathbf{r})$ de la structure.

42. Notons que cette hypothèse exclut de l'espace des solutions tous les régimes transitoires. Il aurait été difficile d'en tenir compte dans tous les cas sans s'intéresser aux mécanismes d'excitation et faire apparaître explicitement les termes de forçage.

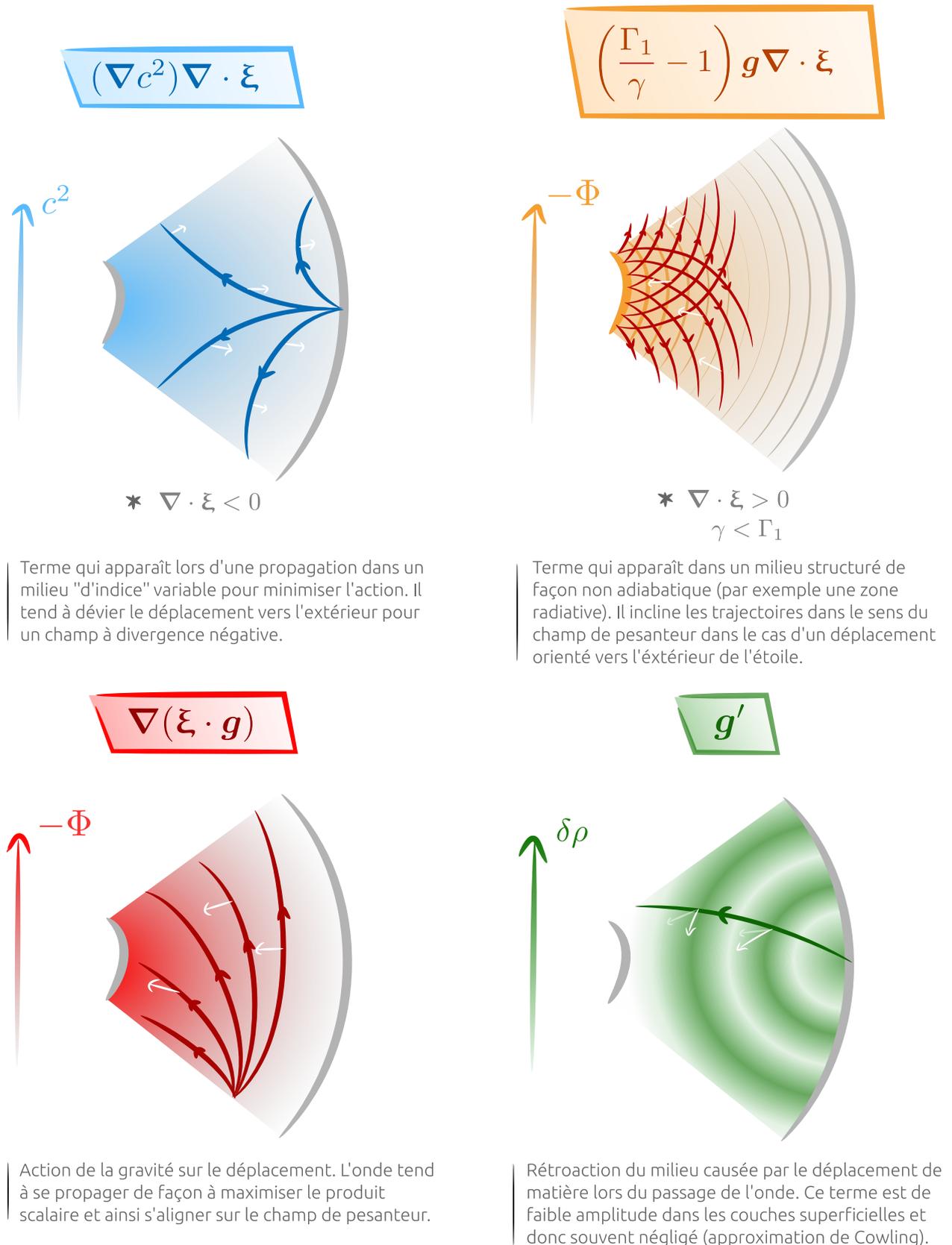


FIGURE 2.6 – Représentation schématique de l'impact des différents termes de l'équation (2.4.41) sur la propagation.

2.4.6 Décomposition radiale et angulaire des perturbations

Commençons par remarquer qu'en l'absence de rotation ou de champ magnétique, la structure à l'équilibre satisfait une symétrie sphérique. On pourra donc écrire les grandeurs à l'équilibre comme des fonctions purement radiales. Afin de pouvoir étudier et caractériser plus facilement leurs perturbations cependant, il est courant d'introduire la décomposition radiale et angulaire des champs radiaux et vectoriels qui interviennent dans les équations. On considérera les harmoniques sphériques normalisées Y_ℓ^m :

$$Y_\ell^m(\theta, \varphi) = (-1)^{\frac{1}{2}(m+|m|)} \sqrt{\frac{(2\ell+1)(\ell-|m|)!}{4\pi(\ell+|m|)!}} P_{\ell,|m|}(\cos\theta) e^{im\varphi} \quad (2.4.46)$$

où $P_{\ell,m}$, les polynômes de Legendre, sont eux-mêmes solutions de

$$-\frac{d}{dx} \left[(1-x^2) \frac{dP_{\ell,m}(x)}{dx} \right] + \frac{m^2}{(1-x^2)} P_{\ell,m}(x) = \ell(\ell+1) P_{\ell,m}(x). \quad (2.4.47)$$

Ainsi, tout champ scalaire $y(\mathbf{r})$ admet pour décomposition en coordonnées sphériques (r, θ, φ) :

$$y(\mathbf{r}) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} y_{\ell m}(r) Y_\ell^m(\theta, \varphi) \quad (2.4.48)$$

tandis qu'un champ vectoriel peut s'écrire (RIEUTORD, 2004) :

$$\mathbf{y}(\mathbf{r}) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} y_r^{\ell m}(r) \mathbf{R}_\ell^m(\theta, \varphi) + y_s^{\ell m}(r) \mathbf{S}_\ell^m(\theta, \varphi) + y_t^{\ell m}(r) \mathbf{T}_\ell^m(\theta, \varphi), \quad (2.4.49)$$

avec

$$\mathbf{R}_\ell^m(\theta, \varphi) \equiv Y_\ell^m(\theta, \varphi) \mathbf{e}_r \quad (2.4.50)$$

$$\mathbf{S}_\ell^m(\theta, \varphi) \equiv \nabla_2 Y_\ell^m(\theta, \varphi) \quad (2.4.51)$$

$$\mathbf{T}_\ell^m(\theta, \varphi) \equiv \nabla_2 \times \mathbf{R}_\ell^m(\theta, \varphi), \quad (2.4.52)$$

les composantes respectivement radiales, solénoïdales, et toroïdales sur la sphère. Il vient en effet immédiatement que $\nabla_2 \times \mathbf{S}_\ell^m = \mathbf{0}$ et $\nabla_2 \cdot \mathbf{T}_\ell^m = 0$ avec ∇_2 le gradient sur la sphère \mathcal{S}_2 . Ainsi, l'étude de toute grandeur faisant intervenir la position \mathbf{r} se ramène à celle de coefficients purement radiaux et identifiés par le doublet (ℓ, m) . Notons que la restriction à des champs scalaires réguliers au centre ($r = 0$) impose naturellement la condition aux limites suivante (RIEUTORD, 2004) :

$$y_{\ell m}(r) \underset{r \rightarrow 0}{=} \mathcal{O}(r^\ell). \quad (2.4.53)$$

Notons que l'application de cette propriété aux champs physiques $\mathbf{r} \cdot \mathbf{y} = \sum_{\ell m} r y_r^{\ell m} Y_\ell^m$, $\nabla \cdot \mathbf{y} = \sum_{\ell m} \left[\frac{1}{r^2} \frac{dr^2 y_r^{\ell m}}{dr} - \frac{\ell(\ell+1)}{r} y_s^{\ell m} \right] Y_\ell^m$ et $\mathbf{r} \cdot (\nabla \times \mathbf{y}) = \sum_{\ell m} \ell(\ell+1) y_t^{\ell m} Y_\ell^m$ implique que :

$$y_r^{\ell m}(r) \underset{r \rightarrow 0}{=} \mathcal{O}(r^{\ell-1}) \quad \text{et} \quad y_r^{00}(r) \underset{r \rightarrow 0}{=} \mathcal{O}(r) \quad (2.4.54)$$

$$y_s^{\ell m}(r) \underset{r \rightarrow 0}{=} \mathcal{O}(r^{\ell-1}) \quad (2.4.55)$$

$$y_t^{\ell m}(r) \underset{r \rightarrow 0}{=} \mathcal{O}(r^\ell). \quad (2.4.56)$$

où l'apparition des facteurs $\ell(\ell+1)$ vient du fait que $\Delta_2 Y_\ell^m = -\ell(\ell+1) Y_\ell^m$ (qui résulte directement de la définition des polynômes de Legendre (2.4.47)).

2.4.7 Conditions aux limites en surface

À la différence de ces conditions centrales, les conditions aux limites en surface ($r = R$) s'obtiennent par des arguments physiques. Par exemple, la pression $p + p'$ et le champ de déplacement ξ doivent être continus sur l'interface avec l'extérieur formée par le bord de l'étoile. On peut alors se demander comment une telle condition se traduit dans le cas d'une grandeur quelconque y . La surface étant déformée par le champ de déplacement ξ , celle-ci ne peut être considérée comme une sphère de rayon R et la valeur de y sur l'interface ne peut être donnée par $y(\mathbf{r}) + y'(\mathbf{r})$ avec $|\mathbf{r}| = R$. Il est en effet nécessaire d'estimer $y + y'$ en $\mathbf{r} + \xi$, ce qui correspond à $(y + y') + \xi \cdot \nabla(y + y') \simeq y + \delta y$, alors estimé en un point de la surface à l'équilibre (i.e. vérifiant $|\mathbf{r}| = R$). La condition résultante peut donc être écrite $y + \delta y \underset{r=R}{=} y_{\text{ext}} + \delta y_{\text{ext}}$ avec y_{ext} la valeur de y à l'extérieur de l'étoile. On remarquera cependant que, puisque cette condition doit être aussi bien vérifiée dans la structure à l'équilibre, $y \underset{r=R}{=} y_{\text{ext}}$ et donc que la continuité de y sur l'interface sera satisfaite tant que :

$$\delta y \underset{r=R}{=} \delta y_{\text{ext}}. \quad (2.4.57)$$

La continuité du champ de déplacement fournit immédiatement $\delta \mathbf{r} \underset{r=R}{=} \delta \mathbf{r}_{\text{ext}}$, ou autrement dit $\xi \underset{r=R}{=} \xi_{\text{ext}}$, ce qui nous assure que la perturbation lagrangienne possède bien la même définition sur les deux faces de la surface. Il sera alors possible d'explicitier un peu plus la condition (2.4.57) en écrivant :

$$(y'(r = R) - y'_{\text{ext}}(r = R)) + \xi \cdot (\nabla y|_{r=R} - \nabla y_{\text{ext}}|_{r=R}) = 0. \quad (2.4.58)$$

Notons que lorsque la pression extérieure est nulle (condition de surface libre), celle-ci vérifie simplement $\delta p \underset{r=R}{=} 0$ (LEDOUX et WALRAVEN, 1958). Puisque l'on a ici supposé que $\delta p = c^2 \delta \rho$, cela revient à considérer que $\nabla \cdot \xi \underset{r=R}{=} 0$, ou de façon explicite sur les coefficients :

$$\left. \frac{d\xi_r^{\ell m}}{dr} \right|_R + \frac{2}{R} \xi_r^{\ell m}(R) - \frac{\ell(\ell+1)}{R} \xi_s^{\ell m}(R) = 0. \quad (2.4.59)$$

D'autre part, la perturbation du potentiel de gravitation à l'extérieur de l'étoile vérifie dans le vide l'équation de Laplace : $\Delta \Phi'_{\text{ext}} = 0$. La seule solution valide (qui tend vers 0 à l'infini) pour ces coefficients s'écrit donc $\Phi'_{\text{ext}, \ell m} \propto 1/r^{\ell+1}$. La continuité de ce potentiel à travers la surface impose donc en $r = R$:

$$\Phi'_{\ell m}(R) = \Phi'_{\text{ext}, \ell m}(R) = \frac{\text{cste}}{R^{\ell+1}}. \quad (2.4.60)$$

De plus, le gradient de Φ' doit lui-même être continu sur l'interface :

$$\begin{aligned} \frac{\partial \Phi'}{\partial r} + \xi \cdot \nabla \left(\frac{d\Phi}{dr} \right) \underset{r=R}{=} \frac{d\Phi'_{\text{ext}}}{dr} + \xi \cdot \nabla \left(\frac{d\Phi_{\text{ext}}}{dr} \right) \\ \Rightarrow \frac{d\Phi'_{\ell m}}{dr} + \xi_r^{\ell m} \frac{d^2 \Phi}{dr^2} \underset{r=R}{=} -\frac{\ell+1}{r} \Phi'_{\text{ext}, \ell m} + \xi_r^{\ell m} \frac{d^2 \Phi_{\text{ext}}}{dr^2} \end{aligned} \quad (2.4.61)$$

et les champs à l'équilibre sont solutions des équations de Laplace et Poisson :

$$\Delta \Phi_{\text{ext}} = \frac{d^2 \Phi_{\text{ext}}}{dr^2} + \frac{2}{r} \frac{d\Phi_{\text{ext}}}{dr} = 0 \quad (2.4.62)$$

$$\Delta \Phi = \frac{d^2 \Phi}{dr^2} + \frac{2}{r} \frac{d\Phi}{dr} = 4\pi G \rho. \quad (2.4.63)$$

Puisque à l'équilibre, l'interface n'est autre que la sphère de rayon R , la continuité du gradient impose $d\Phi_{\text{ext}}/dr \Big|_{r=R} = d\Phi/dr$ et donc :

$$\frac{d^2\Phi_{\text{ext}}}{dr^2} - \frac{d^2\Phi}{dr^2} \Big|_{r=R} = \Delta\Phi_{\text{ext}} - \Delta\Phi \Big|_{r=R} = -4\pi G\rho. \quad (2.4.64)$$

Finalement, l'injection de cette relation dans (2.4.61) et la substitution de $\Phi'_{\ell m}$ par $\Phi'_{\text{ext}, \ell m}$ (cf. équation (2.4.60)) aboutit à :

$$\frac{d\Phi'_{\ell m}}{dr} \Big|_R + \frac{\ell+1}{R}\Phi'_{\ell m}(R) = -4\pi G\rho(R)\xi_r^{\ell m}(R). \quad (2.4.65)$$

Il est courant de trouver cette condition sous la forme

$$\frac{d\Phi'_{\ell m}}{dr} \Big|_R + \frac{\ell+1}{R}\Phi'_{\ell m}(R) = 0. \quad (2.4.66)$$

en supposant donc que $\rho(R) = 0$ (UNNO et al., 1989). Munie des conditions aux limites (2.4.59) et (2.4.65), l'équation (2.4.45) définit le problème aux valeurs propres (pour des oscillations adiabatiques) dans son entièreté.

2.4.8 Géométrie particulière du champ de déplacement

Il est temps de faire une remarque importante afin de simplifier l'écriture du champ de déplacement ξ . En partant de la décomposition du champ sur les harmoniques sphériques, on écrira :

$$\xi(\mathbf{r}) = \xi_r(\mathbf{r}) + \xi_s(\mathbf{r}) + \xi_t(\mathbf{r}) \quad (2.4.67)$$

où les expressions des composantes radiale, solénoïdale et toroïdale du déplacement s'obtiennent par identification avec l'équation (2.4.49). Considérons maintenant le rotationnel de l'équation (2.4.45) :

$$\omega^2 \nabla \times \xi = \nabla \times (\mathcal{L}\xi) = -\nabla \times \left(\left[\nabla^2 c^2 + \left(\frac{\Gamma_1}{\gamma} - 1 \right) \mathbf{g} \right] \nabla \cdot \xi \right). \quad (2.4.68)$$

qui s'obtient par élimination de tous ses gradients. Puisque le champ $\nabla^2 c^2 + (\Gamma_1/\gamma - 1) \mathbf{g}$ est porté par le vecteur \mathbf{r} (la structure à l'équilibre satisfait une symétrie sphérique), on peut remarquer que

$$\mathbf{r} \cdot (\nabla \times \xi) = 0 \quad \text{si} \quad \omega^2 \neq 0, \quad (2.4.69)$$

ou, écrit sur la décomposition en harmoniques sphériques :

$$\sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} \ell(\ell+1) \xi_t^{\ell m}(r) Y_\ell^m(\theta, \varphi) = -\Delta_2 \mathcal{T}(r, \theta, \varphi) = 0 \quad \text{si} \quad \omega^2 \neq 0. \quad (2.4.70)$$

Le champ scalaire $\mathcal{T} = \sum_{\ell m} \xi_t^{\ell m} Y_\ell^m$ ainsi défini ne peut donc varier que radialement : $\mathcal{T}(r, \theta, \varphi) = \mathcal{T}(r)$. Cependant, il est également relié au vecteur ξ_t via la décomposition de Chandrasekhar-Kendall (CHANDRASEKHAR et KENDALL, 1957; AIZENMAN et SMEYERS, 1977) :

$$\xi_t(\mathbf{r}) = \nabla \times (\mathcal{T}(r) \mathbf{e}_r) = \mathbf{0}, \quad (2.4.71)$$

puisque $\mathcal{T}(r) \mathbf{e}_r = \nabla(\int \mathcal{T} dr)$. Autrement dit, le champ de déplacement n'a pas de composante

toroïdale⁴³ ! Ce résultat important n'est toutefois pas systématique. Son obtention repose notamment sur l'hypothèse d'une structure sphérique à l'équilibre et ainsi l'absence de champ magnétique ou de rotation dans cette structure.

Dans le cadre considéré ici, cette hypothèse reste tout à fait valide et l'on se limitera donc aux deux premiers termes de l'équation (2.4.67). La postérité ayant gardée la dénomination de composante horizontale pour désigner la partie solénoïdale du champ, on notera finalement :

$$\begin{aligned}\boldsymbol{\xi}(\mathbf{r}) &= \boldsymbol{\xi}_r(\mathbf{r}) + \boldsymbol{\xi}_h(\mathbf{r}) \\ &= \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} \xi_r^{\ell m}(r) \mathbf{R}_\ell^m(\theta, \varphi) + \xi_h^{\ell m}(r) \mathbf{S}_\ell^m(\theta, \varphi)\end{aligned}\quad (2.4.72)$$

ou encore en base sphérique :

$$\boldsymbol{\xi}(\mathbf{r}) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} \xi_r^{\ell m} Y_\ell^m \mathbf{e}_r + \xi_h^{\ell m} \frac{\partial Y_\ell^m}{\partial \theta} \mathbf{e}_\theta + \frac{\xi_h^{\ell m}}{\sin \theta} \frac{\partial Y_\ell^m}{\partial \varphi} \mathbf{e}_\varphi. \quad (2.4.73)$$

2.4.9 Séparabilité des solutions

L'intérêt de la décomposition du champ de déplacement (2.4.72) apparaît tout de suite à la vue des propriétés d'orthogonalité suivantes (RIEUTORD, 2004) :

$$\int \mathbf{R}_{\ell'}^{m'*} \cdot \mathbf{S}_\ell^m d\Omega = 0 \quad (2.4.74)$$

$$\int \mathbf{R}_{\ell'}^{m'*} \cdot \mathbf{R}_\ell^m d\Omega = \int Y_{\ell'}^{m'*} Y_\ell^m d\Omega = \delta_{\ell\ell'} \delta_{mm'} \quad (2.4.75)$$

$$\int \mathbf{S}_{\ell'}^{m'*} \cdot \mathbf{S}_\ell^m d\Omega = \ell(\ell+1) \delta_{\ell\ell'} \delta_{mm'} \quad (2.4.76)$$

avec δ_{ij} le symbole de Kronecker et Ω l'angle solide et où les intégrales portent sur l'intégralité de sphère de rayon r (cela sera le cas de toutes les intégrales d'angle solide qui suivront). En effet, il est désormais possible de séparer le problème (2.4.45) en deux parties distinctes, et ce pour chaque couple (ℓ, m) :

$$\mathcal{L}_r^{\ell m} \boldsymbol{\xi} = \omega^2 \xi_r^{\ell m} \quad (2.4.77)$$

$$\mathcal{L}_h^{\ell m} \boldsymbol{\xi} = \omega^2 \xi_h^{\ell m} \quad (2.4.78)$$

avec

$$\mathcal{L}_r^{\ell m} \boldsymbol{\xi} \equiv \int \mathbf{R}_\ell^{m*} \cdot \mathcal{L} \boldsymbol{\xi} d\Omega \quad (2.4.79)$$

$$\ell(\ell+1) \mathcal{L}_h^{\ell m} \boldsymbol{\xi} \equiv \int \mathbf{S}_\ell^{m*} \cdot \mathcal{L} \boldsymbol{\xi} d\Omega. \quad (2.4.80)$$

deux opérateurs associant à $\boldsymbol{\xi}$ une fonction radiale.

Si la forme (2.4.44) de l'opérateur \mathcal{L} est utile pour mettre en évidence le lien entre l'équation d'oscillation et un problème aux valeurs propres, sa forme intégral-différentielle le rend peu pratique à utiliser pour obtenir une décomposition sur les harmoniques sphériques. On lui préférera son expression en fonction de $\nabla \Phi'$ avec la condition

$$\Delta \Phi' = 4\pi G \rho' = -4\pi G \nabla \cdot (\rho \boldsymbol{\xi}). \quad (2.4.81)$$

43. Si sa fréquence d'oscillation est non nulle. En réalité, le déplacement peut posséder une composante toroïdale non oscillatoire qui s'apparente à un mouvement constant contenu dans une coquille de rayon r . Pour un mode dipolaire ($\ell = 1$) ce mouvement s'apparente exactement à celui engendré par une rotation en couches (*shellular rotation* [en]) de faible amplitude : $\boldsymbol{\xi}_t(\mathbf{r}) = \delta \boldsymbol{\Omega}(r) \times \mathbf{r}$ (cf. AIZENMAN et SMEYERS, 1977)

Les différents opérateurs intervenant dans \mathcal{L} se décomposent alors de la façon suivante :

$$\nabla (\nabla \cdot \xi) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} \frac{d\Lambda_{\ell m}}{dr} \mathbf{R}_{\ell}^m + \frac{\Lambda_{\ell m}}{r} \mathbf{S}_{\ell}^m \quad (2.4.82)$$

$$\left[\nabla^2 c^2 + \left(\frac{\Gamma_1}{\gamma} - 1 \right) \mathbf{g} \right] \nabla \cdot \xi = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} \left[\frac{dc^2}{dr} - \left(\frac{\Gamma_1}{\gamma} - 1 \right) g \right] \Lambda_{\ell m} \mathbf{R}_{\ell}^m \quad (2.4.83)$$

$$\nabla (\xi \cdot \mathbf{g}) = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} -\frac{d g \xi_r^{\ell m}}{dr} \mathbf{R}_{\ell}^m - \frac{g \xi_r^{\ell m}}{r} \mathbf{S}_{\ell}^m \quad (2.4.84)$$

$$\nabla \Phi' = \sum_{\ell=0}^{+\infty} \sum_{m=-\ell}^{+\ell} \frac{d\Phi'_{\ell m}}{dr} \mathbf{R}_{\ell}^m + \frac{\Phi'_{\ell m}}{r} \mathbf{S}_{\ell}^m \quad (2.4.85)$$

avec $g \equiv |\mathbf{g}|$ et

$$\Lambda_{\ell m} \equiv \int Y_{\ell}^{m*} (\nabla \cdot \xi) d\Omega = \frac{1}{r^2} \frac{dr^2 \xi_r^{\ell m}}{dr} - \frac{\ell(\ell+1)}{r} \xi_h^{\ell m}. \quad (2.4.86)$$

Les deux opérateurs \mathcal{L}_r et \mathcal{L}_h s'obtiennent directement :

$$\mathcal{L}_r^{\ell m} \xi = \frac{d}{dr} \left(\Phi'_{\ell m} - c^2 \Lambda_{\ell m} + g \xi_r^{\ell m} \right) + \left(\frac{\Gamma_1}{\gamma} - 1 \right) g \Lambda_{\ell m} \quad (2.4.87)$$

$$\mathcal{L}_h^{\ell m} \xi = \frac{1}{r} \left(\Phi'_{\ell m} - c^2 \Lambda_{\ell m} + g \xi_r^{\ell m} \right). \quad (2.4.88)$$

sous la condition

$$\Delta \Phi'_{\ell m} - \frac{\ell(\ell+1)}{r^2} \Phi'_{\ell m} = -4\pi G \left(\rho \Lambda_{\ell m} + \frac{d\rho}{dr} \xi_r^{\ell m} \right), \quad (2.4.89)$$

fournie par l'équation (2.4.81). De façon remarquable, on voit que les opérateurs $\mathcal{L}_r^{\ell m}$ et $\mathcal{L}_h^{\ell m}$ ne font pas intervenir ξ dans son intégralité mais seulement $\xi^{\ell m} \equiv \xi_r^{\ell m} \mathbf{R}_{\ell}^m + \xi_h^{\ell m} \mathbf{S}_{\ell}^m$. Autrement dit, les $\xi^{\ell m}$ peuvent être étudiés indépendamment et sont eux-mêmes solutions du problème aux valeurs propres :

$$\mathcal{L} \xi^{\ell m} = \omega_{\ell m}^2 \xi^{\ell m}. \quad (2.4.90)$$

avec \mathcal{L} donné par l'équation (2.4.44). On omettra par la suite de renseigner le doublet (ℓ, m) pour alléger les notations (lorsque cela ne prête pas à confusion), on gardera seulement à l'esprit que chaque vecteur et valeur propres correspondent à un doublet spécifique. Remarquons que l'équation à laquelle on vient d'aboutir s'écrit alors de la même façon que le problème initial (2.4.44). Elle est néanmoins beaucoup plus simple. Si la géométrie du champ ξ sur les variables (θ, φ) pouvait être arbitrairement complexe dans le premier cas, elle désormais entièrement spécifiée par le choix du doublet (ℓ, m) auquel on s'intéresse. Ainsi, le problème aux valeurs propres (2.4.90) se réduit à l'étude de fonctions purement radiales, et ce par la résolution des équations (2.4.87), (2.4.88) (en remplaçant les opérateurs à gauches par leurs expressions respectives (2.4.77)-(2.4.78)) et (2.4.89). Les fréquences propres (et vecteurs propres) du nouveau problème aux valeurs propres (2.4.90) pourront être indexées à l'aide de l'entier n , que l'on appellera *ordre radial*. Ainsi, tout mode d'une étoile est *in fine* caractérisé par trois nombres quantiques : l'ordre radial n , le degré ℓ et l'ordre azimutal m , lesquelles spécifient intégralement sa géométrie. Notons que certains triplets vérifient des propriétés géométriques particulières. Ainsi, la définition (2.4.46) nous permet immédiatement de dire que $\partial_{\varphi} \xi = \mathbf{0}$ dans le cas particulier où $m = 0$. Ces modes propres possèdent donc une symétrie azimutale. Les modes vérifiant $\ell = 0$ possèdent, quant à eux, une symétrie sphérique et sont ainsi qualifiés de modes radiaux.

2.4.10 L'équation d'onde

On réfère par équation d'onde à la relation qui découle du système (2.4.87)-(2.4.89), liant une unique variable à la fréquence propre ω . Au regard dudit système (d'ordre 4 puisque l'équation (2.4.89) contient une dérivée seconde), l'obtention d'une telle équation semble complexe. On propose néanmoins l'obtention d'une simplification en annexe B.3, en réduisant les équations (2.4.87) et (2.4.88) sous la forme suivante :

$$\frac{d^2 \Xi}{d\tau^2} + (\omega^2 - V)\Xi = r \frac{d^2 \Psi}{d\tau^2} + Wr\Psi \quad (2.4.91)$$

avec comme variable la profondeur acoustique⁴⁴ $\tau \equiv \int_r^R dr/c$, Ξ et Ψ deux fonctions respectivement proportionnelles à la divergence du champ de déplacement Λ et la perturbation du potentiel gravitationnel Φ' et où V et W désignent deux potentiels dont les expressions peuvent être déduites de l'annexe B.3. Mais cette relation reste difficile à appréhender (sans parler d'en extraire des approximations pertinentes).

La variable encombrante (car ne s'exprimant pas trivialement en fonction des quantités ξ_r et ξ_{hr} , cf. relation B.1.13) est ici la perturbation du potentiel Φ' . Une approximation courante consiste tout simplement à négliger cette grandeur. Au regard de l'expression (2.4.34), cette hypothèse se justifie particulièrement bien à haute fréquence, Φ' s'exprimant comme l'intégrale sur l'espace d'une quantité oscillante de faible longueur d'onde. On peut donc s'attendre à une quasi annulation de cette perturbation à grande échelle. Cette approximation est connue sous le nom d'*approximation de Cowling* (COWLING, 1941) et est largement utilisée pour étudier les modes non radiaux ($\ell \neq 0$). Sous cette limite, il est alors courant de trouver l'équation d'onde sous la forme idéale qui suit :

$$\frac{d^2 \xi_r}{d\tau^2} + \omega^2 \left(1 - \frac{S_\ell^2}{\omega^2}\right) \left(1 - \frac{\mathcal{N}^2}{\omega^2}\right) \xi_r \simeq 0, \quad (2.4.92)$$

où les fréquences de Lamb S_ℓ et de Brunt-Väisälä \mathcal{N} :

$$S_\ell^2 \equiv \frac{\ell(\ell+1)c^2}{r^2} \quad (2.4.93)$$

$$\mathcal{N}^2 \equiv \left(\frac{\Gamma_1}{\gamma} - 1\right) \frac{g^2}{c^2} \quad (2.4.94)$$

définissent les régions de propagation ou d'évanescence de l'onde. En effet, l'inspection de l'équation (2.4.92) permet de voir que la variable ξ_r adopte un comportement oscillatoire lorsque $\omega^2 - S_\ell^2$ et $\omega^2 - \mathcal{N}^2$ sont de même signe, là où elle est exponentiellement atténuée dans le cas inverse. Cette relation d'onde définit ainsi les cavité p ($\omega^2 > S_\ell^2$ et $\omega^2 > \mathcal{N}^2$) des modes de pression et cavité g ($\omega^2 < S_\ell^2$ et $\omega^2 < \mathcal{N}^2$) des modes de gravité représentées en FIGURE 2.7.

Cependant, au-delà de l'intérêt certain qu'apporte son aspect intuitif, il est difficile de concevoir dans quel cas l'équation (2.4.92) constitue une approximation valable de l'équation d'onde (2.4.91) (munie de la contrainte (2.4.89)). Sous l'approximation de Cowling, cette dernière prend plutôt la forme d'une équation de Schrödinger stationnaire (cf. annexe B.3) :

$$\frac{d^2 \Xi}{d\tau^2} + (\omega^2 - V)\Xi = 0. \quad (2.4.95)$$

Dans cette équation, le potentiel acoustique V rassemble les contraintes à l'apparition de modes harmoniques idéaux (cas où $V = 0$). Pour pouvoir se ramener à l'expression (2.4.92), il serait alors nécessaire que les termes qui apparaissent dans le potentiel V (cf. annexe B.3) et homogènes à des fréquences au carré ($(g/c)^2$, $(c/r)^2$, g/r , dg/dr , ...) puissent être négligés devant

44. Il s'agit simplement du temps que met le son à atteindre la position r depuis la surface de l'étoile.

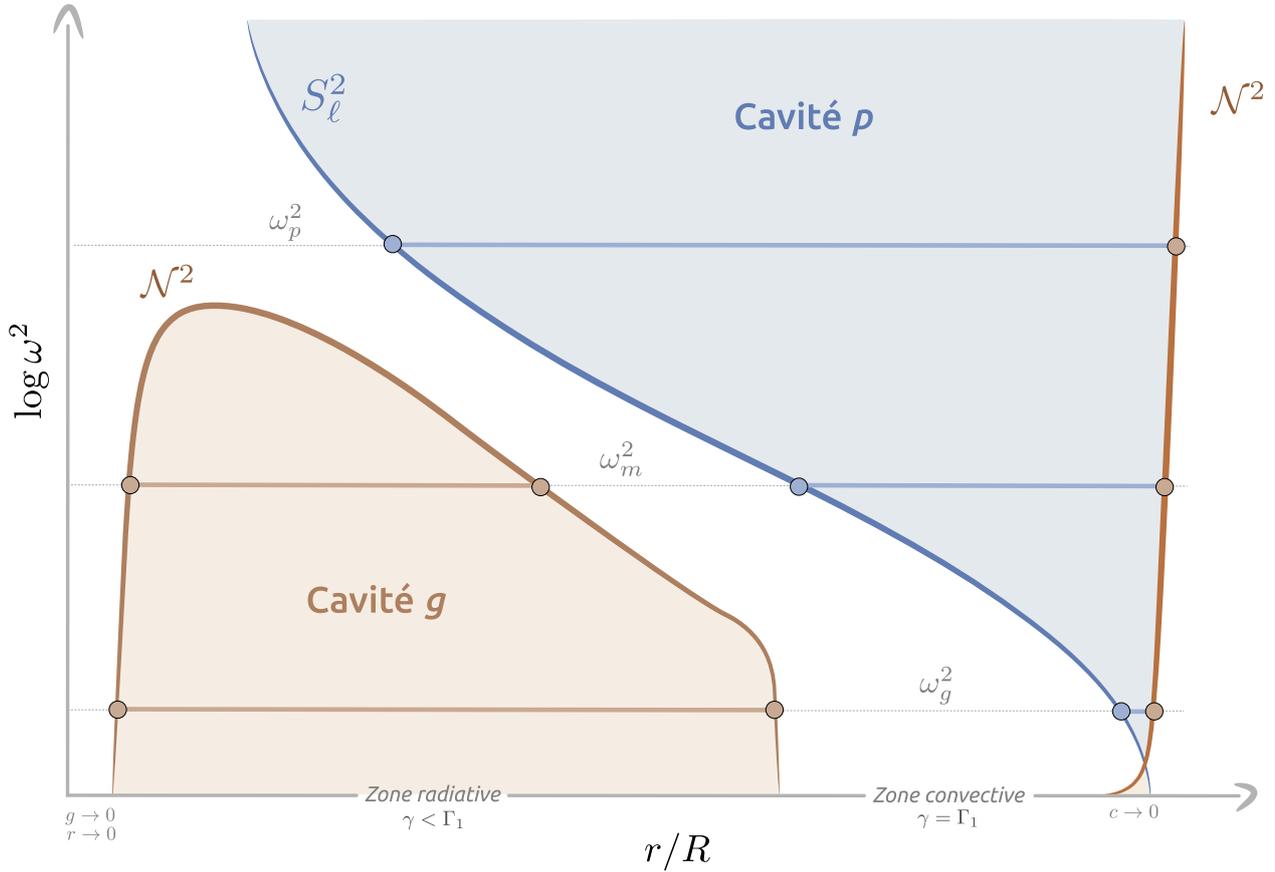


FIGURE 2.7 – Inspiré de UNNO et al., 1989. Représentation schématique des cavités p ($\omega^2 > S_\ell^2$ et $\omega^2 > N^2$, en bleu) et g ($\omega^2 < S_\ell^2$ et $\omega^2 < N^2$, en marron) engendrées par l'équation d'onde (2.4.92) pour plusieurs modes propres. Les fréquences propres ω_p , ω_g et ω_m qui y figurent sont respectivement associées à un mode de pression, de gravité et un mode mixte. Les régions dans lesquelles les modes adoptent un comportement oscillatoire sont représentées par un trait horizontal plus épais (bleu pour la cavité p, marron pour la cavité g). Les informations nécessaires à la compréhension des comportements des fréquences de Lamb S_ℓ (courbe bleue) et de Brunt-Väisälä \mathcal{N} (courbe marron) sont indiquées sur l'axe du bas (à comparer avec les équations (2.4.93)-(2.4.94)).

les fréquences de Lamb et de Brunt-Väisälä⁴⁵ (qui y apparaissent également). Or, ces fréquences sont précisément de l'ordre de $(c/r)^2$ et $(g/c)^2$ par définition (voir (2.4.93) & (2.4.94)). On se concentrera donc sur l'étude de l'équation de Schrödinger (2.4.95) dans la partie qui suit.

L'expression du potentiel s'exprime explicitement dans la limite des hautes fréquences ($\mathcal{L} \rightarrow 0$) de la façon suivante (cf. annexe B.3) :

$$\begin{aligned}
 V(\tau) \underset{\mathcal{L} \rightarrow 0}{\sim} & \left(\frac{\Gamma_1}{\gamma} - 1 \right) \frac{2g}{r} + \frac{dg}{dr} - \frac{2c}{r} \frac{dc}{dr} - \frac{d}{d\tau} \left(\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} \right) \\
 & + \left(\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} \right) \left(\frac{dg}{dr} - \frac{2g}{r} - \omega^2 \right)^{-1} \frac{d}{d\tau} \left(\frac{dg}{dr} - \frac{2g}{r} \right) \\
 & + \frac{1}{4} \left[\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} - \frac{2c}{r} - \left(\frac{dg}{dr} - \frac{2g}{r} - \omega^2 \right)^{-1} \frac{d}{d\tau} \left(\frac{dg}{dr} - \frac{2g}{r} \right) \right]^2 \\
 & + \frac{1}{2} \frac{d}{d\tau} \left[\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} - \frac{2c}{r} - \left(\frac{dg}{dr} - \frac{2g}{r} - \omega^2 \right)^{-1} \frac{d}{d\tau} \left(\frac{dg}{dr} - \frac{2g}{r} \right) \right].
 \end{aligned} \tag{2.4.96}$$

45. Cela ne constitue de plus qu'un prérequis, il n'est pas certain que les termes restant se factorisent de la même façon qu'en équation (2.4.92).

où \mathcal{L} désigne le rapport de la fréquence de Lamb par la fréquence d'oscillation. Les nombreux termes qui apparaissent rendent l'évaluation d'un comportement global assez incertaine mais l'on peut tirer quelques enseignements d'une observation qualitative de cette expression. Tout d'abord, remarquons que les termes qui divergent au centre de l'étoile ne se compensent pas. En cet endroit, le comportement du potentiel sera vraisemblablement dominé par :

$$V \underset{\substack{\mathcal{L} \rightarrow 0 \\ \tau \rightarrow \mathcal{T}}}{\sim} \frac{1}{4} \left[-\frac{2c}{r} \right]^2 + \frac{1}{2} \frac{d}{d\tau} \left[-\frac{2c}{r} \right] \quad (2.4.97)$$

$$\underset{\substack{\mathcal{L} \rightarrow 0 \\ \tau \rightarrow \mathcal{T}}}{\sim} \frac{2c^2}{r^2} = \mathcal{O}(r^{-2})$$

avec $\mathcal{T} \equiv \int_0^R dr/c$, le rayon acoustique total de l'étoile. Ainsi, on peut s'attendre à un point tournant (point auquel $V = \omega^2$) central pour toute fréquence ω finie, y compris pour les modes radiaux⁴⁶ (ce qui n'était pas le cas de l'équation (2.4.92) puisque la fréquence de Lamb serait alors nulle). Les autres termes sont divergeant en surface (g^2/c^2 par exemple) ou résultent du produit d'un terme divergeant et d'un terme s'annulant ($g/r = (g/c)(c/r)$, par exemple), ce qui leur confère une limite finie en cet endroit. Donner un équivalent précis du potentiel en surface est en revanche bien plus ardu, et nécessiterait en réalité de développer un à un les termes apparaissant dans l'équation (2.4.97). On peut néanmoins affirmer par une analyse numérique que ce potentiel diverge en surface et définit donc bien un point tournant superficiel pour toute fréquence propre. L'analyse des points tournants internes est moins triviale même si l'on peut supposer que les termes en $d\gamma/d\tau$ seront particulièrement sensibles à la région de transition entre les zones radiative et convective (voir chapitre 3) et sont ainsi susceptible d'annuler $\omega^2 - V$ à cet endroit. Notons enfin la présence de termes en d^2c/dr^2 , particulièrement sensibles à ce qui ce produit dans la région d'ionisation au travers de leur composante en $d^2\Gamma_1/dr^2$.

2.4.11 Analyse JWKB

Une approximation de la solution à l'équation (2.4.95) peut être obtenue dans l'approximation dite JWKB (JEFFREYS, 1925; WENTZEL, 1926; KRAMERS, 1926; BRILLOUIN, 1926; LANGER, 1937). Il s'agit de proposer des approximations à l'équation⁴⁷ :

$$\varepsilon^2 y'' + Q(x)y = 0 \quad (2.4.98)$$

dans la limite où $\varepsilon \rightarrow 0$ et ce dans les trois régions d'intérêt que sont : (I) $Q > 0$, (II) $Q \rightarrow 0$ et (III) $Q < 0$. En comparant les équations (2.4.95) et (2.4.98), on voit que notre équation d'onde se met sous cette forme standard en effectuant les transformations suivantes :

$$y \rightarrow \Xi, \quad (2.4.99)$$

$$x \rightarrow \tau/\mathcal{T} \equiv t, \quad (2.4.100)$$

$$Q \rightarrow 1 - \frac{V}{\omega^2}, \quad (2.4.101)$$

$$\varepsilon \rightarrow (\omega\mathcal{T})^{-1}. \quad (2.4.102)$$

Ainsi, la limite $\varepsilon \rightarrow 0$ correspond au régime où $\omega \rightarrow \infty$ ⁴⁸, et les trois régions d'intérêt correspondent respectivement à (I) $V < \omega^2$ (région de propagation), (II) $V \sim \omega^2$ (point tournant)

46. On rappelle que \mathcal{L} est strictement nul pour les modes radiaux, ce qui fait de l'équation (2.4.97) une expression exacte et non plus asymptotique dans ce cas.

47. À noter que l'on a adopté ici la convention $Q \rightarrow -Q$ par rapport à l'ouvrage de référence de BENDER et ORSZAG, 1978 pour éviter de faire apparaître un signe "−" dans la seule région qui va nous intéresser.

48. Pour donner un ordre d'idée, $\varepsilon \sim 0.015$ pour les modes les plus visibles du soleil ($\nu \sim 3000$ mHz), rendant relativement bien vérifiés les développements qui suivent.

et (III) $V > \omega^2$ (région évanescence). On ne s'intéressera par la suite qu'à l'approximation dans la région de propagation de l'onde (I), laquelle constitue la majeure partie de l'étoile. On détaillera pas la façon dont l'approximation JWKB est obtenue (celle-ci est détaillée de façon exhaustive dans **BENDER et ORSZAG, 1978**) et se contentera ici de discuter son expression au premier ordre :

$$y \propto Q^{-1/4} \sin \left(\frac{1}{\varepsilon} \int_x^B \sqrt{Q(x')} dx' + \frac{\pi}{4} \right) \quad (2.4.103)$$

où la constante de proportionnalité se déduit des conditions aux limites sous la contrainte :

$$\frac{1}{\varepsilon} \int_A^B \sqrt{Q(x)} dx = \left(n - \frac{1}{2} \right) \pi + \mathcal{O}(\varepsilon) \quad (2.4.104)$$

avec n un entier strictement positif et où A et B les points tournants inférieur et supérieur de la cavité. On remplaçant chacun des termes par leur expression dans le formalisme sismique :

$$\Xi \propto \left[1 - \frac{V}{\omega^2} \right]^{-1/4} \sin \left(\omega \int_{\tau_A}^{\tau} \sqrt{1 - \frac{V(\tau')}{\omega^2}} d\tau' + \frac{3\pi}{4} \right) \quad (2.4.105)$$

où l'on a substitué les bornes de l'intégrale à l'aide de la contrainte

$$\omega \int_{\tau_A}^{\tau_B} \sqrt{1 - \frac{V(\tau')}{\omega^2}} d\tau' = \left(n - \frac{1}{2} \right) \pi + \mathcal{O}((\omega\mathcal{T})^{-1}). \quad (2.4.106)$$

En développant la racine en série⁴⁹ :

$$\begin{aligned} \sqrt{1 - \frac{V}{\omega^2}} &= 1 + \sum_{k=1}^{\infty} \frac{\frac{1}{2} \left(\frac{1}{2} - 1 \right) \cdots \left(\frac{1}{2} - (k-1) \right)}{k!} \left(-\frac{V}{\omega^2} \right)^k \\ &= 1 + \sum_{k=1}^{\infty} \frac{\left[k - \frac{3}{2} \right] \left[\left(k - \frac{3}{2} \right) - 1 \right] \cdots \left[\left(k - \frac{3}{2} \right) - (k-1) \right]}{k!} \left(\frac{V}{\omega^2} \right)^k \\ &= 1 + \sum_{k=1}^{\infty} \frac{(k-3/2)!}{(-3/2)! k!} \left(\frac{V}{\omega^2} \right)^k, \end{aligned} \quad (2.4.107)$$

et de la même façon :

$$\left[1 - \frac{V}{\omega^2} \right]^{-1/4} = 1 + \sum_{k=1}^{\infty} \frac{(k-5/4)!}{(-5/4)! k!} \left(\frac{V}{\omega^2} \right)^k, \quad (2.4.108)$$

on peut mettre l'approximation (2.4.105) sous la forme :

$$\Xi \propto \mathcal{A}_{\infty} \sin(\omega\tau + \psi_{\infty}) \quad (2.4.109)$$

avec l'introduction d'une amplitude et d'une phase :

$$\mathcal{A}_{\infty}(\omega, \tau) \equiv 1 + \sum_{k=1}^{\infty} \binom{k}{-5/4} \left(\frac{V}{\omega^2} \right)^k \quad (2.4.110)$$

$$\psi_{\infty}(\omega, \tau) \equiv -\omega\tau_A + \frac{3\pi}{4} + \sum_{k=1}^{\infty} \binom{k}{-3/2} \frac{1}{\omega^{2k}} \int_{\tau_A}^{\tau} V^k d\tau. \quad (2.4.111)$$

49. On adopte ici la convention selon laquelle $\alpha!$ correspond à l'application de la fonction Gamma d'Euler sur $\alpha + 1$, qu'il soit entier ou non. Cela évite l'introduction d'une notation pouvant être confondue avec le premier exposant adiabatique Γ_1 .

Compte tenu de l'expression de V , la formule (2.4.109) reste encore une fois très complexe et l'on peut se demander s'il est possible d'en donner une approximation dans les régions où la propagation est peu contrainte ($V \ll \omega^2$). Dans un tel cas, il suffit de tronquer les développements à l'ordre souhaité. Par exemple, à l'ordre 0, la dépendance de Ξ en ω est simplement donnée par :

$$\Xi \propto \sin\left(\omega(\tau - \tau_A) + \frac{3\pi}{4}\right) \quad (2.4.112)$$

et l'on retrouve ainsi la solution harmonique.

2.4.12 Relation asymptotique des fréquences propres

On peut remarquer que la contrainte (2.4.106) donnée plus haut nous permet également de dériver une approximation des fréquences propres dans leur régime asymptotique. En effet, en supposant alors que $V/\omega^2 \ll 1$, on obtient l'approximation suivante :

$$\omega \int_{\tau_A}^{\tau_B} \sqrt{1 - \frac{V}{\omega^2}} d\tau = \omega \int_{\tau_A}^{\tau_B} \left[1 - \frac{V}{2\omega^2}\right] d\tau + \mathcal{O}\left(\frac{1}{\omega\mathcal{T}}\right) = \omega(\tau_B - \tau_A) + \mathcal{O}\left(\frac{1}{\omega\mathcal{T}}\right). \quad (2.4.113)$$

Écrite sur les fréquences $\nu \equiv \omega/2\pi$, l'équation (2.4.106) devient au premier ordre :

$$\nu \simeq \left(\frac{\mathcal{T}}{\tau_B - \tau_A}\right) \left(n - \frac{1}{2}\right) \Delta\nu \quad (2.4.114)$$

avec $\Delta\nu = (2\mathcal{T})^{-1}$, la grande séparation. Dans cette équation, la quantité $\mathcal{T}/(\tau_B - \tau_A)$ est une quantité proche de l'unité et qui s'en rapproche d'autant plus au fur et à mesure que $\nu \rightarrow \infty$ et que les points tournants se rapprochent de leurs limites $\tau_A \rightarrow 0$ et $\tau_B \rightarrow \mathcal{T}$. On essaiera ici d'une donner une approximation dans un cadre très schématique et typique des modes de pression : le point tournant τ_B est provoqué par la fréquence de Lamb S_ℓ et le point tournant superficiel est causé par la fréquence de coupure ω_c . Au centre de l'étoile, la variable τ peut être approximée par le développement de Taylor :

$$\tau \underset{r \rightarrow 0}{\sim} \mathcal{T} + r \frac{d\tau}{dr} = \mathcal{T} - \frac{r}{c}. \quad (2.4.115)$$

On pourra ainsi approcher la fréquence de Lamb par :

$$S_\ell^2 = \frac{\ell(\ell+1)c^2}{r^2} \underset{\tau \rightarrow \mathcal{T}}{\sim} \frac{\ell(\ell+1)}{(\mathcal{T} - \tau)^2} \quad (2.4.116)$$

et la solution de l'équation⁵⁰ $S_\ell^2 = \omega^2$ par :

$$\tau_B \simeq \mathcal{T} - \frac{\ell + 1/2}{\omega} \quad (2.4.117)$$

en procédant à l'approximation classique $\sqrt{\ell(\ell+1)} \simeq \ell + 1/2$. En réutilisant le développement de [GOUGH, 1993](#), une approximation pertinente pour la fréquence de coupure en surface se trouve être donnée par :

$$\omega_c \underset{\tau \rightarrow 0}{\sim} \frac{N_{\text{eff}} + 1}{\tau}. \quad (2.4.118)$$

où N_{eff} désigne un indice polytropique effectif, dans le sens où le polytrope d'exposant $\gamma_{\text{eff}} = 1 + 1/N_{\text{eff}}$ approxime raisonnablement bien la stratification en pression et densité de cette région.

50. À noter que l'on a trouvé un équivalent du potentiel V au centre en (2.4.97) : $V \sim 2c^2/r^2 \sim 2/(\mathcal{T} - \tau)^2$, même dans le cas d'une fréquence de Lamb nulle. Il serait donc plus cohérent de chercher le point tournant comme solution de $(\ell^2 + \ell + 2)/(\mathcal{T} - \tau)^2 = \omega^2$, ce qui mène à $\tau_B \simeq \mathcal{T} - \sqrt{\ell^2 + \ell + 2}/\omega$. L'équivalent obtenu pour des degrés ℓ suffisamment grand reste néanmoins le même.

La solution de l'équation $\omega_c^2 = \omega^2$ s'obtient alors immédiatement :

$$\tau_A \simeq \frac{N_{\text{eff}} + 1}{\omega}, \quad (2.4.119)$$

tout comme la taille de cavité :

$$\tau_B - \tau_A \simeq \mathcal{T} - \frac{\ell + N_{\text{eff}} + 3/2}{\omega}. \quad (2.4.120)$$

Toujours sous l'hypothèse des hautes fréquences $\varepsilon \rightarrow 0$, il vient au premier ordre :

$$\frac{\mathcal{T}}{\tau_B - \tau_A} \simeq 1 + \frac{\ell + N_{\text{eff}} + 3/2}{2\omega/\Delta\nu}. \quad (2.4.121)$$

On pourra alors expliciter complètement ce terme en procédant à l'approximation d'ordre 0 : $\omega/\Delta\nu \simeq (n - 1/2)$ dans la correction d'ordre 1 à l'aide de la relation (2.4.114) :

$$\frac{\mathcal{T}}{\tau_B - \tau_A} \simeq 1 + \frac{\ell + N_{\text{eff}} + 3/2}{2} (n - 1/2)^{-1}. \quad (2.4.122)$$

En remplaçant $\mathcal{T}/(\tau_B - \tau_A)$ par cette expression dans l'équation (2.4.114), la relation asymptotique des fréquences devient alors au premier ordre :

$$\nu \simeq \left(n + \frac{\ell}{2} + \frac{N_{\text{eff}}}{2} + \frac{1}{4} \right) \Delta\nu, \quad (2.4.123)$$

laquelle correspond exactement à celle donnée en (1.2.1). Si cette expression décrit assez fidèlement la répartition des fréquences pour les modes de pression et permet d'expliquer les régularités observées dans les spectres en puissance des pulsateurs de type solaire, celle-ci reste une approximation de premier ordre. On dérivera des formules bien plus exactes des fréquences propres dans la section qui suit, lesquelles nous permettront d'appréhender les effets plus fins qui nous intéressent.

Sur un autre sujet, on notera que l'approximation de la profondeur acoustique du point tournant superficiel (2.4.119) permet également de donner un développement explicite de la phase ψ_∞ du mode (cf. équation 2.4.111) à l'ordre 1 :

$$\psi_\infty \simeq -\omega\tau_A + \frac{3\pi}{4} \simeq \frac{3\pi}{4} - N_{\text{eff}} - 1. \quad (2.4.124)$$

On voit alors que la phase de l'oscillation est indépendante de la fréquence propre ω à cet ordre.

2.5 La théorie des perturbations

Contenu de la section

2.5.1	Symétrie de l'opérateur \mathcal{L}	82
2.5.2	Forme intégrale des fréquences propres	84
2.5.3	Principe variationnel	84
2.5.4	Perturbation des fréquences	87
2.5.5	Inversion sismique	88
2.5.6	Le glitch sismique	89

On s'est penché jusqu'ici sur le système d'équations dont les fréquences d'oscillations ω sont les valeurs propres, ainsi que sur des approximations asymptotiques à ces équations. On va voir dans cette section que ces fréquences peuvent également s'exprimer comme des intégrales portant sur la totalité de la structure de l'étoile. Ce formalisme est la base d'un ensemble de méthodes visant à mettre en lumière comment une différence de fréquences entre deux étoiles (ou entre un modèle et une étoile!) est liée à une différence de structure interne. Ces méthodes fournissent donc des outils puissants pour sonder une étoile et apporter des informations qui semblent, à priori, inatteignables par des seules contraintes de surface.

2.5.1 Symétrie de l'opérateur \mathcal{L}

De nombreuses simplifications au problème résultent du fait que l'opérateur est auto-adjoint, à savoir, pour deux champs de déplacement ξ et $\tilde{\xi}$:

$$\langle \tilde{\xi}, \mathcal{L}\xi \rangle = \langle \mathcal{L}\tilde{\xi}, \xi \rangle \quad (2.5.1)$$

vis-à-vis du produit scalaire

$$\langle \mathbf{u}, \mathbf{v} \rangle \equiv \frac{1}{M} \int_V \rho \mathbf{u}^* \cdot \mathbf{v} dV \quad (2.5.2)$$

où * dénote l'opération de conjugaison. La manière la plus directe de démontrer cette propriété consiste à partir de la forme de \mathcal{L} mettant en facteur $1/\rho$:

$$\mathcal{L}\xi = \frac{1}{\rho} \left[-\nabla (\Gamma_1 p \nabla \cdot \xi + \xi \cdot \nabla p) + (\nabla \cdot \xi) \nabla p + \frac{1}{\gamma p} (\xi \cdot \nabla p) \nabla p + \rho \nabla \Phi' \right], \quad (2.5.3)$$

ce qui donne :

$$\begin{aligned} \langle \tilde{\xi}, \mathcal{L}\xi \rangle &= \frac{1}{M} \int_V -\tilde{\xi}^* \cdot \nabla (\Gamma_1 p \nabla \cdot \xi + \xi \cdot \nabla p) + (\nabla \cdot \xi) (\tilde{\xi}^* \cdot \nabla p) \\ &\quad + \frac{1}{\gamma p} (\xi \cdot \nabla p) (\tilde{\xi}^* \cdot \nabla p) + \rho \tilde{\xi}^* \cdot \nabla \Phi' dV. \end{aligned} \quad (2.5.4)$$

La symétrie des premiers termes se montre facilement par intégration par partie :

$$\begin{aligned} \int_V -\tilde{\xi}^* \cdot \nabla (\Gamma_1 p \nabla \cdot \xi + \xi \cdot \nabla p) dV &= \oint_{\partial V} (-\Gamma_1 p \nabla \cdot \xi - \xi \cdot \nabla p) \tilde{\xi}^* \cdot d\mathbf{S} \\ &\quad + \int_V \Gamma_1 p (\nabla \cdot \xi) (\nabla \cdot \tilde{\xi}^*) + (\nabla \cdot \tilde{\xi}^*) (\xi \cdot \nabla p) dV. \end{aligned} \quad (2.5.5)$$

En effet, puisque $-\Gamma_1 p \nabla \cdot \xi = \delta p = 0$ sur la surface de l'étoile en vertu des conditions aux limites, le terme de surface se réduit à :

$$-\oint_{\partial V} (\xi \cdot \nabla p) \tilde{\xi}^* \cdot d\mathbf{S} = \oint_{\partial V} \frac{1}{|\nabla p|} (\xi \cdot \nabla p) (\tilde{\xi}^* \cdot \nabla p) dS, \quad (2.5.6)$$

la surface étant elle-même une isobare. À ce stade, il reste seulement à travailler sur la perturbation de l'énergie gravitationnelle, $\int_V \rho \tilde{\xi}^* \cdot \nabla \Phi' dV$. Une nouvelle intégration par partie permet d'écrire :

$$\int_V \rho \tilde{\xi}^* \cdot \nabla \Phi' dV = \oint_{\partial V} \Phi' \rho \tilde{\xi}^* \cdot d\mathbf{S} + \frac{1}{4\pi G} \int_V \Phi' \Delta \tilde{\Phi}'^* dV, \quad (2.5.7)$$

ce que la première identité de Green permet de réexprimer comme :

$$\int_V \rho \tilde{\xi}^* \cdot \nabla \Phi' dV = \oint_{\partial V} \Phi' \left(\rho \tilde{\xi}^* + \frac{\nabla \tilde{\Phi}'^*}{4\pi G} \right) \cdot d\mathbf{S} - \frac{1}{4\pi G} \int_V (\nabla \Phi') \cdot (\nabla \tilde{\Phi}'^*) dV. \quad (2.5.8)$$

Ici, on peut remarquer que le champ apparaissant dans le terme de surface est solénoïdal (cf. (2.4.81)); sa composante normale est donc conservée au passage de la surface de l'étoile. Ce terme peut donc être astucieusement réexprimé comme une intégrale de surface pour le volume externe \bar{V} :

$$-\oint_{\partial \bar{V}} \Phi' \left(\rho \tilde{\xi}^* + \frac{\nabla \tilde{\Phi}'^*}{4\pi G} \right) \cdot d\mathbf{S} = -\frac{1}{4\pi G} \int_{\bar{V}} (\nabla \Phi') \cdot (\nabla \tilde{\Phi}'^*) + \Phi' \Delta \tilde{\Phi}'^* dV \quad (2.5.9)$$

puisque la densité est nulle à l'extérieur de l'étoile. Le deuxième terme s'annule puisque $\Delta \tilde{\Phi}'^* = 0$ dans ce domaine et l'on écrira finalement en rassemblant les expressions (2.5.5), (2.5.6), (2.5.8) et (2.5.9) :

$$\begin{aligned} \langle \tilde{\xi}, \mathcal{L}\xi \rangle &= \frac{1}{M} \int_V \Gamma_1 p (\nabla \cdot \xi) (\nabla \cdot \tilde{\xi}^*) + (\nabla \cdot \tilde{\xi}^*) (\xi \cdot \nabla p) + (\nabla \cdot \xi) (\tilde{\xi}^* \cdot \nabla p) dV \\ &+ \frac{1}{M} \int_V \frac{1}{\gamma p} (\xi \cdot \nabla p) (\tilde{\xi}^* \cdot \nabla p) dV + \frac{1}{M} \oint_{\partial V} \frac{1}{|\nabla p|} (\xi \cdot \nabla p) (\tilde{\xi}^* \cdot \nabla p) dS \\ &- \frac{1}{4\pi GM} \int_{V \cup \bar{V}} (\nabla \Phi') \cdot (\nabla \tilde{\Phi}'^*) dV. \end{aligned} \quad (2.5.10)$$

Cette expression est parfaitement symétrique en ξ et $\tilde{\xi}$, ce qui démontre la propriété (2.5.1). Ce résultat est central pour ce qui est de l'étude des différences de structure entre deux étoiles comme on le verra dans le paragraphe 2.5.3. On se propose néanmoins de se pencher sur un premier corollaire. Supposons que ξ et $\tilde{\xi}$ soient maintenant deux vecteurs propres associés aux valeurs propres ω^2 et $\tilde{\omega}^2$. On a ainsi

$$\omega^2 \langle \tilde{\xi}, \xi \rangle = \langle \tilde{\xi}, \mathcal{L}\xi \rangle = \langle \mathcal{L}\tilde{\xi}, \xi \rangle = (\tilde{\omega}^2)^* \langle \tilde{\xi}, \xi \rangle \quad (2.5.11)$$

ou écrit autrement :

$$(\omega^2 - (\tilde{\omega}^2)^*) \langle \tilde{\xi}, \xi \rangle = 0. \quad (2.5.12)$$

Commençons par noter que dans le cas où $\tilde{\xi} = \xi$, ce résultat indique que $\omega^2 = (\omega^2)^*$ et donc que toute valeur propre est réelle. Dans le cas général, cette équation montre en plus l'orthogonalité des vecteurs propres associés à des valeurs propres distinctes. Bien que cela ne prouve pas complètement l'orthogonalité entre deux vecteurs propres quelconque associés à deux triplets (n, ℓ, m) distincts (il pourrait exister une dégénérescence), il est d'usage de la supposer (UNNO et al., 1989).

2.5.2 Forme intégrale des fréquences propres

En plus de démontrer la symétrie de l'opérateur \mathcal{L} , la formule (2.5.10) peut mener à une expression sous forme intégrale des fréquences propres d'oscillations. En effet, si $\tilde{\xi} = \xi$ et si ce champ de déplacement a pour fréquence propre ω^2 , le terme de gauche s'écrit $\omega^2 \langle \tilde{\xi}, \xi \rangle$ et les intégrales peuvent être explicitées en termes de fonctions radiales en remarquant que :

$$\int (\xi^* \cdot \xi) d\Omega = \xi_r^2 + \ell(\ell + 1)\xi_h^2, \quad (2.5.13)$$

$$\int (\nabla \cdot \xi)(\nabla \cdot \xi^*) d\Omega = \Lambda^2, \quad (2.5.14)$$

$$2 \int \text{Re} \{ (\nabla \cdot \xi^*)(\xi \cdot \nabla p) \} d\Omega = 2\xi_r \Lambda \frac{dp}{dr}, \quad (2.5.15)$$

$$\int (\xi \cdot \nabla p)(\xi^* \cdot \nabla p) d\Omega = \left(\xi_r \frac{dp}{dr} \right)^2. \quad (2.5.16)$$

Le seul terme ne s'exprimant pas facilement en fonction de ξ_r et ξ_h est une nouvelle fois la perturbation de l'énergie gravitationnelle, laquelle fait intervenir la perturbation du potentiel Φ' . Il est alors courant de procéder à l'approximation de Cowling (cf. paragraphe 2.4.10) et de simplement négliger ce terme. Il est toutefois possible d'en donner une expression analytique exacte, chose qui est faite dans l'annexe B.1. On obtient l'expression suivante pour le terme de Cowling :

$$\int_V \rho \tilde{\xi}^* \cdot \nabla \Phi' dV = -4\pi G \int_0^R \left(r^\ell \int_r^R s^{-\ell+1} \rho' ds \right)^2 dr, \quad (2.5.17)$$

où ρ' désigne la perturbation eulérienne de la densité, liée au déplacement ξ_r et à sa divergence par :

$$\rho' = -\rho \Lambda - \xi_r \frac{d\rho}{dr}. \quad (2.5.18)$$

Il est alors possible d'obtenir une expression explicite de la fréquence propre ω^2 sous forme intégrale à l'aide des projections (2.5.13)-(2.5.16) en plus de (2.5.17) :

$$\begin{aligned} \omega^2 \int_0^R \rho r^2 (\xi_r^2 + \ell(\ell + 1)\xi_h^2) dr &= \int_0^R \rho c^2 r^2 \Lambda^2 + 2r^2 \xi_r \Lambda \frac{dp}{dr} + \frac{r^2}{\rho} \frac{d\rho}{dr} \frac{dp}{dr} \xi_r^2 dr \\ &\quad - \int_0^R 4\pi G \left(r^\ell \int_r^R s^{-\ell+1} \rho' ds \right)^2 dr + GM\rho(R)\xi_r^2|_R \end{aligned} \quad (2.5.19)$$

où l'on a utilisé les propriétés $\frac{1}{\gamma p} \frac{dp}{dr} = \frac{1}{\rho} \frac{d\rho}{dr}$ et $\frac{dp}{dr} \Big|_R = -\frac{GM\rho(R)}{R^2}$. Si l'équation (2.5.19) a perdu la symétrie qui transparaissait de la relation (2.5.10), elle est néanmoins beaucoup plus adaptée à l'objectif que l'on s'est fixé, à savoir obtenir l'expression d'une différence des fréquences propres pour deux structures distinctes.

2.5.3 Principe variationnel

Imaginons désormais que l'on dispose de deux étoiles dont les structures à l'équilibre diffèrent et ce, de façon suffisamment faible pour que l'on puisse écrire cette différence comme une perturbation de la structure interne. En nommant ces étoiles A et B, on pourrait écrire par exemple $\rho^B = \rho^A + \delta\rho$ ou encore $p^B = p^A + \delta p$. Notons que la nature de ce que l'on appelle perturbation ici est très différente de ce à quoi on référerait au début de la section précédente. Il n'est pas question de perturbations dynamiques d'une grandeur physique mais d'une différence strictement statique entre deux structures à l'équilibre (accordement, le symbole " δ " désignera une différence entre structures à partir de là et non plus la perturbation lagrangienne définie en section 2.4). Ceci étant dit, la notation $\rho^B = \rho^A + \delta\rho$ reste ambiguë. En effet, convient d'explicitier

clairement quels “points” des deux étoiles on choisit pour faire cette différence. On pourrait par exemple choisir de parler de différences au même rayon r par exemple, ou encore à même masse cumulée $m(r)$. Si les profils de densité $\rho(r' < r)$ diffèrent entre nos deux étoiles, ces deux critères ont malheureusement toutes les chances de ne pas correspondre. Autrement dit, de manière générale, il est clair que $r^A = r^B \not\Rightarrow m^A(r^A) = m^B(r^B)$ ⁵¹. Pour éviter toute confusion par la suite, on choisira la notation introduite par CHRISTENSEN-DALSGAARD et THOMPSON, 1997 : $\rho^B = \rho^A + \delta_x \rho$ pour parler de perturbation à x fixé (x doit évidemment être monotone aussi bien dans A que dans B). La FIGURE 2.8 donne un exemple simple illustrant les écarts qui peuvent résulter d’un choix de variable différent, et montre ainsi l’ambiguïté autour de la notation “ δ ”. Enfin, une dernière précision doit être apportée pour les variables qui ne varient pas dans la même gamme de valeurs entre les deux étoiles, ce qui risque d’arriver si $R^A \neq R^B$ ou $M^A \neq M^B$. Alors les points $R^A < r < R^B$ posent problème pour faire la comparaison, puisqu’ils se situent en dehors de l’étoile A. Ainsi, il est préférable de toujours se ramener à des variables normalisées pour étudier les perturbations comme $x = r/R$ ou m/M . Ces dernières nous évitent cette préoccupation puisqu’elles varient toujours entre 0 et 1.⁵²

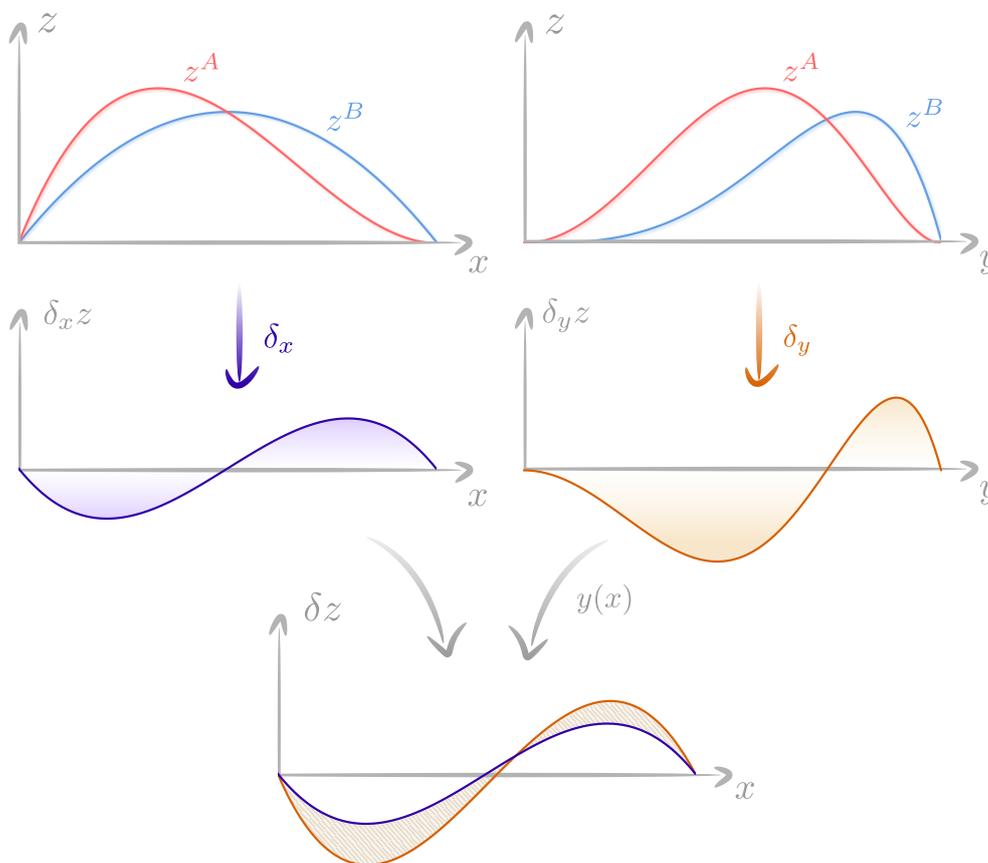


FIGURE 2.8 – Illustration des écarts entre les perturbations $\delta_x z$ et $\delta_y z$ pour les fonctions $z^A(x) = 12x(1-x)^2$, $z^B(x) = 6x(1-x)$, $y^A(x) = \sqrt{x}$ et $y^B(x) = \sqrt[3]{x}$. Cette exemple simple permet de donner analytiquement l’écart entre les différences de structure (aire hachurée dans le dernier graphique) $\delta_x z(x) - \delta_y z(x) = 6x(1 - \sqrt{x} - x + x^2) \neq 0$, ce qui illustre l’intérêt de lever l’ambiguïté autour de ce choix de variable. En particulier, on notera que la zone hachurée est du même ordre de grandeur que la perturbation elle-même ($\delta_x z(x) = 6x(1-x)(1-2x)$) et non pas négligeable devant cette dernière.

Puisque ces deux étoiles diffèrent par leurs structures internes respectives, elles présenteront généralement des fréquences d’oscillation et des champs de déplacement différents. On pourra

51. Il est tout aussi clair que $m^A = m^B \not\Rightarrow r^A(m^A) = r^B(m^B)$.

52. Formellement, la fonction $\delta_x \rho$ peut être définie comme $\delta_x \rho \equiv r \mapsto \rho^B((x^B)^{-1}(x^A(r))) - \rho^A(r)$. Dès lors, les contraintes sur x sus-citées apparaissent comme des conditions d’existence et d’unicité de la fonction $(x^B)^{-1}(x^A)$.

noter $\omega^B = \omega^A + \delta\omega$ (il n'est pas nécessaire ici de préciser la variable sur laquelle porte le "δ", puisque ω est une constante pour les étoiles A et B) et $\xi^B = \xi^A + \delta_x \xi$. La question est maintenant de trouver quel lien il y a entre les variations de structure interne et les perturbations des fréquences d'oscillation. Avant de déterminer explicitement ce lien à l'aide de l'équation (2.5.19), on va démontrer une propriété propre au problème considéré. Commençons par perturber l'équation (2.4.90) :

$$\delta\omega^2 \langle \xi, \xi \rangle + \omega^2 \delta \langle \xi, \xi \rangle = \delta \langle \xi, \mathcal{L}\xi \rangle. \quad (2.5.20)$$

Comme on l'a vu dans le paragraphe précédent, les produits scalaires prennent la forme d'intégrales pour les fonctions ξ et $\mathcal{L}\xi$. Supposons pour le moment qu'il existe une variable normalisée, x , telle que δ_x commute avec ces intégrales. Puisque l'on se limite ici à de faibles perturbations, on se contentera de conserver les termes d'ordre 1. Ainsi :

$$\delta \langle \xi, \xi \rangle = \langle \delta_x \xi, \xi \rangle + \langle \xi, \delta_x \xi \rangle \quad (2.5.21)$$

$$\delta \langle \xi, \mathcal{L}\xi \rangle = \langle \delta_x \xi, \mathcal{L}\xi \rangle + \langle \xi, \delta_x (\mathcal{L}\xi) \rangle. \quad (2.5.22)$$

Le dernier terme apparaissant ici se décompose en deux contributions distinctes au premier ordre :

$$\delta_x (\mathcal{L}\xi) = [\delta_x \mathcal{L}] \xi + \mathcal{L} \delta_x \xi. \quad (2.5.23)$$

La première porte sur les différences de structure interne qui modifient l'opérateur \mathcal{L} lui-même puisqu'il les fait intervenir. La seconde résulte de l'application de l'opérateur à la variation de la solution propre entre nos deux structures. La symétrie de \mathcal{L} nous permet alors d'écrire :

$$\delta \langle \xi, \mathcal{L}\xi \rangle = \langle \delta_x \xi, \mathcal{L}\xi \rangle + \langle \mathcal{L}\xi, \delta_x \xi \rangle + \langle \xi, [\delta_x \mathcal{L}] \xi \rangle \quad (2.5.24)$$

ce qui, réinjecté dans l'expression (2.5.20), mène à :

$$\delta\omega^2 \langle \xi, \xi \rangle = \langle \delta_x \xi, \mathcal{L}\xi - \omega^2 \xi \rangle + \langle \mathcal{L}\xi - (\omega^2)^* \xi, \delta_x \xi \rangle + \langle \xi, [\delta_x \mathcal{L}] \xi \rangle. \quad (2.5.25)$$

Les deux premiers termes de cette équation sont identiquement nulles puisque ξ est vecteur propre de \mathcal{L} et que la valeur propre ω^2 est réelle (cf. paragraphe 2.5.1). Finalement, la variation des fréquences d'oscillation s'écrit comme :

$$\delta\omega^2 = \frac{\langle \xi, [\delta_x \mathcal{L}] \xi \rangle}{\langle \xi, \xi \rangle} \quad (2.5.26)$$

et ne fait intervenir que les perturbations de structure interne et non pas $\delta_x \xi$. Cette propriété est connue sous le nom de principe variationnel (CHANDRASEKHAR, 1964) et constitue le lien recherché entre perturbation des fréquences $\delta\omega^2$ et perturbation de la structure à l'équilibre via $\delta_x \mathcal{L}$. Il reste maintenant à déterminer quelle perturbation peut commuter avec le produit scalaire. À la vue de la définition (2.5.2), on peut écrire en découplant les variables radiales et angulaire (ce qui est possible avec tous les opérateurs que l'on a considéré jusqu'ici) :

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{M} \int_0^R \langle \mathbf{u}, \mathbf{v} \rangle_2 4\pi r^2 dr = \int_0^1 \langle \mathbf{u}, \mathbf{v} \rangle_2 d(m/M) \quad (2.5.27)$$

avec

$$\langle \mathbf{u}, \mathbf{v} \rangle_2 \equiv \frac{1}{4\pi} \int (\mathbf{u}^* \cdot \mathbf{v}) d\Omega \quad (2.5.28)$$

le produit scalaire normalisé sur la sphère. Il est clair que la perturbation de structure qui commute avec cette intégrale sans faire apparaître de terme additionnel est $\delta_{m/M}$. Pour un choix de variable normalisée différent, le principe variationnel ne s'applique pas strictement et des termes supplémentaires peuvent apparaître. Cela se voit directement avec l'équation (2.5.27). Pour la

variable $x = r/R$ par exemple, on obtient :

$$\delta \langle \mathbf{u}, \mathbf{v} \rangle = \left[3 \frac{\delta R}{R} - \frac{\delta M}{M} \right] \langle \mathbf{u}, \mathbf{v} \rangle + \frac{R}{M} \int_0^1 \left[\delta_x (\langle \mathbf{u}, \mathbf{v} \rangle_2) + \frac{\delta_x \rho}{\rho} \langle \mathbf{u}, \mathbf{v} \rangle_2 \right] 4\pi r^2 dx. \quad (2.5.29)$$

À noter que les termes proportionnels à $\langle \mathbf{u}, \mathbf{v} \rangle$ ne sont pas embêtants pour la perturbation d'une égalité puisqu'ils se compensent à gauche et à droite. On voit malgré tout apparaître un nouveau terme en $\delta_x \rho / \rho$ à l'intérieur de l'intégrale qui n'apparaît pas pour une perturbation à m/M fixée :

$$\delta \langle \mathbf{u}, \mathbf{v} \rangle = \int_0^1 \delta_{m/M} (\langle \mathbf{u}, \mathbf{v} \rangle_2) d(m/M). \quad (2.5.30)$$

Les expressions (2.5.29) et (2.5.30) sont bien évidemment égales, les différences apparentes ne résultant que du choix de variable pour définir la perturbation. Elles illustrent néanmoins l'importance de ce choix pour mener les calculs à bien. En particulier, une application naïve du principe variationnel pourrait mener à :

$$\delta \langle \mathbf{u}, \mathbf{v} \rangle = \int_0^1 \delta_x (\langle \mathbf{u}, \mathbf{v} \rangle_2) d(m/M). \quad (2.5.31)$$

Au regard de ce que l'on a dit plus haut, **cette équation est cependant fausse.**

2.5.4 Perturbation des fréquences

Bien qu'il semble que la dérivation de la perturbation $\delta\omega^2$ soit plus simple en utilisant la différence $\delta_{m/M}$ (elle l'est), il est usuel de la mener à l'aide de $\delta_{r/R}$ (à laquelle on réservera la notation δ_x dans ce qui suit). Si l'on a vu que le principe variationnel ne s'applique pas dans ce cas, le découplage entre variables radiale et angulaires suffit à assurer qu'aucun terme supplémentaire en $\delta_x \xi$ n'apparaît dans le calcul. On en proposera la dérivation en annexe B.2 en partant de l'expression (2.5.19), la dérivation étant surtout technique et n'offrant que peu d'éclairage sur la physique en jeu. Les calculs effectués, la perturbation relative des fréquences se décompose sur trois termes :

$$\frac{\delta\omega}{\omega} = \int_0^R \mathcal{K}_{c^2}(r, \omega) \frac{\delta_x c^2}{c^2} + \mathcal{K}_\rho(r, \omega) \frac{\delta_x \rho}{\rho} dr - \mathcal{S}_R(\omega) \frac{\delta R}{R} \quad (2.5.32)$$

avec :

$$\mathcal{K}_{c^2} = \frac{\rho r^2 c^2 \Lambda^2}{2\omega^2 \mathcal{I}}, \quad (2.5.33)$$

$$\mathcal{K}_\rho = \frac{\rho r^2}{2\omega^2 \mathcal{I}} \left[c^2 \Lambda^2 - \omega^2 (\xi_r^2 + \ell(\ell+1)\xi_h^2) - 2g\xi_r \left(\Lambda - \frac{d\xi_r}{dr} \right) - 4\pi G \int_r^R 2\rho\xi_r \left(\Lambda - \frac{d\xi_r}{ds} \right) ds \right. \\ \left. + 8\pi G \left(\rho\xi_r^2 - r^{\ell-1}\xi_r \int_r^R s^{-\ell+1}\rho' ds + (\ell+1)r^{-\ell-2}(\xi_r + \ell\xi_h) \int_0^r s^{2\ell} \int_s^R t^{-\ell+1}\rho' dt ds \right) \right], \quad (2.5.34)$$

$$\mathcal{S}_R = \frac{1}{\omega^2 \mathcal{I}} \int_0^R \rho r^2 \Lambda^2 c^2 dr. \quad (2.5.35)$$

et où \mathcal{I} désigne l'inertie du mode :

$$\mathcal{I} \equiv \int_0^R \rho r^2 (\xi_r^2 + \ell(\ell+1)\xi_h^2) dr, \quad (2.5.36)$$

Dans cette équation, les fonctions \mathcal{K}_{c^2} et \mathcal{K}_ρ désignent les *kernels* respectivement associés à la vitesse du son et la densité dans l'étoile. Le terme de surface \mathcal{S}_R traduit quant à lui un écart de fréquence dû à une simple différence des paramètres globaux, en l'occurrence du rayon δR . Notons que cette manifestation est entièrement due à notre choix de variable de perturbation.

Puisque l'on perturbe à rayon normalisé fixé, l'impact d'un changement de rayon ne peut se manifester qu'en dehors de l'intégrale, par un terme additionnel. Une variation d'un autre paramètre comme la masse, δM , aura naturellement un impact sur les fréquences mais celui-ci peut s'exprimer de façon implicite, par l'intégration de perturbations de structure. Le terme \mathcal{S}_R s'interprète donc comme la manifestation d'une correction entre deux structures possédant un rayon R différent. À l'opposé, l'équation (2.5.32) nous indique que la moindre différence de vitesse du son ou de densité, quelque soit sa position dans l'étoile, peut se traduire par une différence dans les fréquences d'oscillation respectives aux deux structures. Cet effet n'est toutefois pas systématique, les kernels pouvant être localement nuls (ou très faibles) et les différences positives et négatives pouvant se compenser dans l'intégrale.

2.5.5 Inversion sismique

Dans l'équation (2.5.32), le terme de surface est souvent négligé (GOUGH et THOMPSON, 1991), cet effet étant généralement trop fin comparé aux nombreuses incertitudes entourant cette région. Cette relation nous fournit alors le lien désiré dans le paragraphe 2.5.3, liant les différences de structure entre deux étoiles à des différences dans leurs fréquences d'oscillation. Notons que cet écart ne fait pas intervenir toutes les quantités structurelles mais seulement deux, en l'occurrence $\delta_x c^2$ et $\delta_x \rho$. Cette propriété se vérifie presque systématiquement⁵³ (sous l'hypothèse d'une structure à symétrie sphérique, GOUGH et THOMPSON, 1991) même si le choix de ce couple peut être adapté selon les usages (BULDGEN, REESE et DUPRET, 2015; BULDGEN et al., 2017b). La question est alors de savoir comment exploiter cette relation en pratique.

Dans le cas d'une exploitation de données sismiques, on a tout d'abord à notre disposition les fréquences d'oscillation d'un certain nombre de modes de l'astre. Notons \mathcal{D} l'ensemble des triplets (n, ℓ, m) présents dans nos données. On peut dans un second temps générer un modèle de structure pour cette étoile et calculer ces fréquences propres d'oscillations en résolvant le système (2.4.90). Malgré tous nos efforts, il est vraisemblable que les fréquences ainsi générées diffèrent des fréquences observées, nous donnant accès aux $\delta\omega_i/\omega_i$ pour $i \in \mathcal{D}$. Par ailleurs, l'accès à la structure d'équilibre du modèle calculé (ainsi que la détermination des modes propres en son sein) nous donne une connaissance parfaite des $\mathcal{K}_\rho(r, \omega_i)$ et $\mathcal{K}_{c^2}(r, \omega_i)$ pour $i \in \mathcal{D}$ via les équations (2.5.33) et (2.5.34). L'équation (2.5.32) se trouve alors être la formulation d'un problème inverse ayant pour inconnues les fonctions $\delta_x \rho/\rho$ et $\delta_x c^2/c^2$.

Ces deux profils peuvent alors être approximés par une certaine combinaison de fonctions. En effet, on peut exprimer une combinaison de différences de fréquences à l'aide de l'équation (2.5.32) :

$$\sum_{i \in \mathcal{D}} \mathcal{C}_i(r_0) \frac{\delta\omega_i}{\omega_i} = \int_0^R \left[\sum_{i \in \mathcal{D}} \mathcal{C}_i(r_0) \mathcal{K}_{c^2}(r, \omega_i) \right] \frac{\delta_x c^2}{c^2} + \left[\sum_{i \in \mathcal{D}} \mathcal{C}_i(r_0) \mathcal{K}_\rho(r, \omega_i) \right] \frac{\delta_x \rho}{\rho} dr \quad (2.5.37)$$

avec les $\mathcal{C}_i(r_0)$ des coefficients quelconques pour $i \in \mathcal{D}$ et $0 < r_0 < R$. Considérons maintenant deux séries de coefficients distinctes, $\mathcal{C}_i^{c^2}(r_0)$ et $\mathcal{C}_i^\rho(r_0)$ chacune vérifiant (au moins de manière approximative) les deux conditions :

$$\sum_{i \in \mathcal{D}} \mathcal{C}_i^{c^2}(r_0) \mathcal{K}_{c^2}(r, \omega_i) \simeq \delta(r - r_0) \quad \text{et} \quad \sum_{i \in \mathcal{D}} \mathcal{C}_i^{c^2}(r_0) \mathcal{K}_\rho(r, \omega_i) \simeq 0, \quad (2.5.38)$$

$$\sum_{i \in \mathcal{D}} \mathcal{C}_i^\rho(r_0) \mathcal{K}_{c^2}(r, \omega_i) \simeq 0 \quad \text{et} \quad \sum_{i \in \mathcal{D}} \mathcal{C}_i^\rho(r_0) \mathcal{K}_\rho(r, \omega_i) \simeq \delta(r - r_0). \quad (2.5.39)$$

53. On pourra noter le cas particulier d'une expression de la forme intégrale (2.5.32) sur des différences d'abondances comme $\delta_x Y$ (différence en hélium) et $\delta_x Z$ (différence en métaux), qui elle peut faire intervenir des variables supplémentaires (et leurs kernels associés BULDGEN et al., 2017a).

Dans ce cas, la relation (2.5.37) nous permet immédiatement d'écrire

$$\sum_{i \in \mathcal{D}} \mathcal{C}_i^{c^2}(r_0) \frac{\delta \omega_i}{\omega_i} \simeq \frac{\delta_x c^2}{c^2}(r_0), \quad (2.5.40)$$

$$\sum_{i \in \mathcal{D}} \mathcal{C}_i^\rho(r_0) \frac{\delta \omega_i}{\omega_i} \simeq \frac{\delta_x \rho}{\rho}(r_0) \quad (2.5.41)$$

ou autrement dit : la combinaison $\sum_i (\delta \omega_i / \omega_i) \mathcal{C}_i^{c^2}$ fournit une approximation du profil $\delta_x c^2 / c^2$ et $\sum_i (\delta \omega_i / \omega_i) \mathcal{C}_i^\rho$ du profil $\delta_x \rho / \rho$. En pratique, l'ampleur de l'approximation dépend de la précision sur l'estimation des fréquences d'oscillations ω_i mais surtout d'à quel point il est possible d'imposer les conditions (2.5.38) et (2.5.39). Si ses critères sont impossibles à satisfaire strictement, tout l'enjeu des méthodes d'inversions dites OLA (*optimally localized average*) (BACKUS et GILBERT, 1968; PIJPERS et THOMPSON, 1992; PIJPERS et THOMPSON, 1994) est d'en proposer une satisfaction optimale. Un des facteurs les plus déterminant est évidemment le nombre de fréquences observées, $|\mathcal{D}|$. À l'instar d'un développement en série, la somme $\sum_{i \in \mathcal{D}} \mathcal{C}_i^{c^2}(r_0) \mathcal{K}_{c^2}(r, \omega_i)$ sera d'autant plus proche d'un Dirac qu'il n'y aura de termes. Elle n'atteindra jamais cette limite cependant, même avec un nombre fréquences très élevé, ce qui pousse en pratique à imposer d'autres "fonctions cibles"⁵⁴ (JEFFREY, 1988), comme un profil gaussien centré en r_0 , $\mathcal{G}_\Delta(r - r_0)$ (PIJPERS et THOMPSON, 1992). La largeur Δ (on souhaite qu'elle soit la plus faible possible pour que le profil gaussien se rapproche au mieux d'un Dirac) qu'il sera possible d'atteindre dépend également de $|\mathcal{D}|$ (pour une position r_0 donnée). Dès lors, plutôt qu'une valeur au point précis r_0 , le terme de droite des équations (2.5.40) et (2.5.41) peut être vu comme l'estimation d'une moyenne locale autour de r_0 . L'information récupérée sera donc nécessairement parcellaire mais d'autant plus fine que l'on disposera de fréquences.

Naturellement, ces méthodes ont été tout d'abord appliquées à des observations solaires, fournies par les réseaux BISON (CHAPLIN et al., 1996) et GONG (HARVEY et al., 1996) ou encore l'instrument embarqué GOLF (GABRIEL et al., 1995; GABRIEL et al., 1997). Le nombre considérable de fréquences observées, et ce pour des degrés ℓ très élevés ($\ell > 1000$, CHRISTENSEN-DALSGAARD, 2002), a permis d'enrichir considérablement notre compréhension de la structure solaire. À l'aide de la méthode présentée plus haut, il a notamment été possible d'identifier très précisément la localisation de la transition entre les régions radiative et convective de l'étoile à l'aide du profil de $\delta_x c^2 / c^2$ (voir FIGURE 2.9) : $r_{\text{BCZ}} / R \simeq 0.713$ (CHRISTENSEN-DALSGAARD et al., 1985; BASU et ANTIA, 1997; ANTIA, BASU et CHITRE, 1998; BASU et ANTIA, 2004; SERENELLI, HAXTON et PEÑA-GARAY, 2011). L'estimation des différences de structure dans les régions plus superficielles du Soleil ont également permis de comparer les prédictions de nos équations d'états les plus avancées (à savoir les équations MHD (HUMMER et MIHALAS, 1988; MIHALAS, DAPPEN et HUMMER, 1988; DAEPPEN et al., 1988) et OPAL (ROGERS, SWENSON et IGLESIAS, 1996)) avec la structure interne elle-même. Enfin, la mesure de séparations rotationnelles (*rotational splittings* _[en]) dans les fréquences d'oscillation $\Delta \omega = \omega_m - \omega_{m=0}$ ont fourni des contraintes sur le profil de rotation interne du Soleil, $\Omega(r, \theta)$, par inversion. Ces résultats ont notamment conduit à une cartographie précise de la rotation différentielle au sein l'enveloppe convective (ANTIA, BASU et CHITRE, 1998; SCHOU et al., 1998; THOMPSON et al., 2003; HOWE, 2009). Beaucoup d'autres contraintes peuvent être obtenues par l'inversion de données stellaires, même s'il est aujourd'hui difficile de trouver un modèle de structure les satisfaisant toutes (BULDGEN et al., 2019a).

2.5.6 Le glitch sismique

On peut alors se demander comment procéder lorsque le nombre de fréquences se fait plus faible ou, de manière équivalente, lorsque les variations de structure sont plus locales. Il s'agit du cadre où l'on se situe par défaut pour l'observation des étoiles autres que le Soleil. Les degrés des oscillations observées ne dépassent pas $\ell = 3$ et la précision sur la mesure de leur fréquence est

54. Cela tient notamment du fait que le problème inverse (2.5.32) est mal posé.

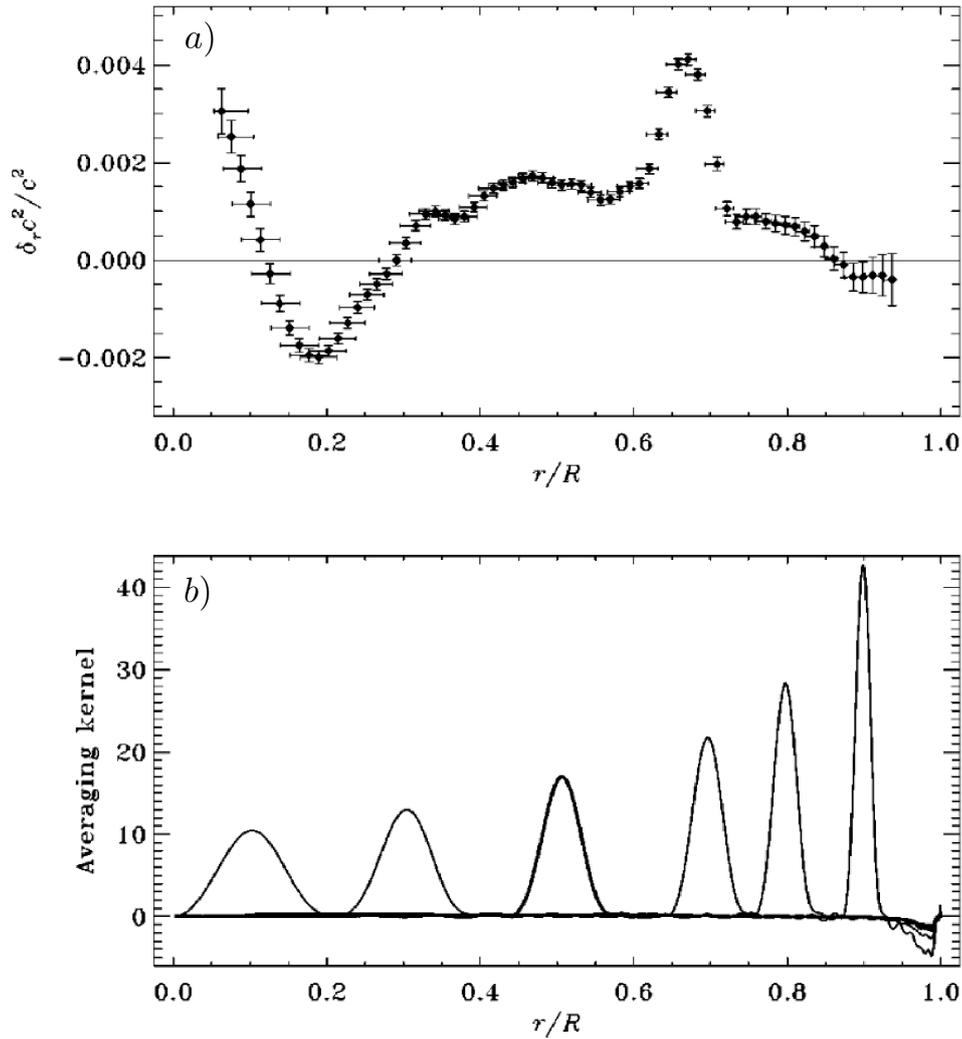


FIGURE 2.9 – Adapté de **BASU et al., 1997**. a) Inversion de la différence du profil de vitesse du son $\delta_x c^2 / c^2$ par rapport au modèle de référence ModelS (**CHRISTENSEN-DALSGAARD et al., 1996**) en exploitant les données du réseau BiSON (**CHAPLIN et al., 1996**). Le “pic” visible à $r/R \simeq 0.7$ trahit la présence de la transition entre les régions radiative et convective du Soleil. b) Exemple de combinaisons de kernels $\sum_{i \in \mathfrak{D}} \mathcal{G}_i^{c^2}(r_0) \mathcal{K}_{c^2}(r, \omega_i)$ pour des fonctions cibles gaussiennes et des profondeurs $r_0 = 0.1, 0.3, 0.5, 0.7, 0.8, 0.9$.

évidemment plus faible que pour le Soleil. Si l’inversion sismique telle que présentée plus haut permet toujours d’obtenir des informations sur la structure (et en particulier sur les quantités intégrées, cf. **REESE et al., 2012**; **BULDGEN et al., 2019b**), il est probable qu’elle ne puisse accéder aux détails les plus fins. En effet, en supposant qu’il soit possible de vérifier exactement les conditions suivantes :

$$\sum_{i \in \mathfrak{D}} \mathcal{G}_i^{c^2}(r_0) \mathcal{K}_{c^2}(r, \omega_i) = \mathcal{G}_\Delta(r_0 - r) \quad \text{et} \quad \sum_{i \in \mathfrak{D}} \mathcal{G}_i^{c^2}(r_0) \mathcal{K}_\rho(r, \omega_i) = 0, \quad (2.5.42)$$

la relation (2.5.37) nous donne⁵⁵ :

$$\sum_{i \in \mathfrak{D}} \mathcal{G}_i^{c^2} \frac{\delta \omega_i}{\omega_i} = \mathcal{G}_\Delta * \frac{\delta_x c^2}{c^2} \quad (2.5.43)$$

55. Dans le cas où la largeur visée Δ n’est pas indépendante de r_0 , la relation (2.5.37) ne se réduit pas à une simple convolution.

où * désigne l'opération de convolution. Il est donc difficile de résoudre précisément des variations d'étendue inférieure à Δ , qui nous donne l'échelle de la convolution par le noyau gaussien. Celle-ci est limitée par le nombre de fréquences disponibles puisque la condition (2.5.42) devient impossible à satisfaire pour des Δ trop faibles.

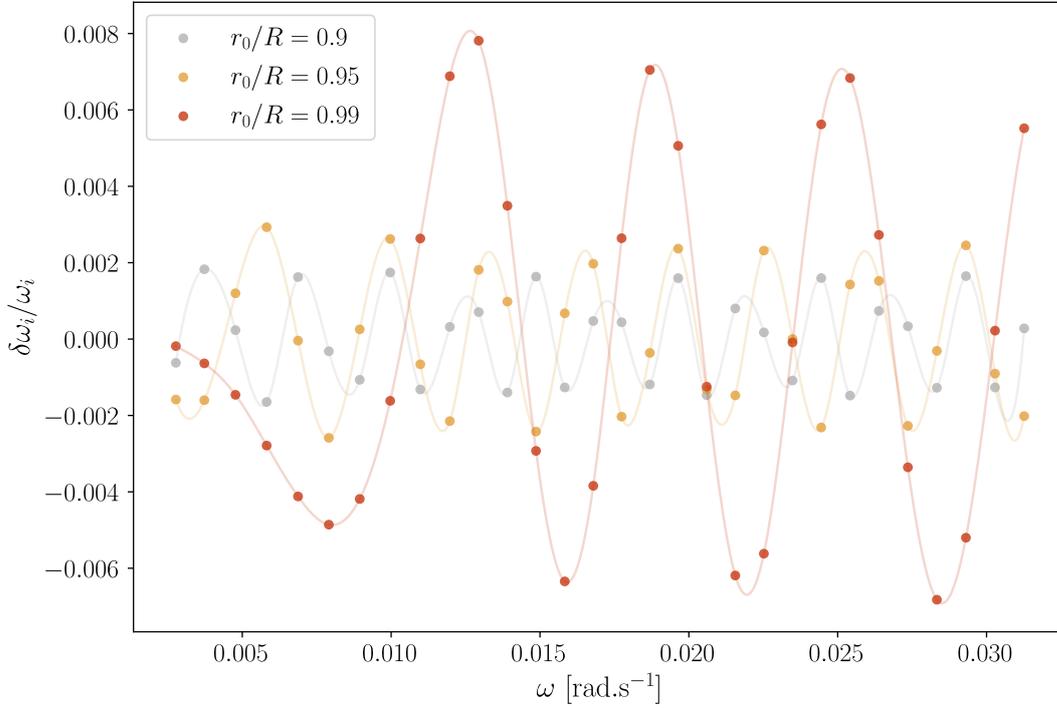


FIGURE 2.10 – Illustration du glitch sismique $\delta\omega_i/\omega_i$ obtenu pour les perturbations de structure $\delta_x\rho/\rho \propto \delta(r_0 - r)$ et $\delta_x c^2/c^2 = 0$ par rapport à un modèle solaire. Les différences de fréquences sont données seulement pour les modes radiaux $\ell = 0$.

Dans un tel contexte, on peut s'interroger sur la signature sismique qu'engendrerait une différence de structure très localisée, et même la plus localisée possible. Pour donner une version idéale de ce problème, supposons que :

$$\frac{\delta_x\rho}{\rho}(r) \propto \delta(r - r_0) \quad \text{et} \quad \frac{\delta_x c^2}{c^2}(r) = 0. \quad (2.5.44)$$

D'après la relation (2.5.32), les variations observées sur les fréquences d'oscillations prennent simplement la forme

$$\frac{\delta\omega_i}{\omega_i} \propto \mathcal{K}_\rho(r_0, \omega_i). \quad (2.5.45)$$

La grandeur $\mathcal{K}_\rho(r_0, \cdot)$ est une fonction oscillante de la fréquence ω , comme le montre la FIGURE 2.10. On y remarquera que la période de cette oscillation (entre autre) ne semble pas indépendante de la position de la perturbation de structure introduite, r_0/R . Moyennant quelques approximations, on verra par la suite qu'il est relativement facile de montrer que la signature sismique peut être approchée par une fonction sinusoïdale (GOUGH, 1990; MONTEIRO, CHRISTENSEN-DALSGAARD et THOMPSON, 1994) :

$$\mathcal{K}_\rho(r_0, \omega_i) = A(\omega_i) \cos(2\omega_i\tau_0 + \phi(\omega_i)). \quad (2.5.46)$$

avec $\tau_0 \equiv \int_{r_0}^R dr/c$, la profondeur acoustique de la perturbation. Dans la suite du manuscrit, on

référer à une telle signature par le terme de *glitch sismique*⁵⁶. Autrement dit, la seule mesure de la période d'oscillation de la fonction $\delta\omega_i/\omega_i$ nous fournit exactement la localisation de la perturbation de structure dans l'étoile⁵⁷. En supposant que l'on mesure cette dernière comme étant égale à π/τ_0^* , il est possible de fournir une estimation de la position, r_0^* , et donc de déterminer la différence de structure avec notre modèle en tous points avec l'expression (2.5.44) (en remplaçant le paramètre r_0 par notre estimation r_0^*). Cela est d'autant plus impressionnant que la variation de structure est on ne peut plus locale (il s'agit d'un Dirac), et donc d'autant plus complexe à résoudre par inversion. On dispose pourtant d'une signature dans les observables de l'étoile au travers des $\delta\omega_i$, et, qui plus est, particulièrement explicite.

Notons qu'il est également possible de tirer des informations de l'amplitude $A(\omega_i)$ ou de la phase $\phi(\omega_i)$ du glitch (voir notamment : GOUGH, 1993 ; MONTEIRO, CHRISTENSEN-DALSGAARD et THOMPSON, 1994 ; BASU, ANTIA et NARASIMHA, 1994). La forme de ces fonctions vont en effet largement dépendre de la forme et de la position de la perturbation de structure (MONTEIRO et THOMPSON, 2005 ; HOUDEK et GOUGH, 2007 ; HOUDEK et GOUGH, 2011). À la différence des méthodes d'inversions présentées dans le paragraphe précédant, lesquelles ne supposaient pas de forme a priori pour la différence de structure (ou même le glitch), il est ici nécessaire de faire un travail théorique préalable afin de déterminer une expression pour le glitch. En effet, des informations sur la structure ne peuvent être déduites que si un lien entre les propriétés physiques et l'apparence du glitch a été établi (ici, entre la période de la signature et la position de la différence de structure).

Bien sûr, en pratique, une différence de structure entre un modèle stellaire et une véritable étoile (ou même entre deux modèles) a peu de chance d'être décrite par les relations (2.5.44). En effet, il semble irréaliste qu'une différence de densité aussi prononcée ($\propto \delta(r - r_0)$) n'engendre aucune contrepartie sur la vitesse du son. Puisque les structures à l'équilibre du modèle et de l'étoiles doivent notamment vérifier les équations (2.4.9)-(2.4.11), la différence entre ces structures ne peut être arbitraire. En supposant malgré tout les hypothèses vérifiées, on risque non seulement de faire une mauvaise estimation de la structure de l'astre observées (laquelle est donnée par les quantités $\rho + \delta_x\rho, c^2 + \delta_x c^2, \dots$), mais surtout d'en faire une estimation non physique.

Finalement, cette exemple très simple permet de mettre en lumière tout ce qui fait l'intérêt de l'exploitation du glitch sismique. Cette méthode permet d'obtenir des contraintes particulièrement forte sur la structure d'un astre dont on observerait les fréquences d'oscillation, et ce, d'autant plus que l'on cherche à apprécier une variation locale. En contrepartie, la procédure doit reposer sur un modèle cohérent (au sens physique) de la perturbation de structure. En effet, le gain d'information retrouvé vis-à-vis de la structure interne se fait au prix d'une injection d'information a priori sur la forme de la différence de structure. Cela peut engendrer des biais conséquents si celle-ci résulte d'une modélisation défailante. Il faut également y ajouter un travail théorique afin de relier l'apparence de la perturbation à des contraintes physiques dans un premier temps, puis à l'expression du glitch sismique lui-même. Cette étape est essentielle à l'ajustement (*fit* [en]) des paramètres nous renseignant sur ces mêmes contraintes. Dans la prochaine partie, on se propose ainsi de voir comment cette méthode peut être employée à la détermination de l'abondance d'hélium mais également d'autres propriétés physiques d'une étoile, tâche qui a constitué l'essentiel de mon travail de thèse.

56. Il existe une certaine ambiguïté autour de ce que désigne réellement le terme de glitch dans la littérature. Certains auteurs l'utilisent pour référer à la signature sismique tandis que d'autres l'emploient pour parler de la cause structurelle de cette signature. Dans ce manuscrit, la distinction sera clairement faite entre les *différences/perturbations de structure* ($\delta\rho, \delta c^2, \dots$) et le *glitch sismique* engendré ($\delta\omega_i$).

57. En réalité, une ambiguïté subsiste entre les profondeurs τ_0 et $\pi/\Delta\omega - \tau_0$, avec $\Delta\omega$ la séparation entre deux fréquences successives (laquelle est à peu près constante, comme on le verra par la suite). En effet, $\cos(2(\pi/\Delta\omega - \tau_0)\omega_i + \phi) \simeq \cos(2\pi n - 2\omega_i\tau_0 + \phi) = \cos(2\omega_i\tau_0 - \phi)$, ce qui rend impossible la distinction de ces deux profondeurs par la seule observation de la période d'oscillation (égale à π/τ_0 dans les deux cas).

Deuxième partie

Partie principale

Chapitre 3

Le glitch sismique

Contenu du chapitre

3.1	Les origines des glitches sismiques	95
3.2	Le glitch de l'ionisation	97
3.3	Modélisation de la perturbation de structure	99
3.4	Expressions usuelles du glitch	100
3.5	Modèle de référence et diagnostic sismique	104
3.6	Limites et discussion	106
3.7	Limits and discussion [en]	110

Ce court chapitre nous permettra de nous attarder sur la notion de glitch sismique au-delà de l'introduction théorique que l'on vient d'en faire. On y abordera la question de l'origine des glitches ainsi que de l'information qu'ils transportent sur la structure stellaire. On donnera des exemples d'utilisation du glitch de l'ionisation dans des travaux précédents, avant de voir comment améliorer ces approches.

3.1 Les origines des glitches sismiques

On a vu dans le chapitre précédent ce qu'était un glitch sismique, l'émergence d'une signature oscillatoire au sein des perturbations des fréquences d'oscillation lors d'une différence de structure suffisamment abrupte. Notons comme cette définition est ancrée dans le formalisme de l'approche perturbative ; parler d'un glitch sismique nécessite de définir des fréquences d'oscillation de référence associées à une structure de référence. Ainsi, définir le glitch sismique comme une propriété propre à un unique astre que l'on aurait observé peut être vu comme un abus de langage. C'est pourtant ce que l'on a fait dans le paragraphe introductif 1.5 (et que l'on fera de nouveau par la suite). Dans un tel cas, la définition du modèle de référence se fait de manière implicite. Par exemple, alors que je faisais référence à l'origine structurelle du glitch de l'ionisation : *Ainsi, dans cette zone, certaines grandeurs physiques comme la température T (ou la vitesse de propagation des ondes, c , qui lui est directement reliée) ont tendance à augmenter "moins vite qu'elles ne le devraient"*. Ici, le "moins vite qu'elles ne le devraient" réfère à ce qu'il se passerait si jamais le processus d'ionisation n'intervenait pas dans la structure. On peut donc voir la définition de la référence comme un modèle où l'ionisation serait absente. Le glitch sismique ainsi défini correspond donc aux différences de fréquences causées par l'impact de l'ionisation sur la structure. On voit bien, par cet exemple, que le modèle de référence ne nécessite pas d'être plausible : il est clair que l'ionisation intervient bel est bien dans la structure de l'étoile. Il se doit seulement d'être un intermédiaire et en particulier, doit permettre le calcul de ses fréquences propres d'oscillation, lesquelles serviront de référence. Si cette considération vis-à-vis de la définition du glitch peut sembler anecdotique, il me semble important de toujours garder à l'esprit ce qu'est le modèle de

référence pour un glitch donné. On verra que cela peut avoir un grand impact sur la méthode et l'exploitation des résultats.

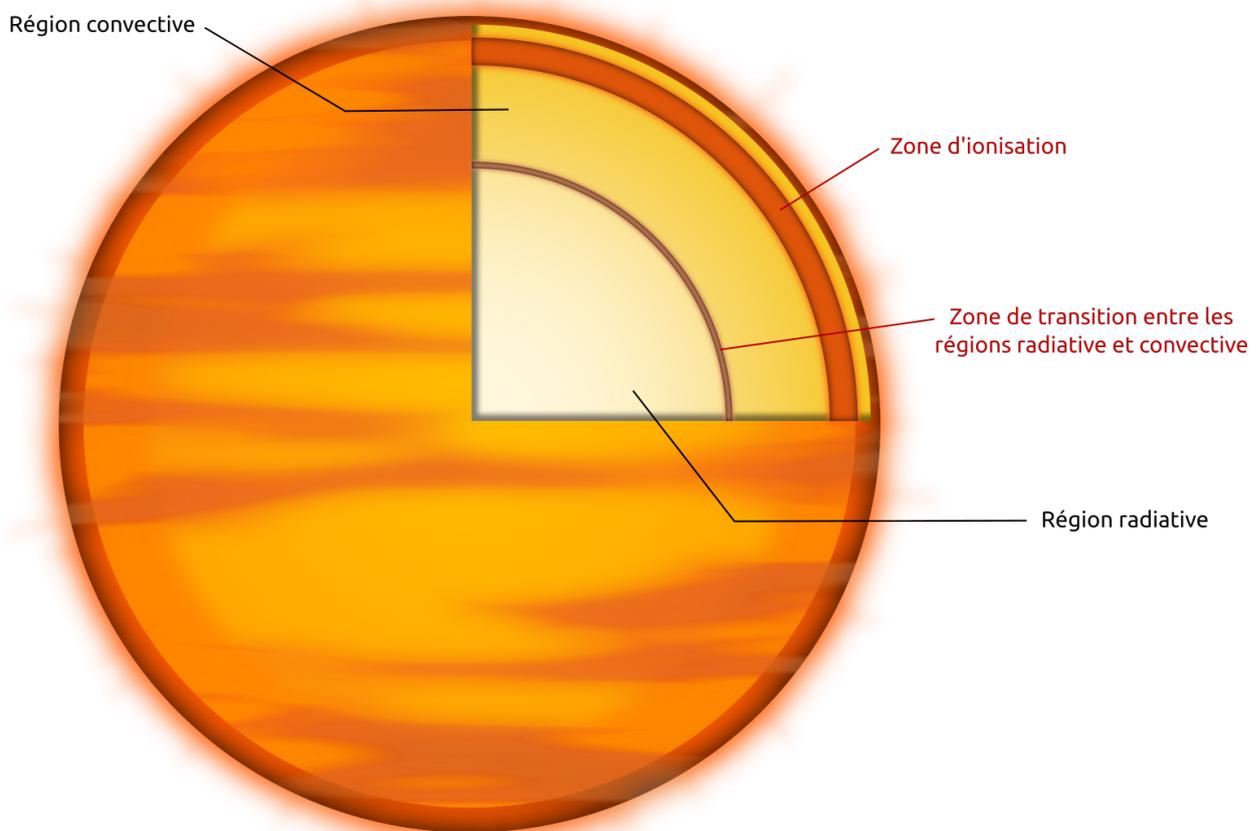


FIGURE 3.1 – Représentation des régions à même de causer un glitch sismique pour une étoile de type solaire. La première est située dans la zone d'ionisation tandis que la seconde provient de la transition brutale entre les régions dominées par le transport radiatif et convectif.

Ce point éclairci, les pulsateurs de type solaire disposent de deux régions à même de causer un glitch sismique. Celles-ci sont représentées en FIGURE 3.1. La première, la plus profonde, est située à l'interface des régions radiative et convective de l'étoile. À cet endroit, au niveau de la base de cette zone convective (que l'on abrégiera par l'acronyme BCZ pour "Base of the Convection Zone"), le transport d'énergie dominant change sur une courte distance et les variations de structures engendrées sont à l'origine de ce que l'on appellera ici le *glitch de la BCZ*. Ce glitch a notamment été utilisé pour essayer de mieux contraindre le phénomène d'overshoot introduit dans le paragraphe 2.3.6 (MONTEIRO, CHRISTENSEN-DALSGAARD et THOMPSON, 1994; BASU et ANTIA, 1994b; CHRISTENSEN-DALSGAARD, MONTEIRO et THOMPSON, 1995; CHRISTENSEN-DALSGAARD et al., 2011). En effet, de nombreux modèles existent pour expliquer la manière et l'étendue sur laquelle les mouvements convectifs peuvent pénétrer la région radiative. Chacun possède un impact différent sur les propriétés structurelles locales, que ce soit au niveau du mélange des éléments (BASU, ANTIA et NARASIMHA, 1994), des discontinuités des dérivées de la densité ρ et de la vitesse du son c (MONTEIRO, CHRISTENSEN-DALSGAARD et THOMPSON, 1994; CHRISTENSEN-DALSGAARD, MONTEIRO et THOMPSON, 1995; BASU et ANTIA, 1994a; HOUDEK et GOUGH, 2007) ou encore du flux d'énergie entraîné par les mouvements de pénétration (ZAHN, 1991; REMPEL, 2004; ROGERS, GLATZMAIER et JONES, 2006). Un autre attrait du glitch de la BCZ provient de sa capacité à localiser cette discontinuité dans l'étoile, comme on l'a vu dans le paragraphe 2.5.6. Si ces astres ont pour propriété de posséder une enveloppe convective surplombant

leur zone radiative¹, et sont donc à même d'exhiber un tel glitch, la position de la BCZ dépend beaucoup de la masse de l'étoile (MONTEIRO, CHRISTENSEN-DALSGAARD et THOMPSON, 2000). Une localisation précise de cette zone permettrait (en plus des contraintes qu'elle apporte sur les paramètres stellaires, MAZUMDAR, 2005) d'obtenir des renseignements sur les désaccords dans la composition chimique du soleil entre observables spectroscopiques et sismiques (ASPLUND et al., 2009). En effet, ces abondances impactent directement les opacités qui sont elles-même connues pour influencer sur la taille de l'enveloppe convective (BASU et ANTIA, 2008; GOUGH, 2013; VERMA et al., 2017).

La signature nous intéressant dans ce manuscrit reste cependant celle de la région d'ionisation. Située plus en surface que la BCZ, les variations de propriétés abruptes qui y prennent place (cf. section 1.5, paragraphe 2.3.7) engendrent l'apparition du glitch de l'ionisation. Comme on l'a vu dans l'introduction, un intérêt majeur quant à l'exploitation de cette signature réside dans l'information qu'elle contient sur l'abondance en hélium Y. Outre cette estimation, il s'agit là d'une opportunité de lever la dégénérescence masse-hélium et donc d'obtenir des contraintes sur la masse et la physique de l'étoile².

3.2 Le glitch de l'ionisation

Regardons maintenant plus en détail quelle est l'empreinte structurelle de la région d'ionisation. Comme évoqué dans le paragraphe 2.3.7, un traceur tout indiqué se trouve être le premier exposant adiabatique Γ_1 , défini par l'équation (2.3.41). Une représentation de cet exposant dans la zone d'ionisation d'une étoile de type solaire peut être retrouvée en FIGURE 3.2.

Alors que Γ_1 se situe très proche de 5/3 dans la majeure partie de l'étoile, celui-ci entame une diminution très abrupte à l'entrée de la région d'ionisation ($\tau \sim 100$ s). Après s'être rapproché de la valeur 1, l'exposant adiabatique ré-augmente progressivement pour retrouver la valeur 5/3 au sortir de la zone d'ionisation ($\tau \sim 1000$ s). Ces variations ne se font cependant pas de façon monotone. La plus forte variation est aussi la plus superficielle et est portée par l'ionisation de l'hydrogène (en gris). Suivent ensuite les impacts des première (en bleu) et seconde (en violet) ionisations de l'hélium qui semblent moins marquées. À noter que les ionisations des métaux, en proportion bien plus faible que l'hydrogène ou l'hélium, ne produisent pas d'impact visible sur cette courbe³. Essayons de comprendre l'allure de cette fonction par un court raisonnement théorique. Comme souligné dans le paragraphe 2.3.6, de par l'homogénéité chimique et l'isentropie de la région convective⁴, le premier exposant adiabatique est assimilable à l'exposant effectif γ dans une zone d'ionisation. En utilisant la différentielle liant l'énergie $m\varepsilon$ et l'entropie ms par particule dérivée dans l'annexe A.5 :

$$d(m\varepsilon) = Td(ms) - pd \left(\frac{1}{n} \right) \quad (3.2.1)$$

et puisque $dm = ds = 0$:

$$d\varepsilon = \frac{p}{\rho} d \ln \rho. \quad (3.2.2)$$

1. Ce n'est pas le cas de toutes les étoiles, à partir d'une certaine masse (légèrement supérieure à $1 M_{\odot}$ pour les étoiles de séquence principale, voir KIPPENHAHN, WEIGERT et WEISS, 2012) l'enveloppe convective tend à disparaître et un cœur convectif se développe. La structure de ces étoiles plus massives s'en voit donc inversée : un intérieur convectif pour une enveloppe radiative.

2. À noter que des observations récentes ont mis en évidence un autre type de glitch dans les étoiles plus évoluées, lors de leur phase géante rouge (CUNHA et al., 2015; VRARD et CUNHA, 2019). D'origine bien plus profonde, cette signature n'impacte que les modes de gravité, ce qui explique son absence dans les spectres des étoiles de séquence principale. Il existe de nombreuses hypothèses quant à son origine : fusion de l'hydrogène en couche, inhomogénéités laissées par l'enveloppe convective lors de la progression de la fusion en couche ou même l'impact d'un cœur convectif par overshoot (BOSSINI et al., 2015; BOSSINI et al., 2017)

3. Comme on le verra dans le chapitre 4, l'impact des métaux sur le premier exposant adiabatique Γ_1 se fait plutôt de façon indirecte.

4. Les zones les plus superficielles, connues pour être super-adiabatiques ($\gamma > \Gamma_1$), ne seront pas considérées ici.

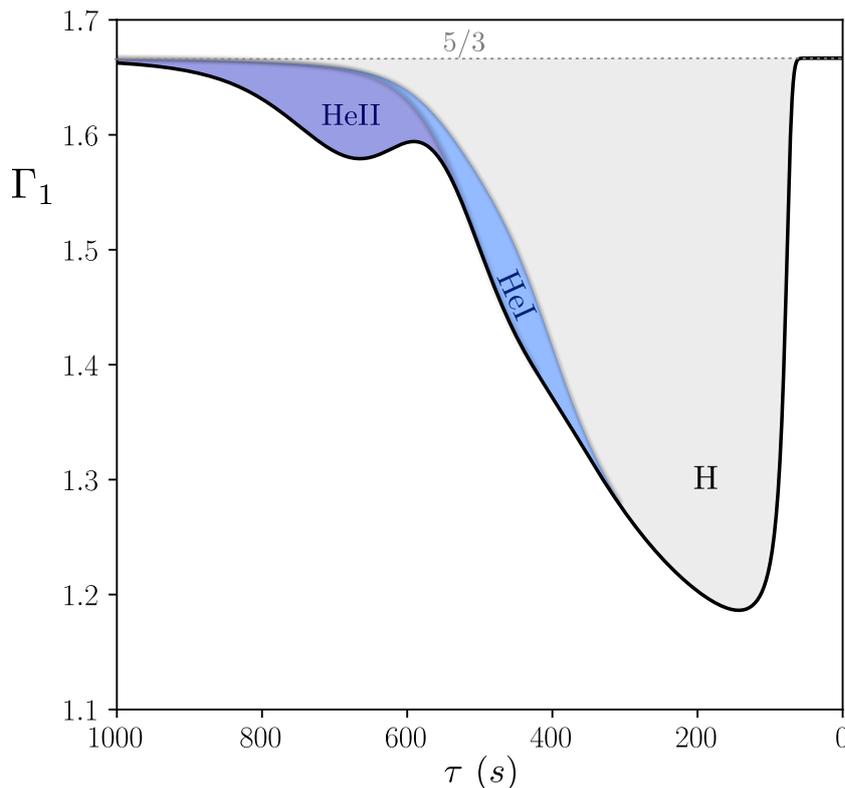


FIGURE 3.2 – Le premier exposant adiabatique, Γ_1 , en fonction de la profondeur acoustique τ (la surface se trouve donc à droite, à $\tau = 0$ s) dans la région d'ionisation pour une étoile de type solaire. Les déviations à la valeur $5/3$ (pointillés gris) causées par l'ionisation sont chacune représentées par une couleur distincte. En gris : ionisation de l'hydrogène (H), en bleu : première ionisation de l'hélium (HeI), en violet : seconde ionisation de l'hélium (HeII).

L'exposant effectif $\gamma \equiv d \ln p / d \ln \rho$ sera donc donné par :

$$\gamma = \frac{1}{\rho} \frac{dp}{d\varepsilon}. \quad (3.2.3)$$

En divisant maintenant l'énergie interne entre l'énergie cinétique (issue de l'agitation thermique) $(3/2)p/\rho$ et une énergie d'ionisation ε_{ion} (cf. annexe A.5), on peut facilement exprimer la dérivée $d\varepsilon/dp$:

$$\frac{d\varepsilon}{dp} = \frac{d\varepsilon}{dp} \frac{d\varepsilon_{\text{ion}}}{d\varepsilon} + \frac{3}{2\rho} \left(1 - \frac{1}{\gamma}\right). \quad (3.2.4)$$

Or, puisque l'équation (3.2.3) donne $d\varepsilon/dp = 1/\gamma\rho$, cette équation n'est autre qu'une relation liant γ à la quantité $d\varepsilon_{\text{ion}}/d\varepsilon$:

$$1 = \frac{d\varepsilon_{\text{ion}}}{d\varepsilon} + \frac{3\gamma}{2} \left(1 - \frac{1}{\gamma}\right), \quad (3.2.5)$$

nous permettant finalement d'écrire :

$$\frac{5}{3} - \gamma = \frac{2}{3} \frac{d\varepsilon_{\text{ion}}}{d\varepsilon}. \quad (3.2.6)$$

Cette équation nous informe directement sur l'allure de la courbe en FIGURE 3.2. En effet, le terme de gauche correspond aux aires colorées (l'écart de la fonction à la valeur $5/3$) et découle donc du rapport entre les variations d'énergies d'ionisation et totale. Remarquons tout d'abord que, puisque la variation d'énergie allouée à l'ionisation est incluse dans la variation d'énergie

totale (et que ε_{cin} et ε_{ion} varient toujours dans le même sens), on a $0 \leq d\varepsilon_{\text{ion}}/d\varepsilon < 1$ et donc⁵ :

$$1 < \gamma \leq \frac{5}{3}. \quad (3.2.7)$$

Les limites ne sont atteintes que dans les deux cas où les variations d'énergie interne⁶ ont pour unique origine l'ionisation ($\gamma \rightarrow 1$) ou au contraire l'énergie cinétique du gaz ($\gamma = 5/3$). Entre ces deux extrêmes se trouve un spectre sur lequel le premier exposant adiabatique varie dans la région d'ionisation. Ainsi, les "puits" creusés par chaque élément sur la FIGURE 3.2 correspondent à des régions où l'ionisation se fait particulièrement intense, jusqu'à contribuer à une part non négligeable de l'énergie localement disponible. Avec ceci à l'esprit, on voit que l'ionisation requérant le plus d'énergie est celle de l'hydrogène (H), ce qui s'explique notamment par sa fraction en nombre dominante : $x_{\text{H}} \sim 92\%$ pour une fraction massique d'hélium $Y \sim 25\%$. Bien que l'énergie nécessaire à ioniser l'hélium (que ce soit pour les transitions HeI ou HeII) est supérieure, son impact se fait moins sentir sur la courbe. Outre sa présence moindre ($x_{\text{He}} \sim 8\%$), c'est aussi l'énergie cinétique qui se fait plus forte au fur et à mesure que l'on s'enfonce dans l'étoile (et donc que τ augmente).

3.3 Modélisation de la perturbation de structure

Comme présentée dans le paragraphe 2.5.6, l'exploitation du glitch doit reposer sur une modélisation de la perturbation de structure, en l'occurrence celle de la région d'ionisation. Si toutes les grandeurs structurelles (comme la pression, la densité, la température, ...) s'en trouvent impactées, il est naturel de s'intéresser aux variations particulièrement notables observées sur le premier exposant adiabatique, Γ_1 . Puisque l'objectif affiché est de procéder à une estimation de l'abondance en hélium, l'impact à modéliser est donc la conséquence d'une variation en hélium⁷ δY sur ce profil. La première zone d'ionisation de l'hélium étant dominée par celle de l'hydrogène (cf. FIGURE 3.2), la discussion a majoritairement tourné autour d'une modélisation de l'impact de la seconde zone d'ionisation de l'hélium (MONTEIRO et THOMPSON, 1998; GOUGH, 2002; BASU et al., 2004; MONTEIRO et THOMPSON, 2005; HOUDEK et GOUGH, 2007; HOUDEK et GOUGH, 2011). Cette signature cause un puits dans le profil de Γ_1 , et il semble naturel de penser que son amplitude est directement liée à l'abondance d'hélium (ce qu'il est facile de vérifier par des résultats numériques (HOUDEK et GOUGH, 2007)). Ainsi une augmentation de l'abondance d'hélium ($\delta Y > 0$) a tendance à creuser le puits HeII, plus d'énergie étant allouée à l'ionisation de cet élément (et réciproquement, voir FIGURE 3.3). Plusieurs modèles utilisés pour reproduire ce comportement (ainsi que leur paramétrisation) sont représentés en FIGURE 3.4.

Chacun des exemples y apparaissant possède ses avantages comme ses inconvénients. Le dirac peut être intéressant puisqu'il ne nécessite que deux paramètres pour être complètement défini, en plus de rendre la dérivation de la signature sismique associée relativement simple (cf. raisonnement du paragraphe 2.5.6 par exemple). D'un autre côté, il reste une approximation plutôt grossière de ce qu'un changement d'hélium induit au niveau de la structure. Le profil triangulaire permet de s'approcher un peu plus de la variation réelle du premier exposant adiabatique au prix d'un paramètre supplémentaire contrôlant l'étendue de la perturbation de structure⁸. Le glitch sismique reste relativement facile à dériver, le profil de perturbation ne résultant que de l'assemblage de deux fonctions affines. Ce profil triangulaire peut aussi être troqué pour

5. Notons que $d\varepsilon_{\text{ion}} = d\varepsilon$, correspondant au cas $\gamma = 1$, n'advient que sous l'hypothèse où $d(p/\rho) = 0$ et donc que $dT = d(1 + x_e) = 0$ (cf. équation (A.5.16)). Or, dans un tel cas, il est clair que l'on a également $d\varepsilon_{\text{ion}}$ d'après la définition (A.5.22). Cela montre pourquoi la limite $\gamma \rightarrow 1$ est de fait inatteignable.

6. On entend ici par variations celles causées par un changement de position dans la structure par exemple puisque $d\varepsilon_{\text{ion}}/d\varepsilon = (d\varepsilon_{\text{ion}}/dr)(d\varepsilon/dr)^{-1}$.

7. Contrairement aux autres profils, on interprétera l'abondance d'hélium Y comme une quantité indépendante de la position r (ou τ), les abondances étant homogènement distribuées dans la zone convective.

8. Un quatrième paramètre, α , est également introduit dans MONTEIRO et THOMPSON, 2005, permettant de prendre en compte une asymétrie dans le profil

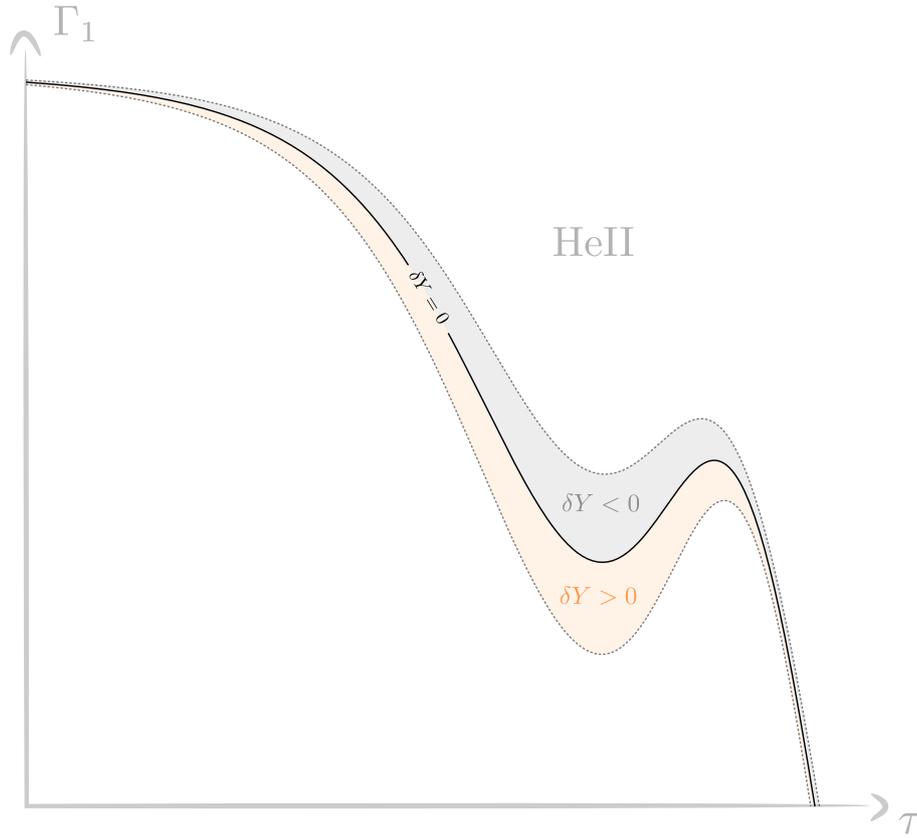


FIGURE 3.3 – Idéalisation de l’impact d’une variation d’hélium δY sur le premier exposant adiabatique Γ_1 dans la seconde région d’ionisation de l’hélium. L’enjeu pour la modélisation de la perturbation de structure est alors de trouver une fonction capable de reproduire adéquatement les aires colorées.

une gaussienne, qui, pour le même nombre de paramètres, permet une restitution plus fidèle de l’impact d’une variation δY en plus d’éviter l’introduction de discontinuités artificielles dans les dérivées de la vitesse du son (HOUDEK et GOUGH, 2007). La dérivation du glitch de l’ionisation sera cependant rendue plus complexe.

3.4 Expressions usuelles du glitch

On se propose ici de voir par quels moyens procéder à une telle dérivation analytique afin d’obtenir une expression du glitch sismique satisfaisant la forme théorique (2.5.46). L’expression obtenue dépendra naturellement du choix d’un modèle pour la perturbation $\delta_x \Gamma_1 / \Gamma_1$ causée par une variation d’abondance δY , on la supposera néanmoins indéterminée pour le moment. À noter que le glitch ne peut être déterminé à partir de la formule (2.5.32) qui, en l’état, dépend des perturbations $\delta_x \rho / \rho$ et $\delta_x c^2 / c^2$ et non pas de $\delta_x \Gamma_1 / \Gamma_1$. Néanmoins, en approximant \mathcal{K}_ρ par sa composante dominante :

$$\mathcal{K}_\rho \simeq \mathcal{K} \equiv \frac{\rho r^2 c^2 \Lambda^2}{2\omega^2 \mathcal{I}}, \quad (3.4.1)$$

on remarque qu’il est possible de factoriser nos deux perturbations de structure :

$$\frac{\delta\omega}{\omega} \simeq \int_0^R \mathcal{K}(r, \omega) \left[\frac{\delta_x \rho}{\rho} + \frac{\delta_x c^2}{c^2} \right] dr. \quad (3.4.2)$$

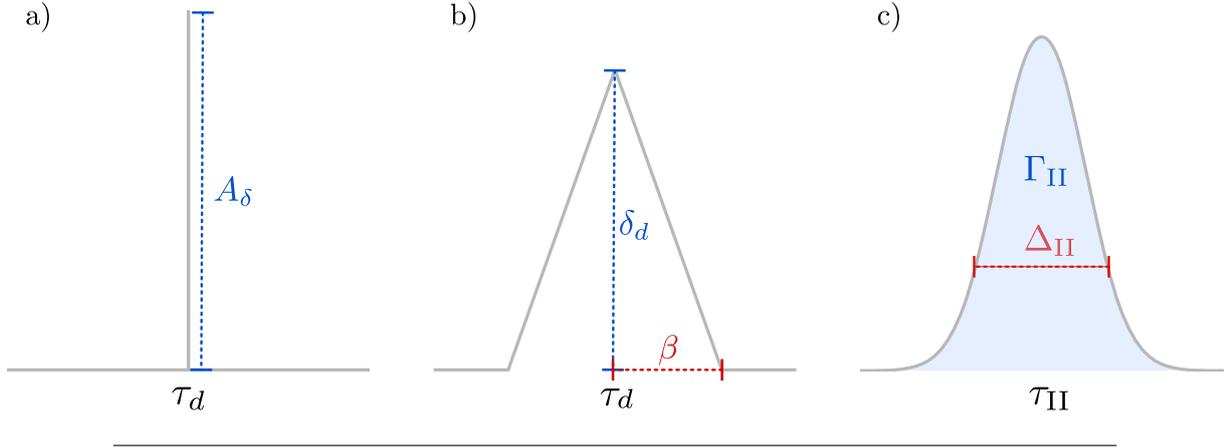


FIGURE 3.4 – Trois exemples de modèles de perturbations de structure (avec leurs paramètres) à l’origine d’un glitch. **a)** Un dirac utilisé par [MONTEIRO, CHRISTENSEN-DALSGAARD et THOMPSON, 1994](#) pour modéliser la perturbation du potentiel acoustique au niveau de la BCZ en présence d’overshoot (puis adopté dans les travaux de [MAZUMDAR et ANTIA, 2001](#); [BASU et al., 2004](#); [BALLOT, TURCK-CHIÈZE et GARCÍA, 2004](#); [PIAU, BALLOT et TURCK-CHIÈZE, 2005](#) pour la seconde région d’ionisation de l’hélium). Le paramètre A_δ détermine alors l’amplitude de la discontinuité tandis que τ_d fixe sa position. **b)** Profil “triangulaire” que l’on retrouve dans les travaux de [MONTEIRO et THOMPSON, 1998](#); [BASU et al., 2004](#); [MONTEIRO et THOMPSON, 2005](#) pour modéliser la signature de la seconde région d’ionisation de l’hélium dans le profil du premier exposant adiabatique (voir [FIGURE 3.2](#)). Un paramètre supplémentaire (β) apparaît pour jouer sur la largeur du “puits”. **c)** Alternative au modèle **b)** proposée par [HOUDEK et GOUGH, 2007](#). La perturbation de l’exposant Γ_1 est dorénavant gaussienne (et donc lisse) pour approcher au plus près la signature réelle de la seconde région d’ionisation de l’hélium (HeII). Les paramètres Γ_{II} , τ_{II} et Δ_{II} contrôlent respectivement l’aire, la position et l’écart-type du profil gaussien.

Maintenant, puisque $\rho c^2 = \Gamma_1 p$, il est facile de faire apparaître la perturbation de structure nous intéressant :

$$\frac{\delta\omega}{\omega} \simeq \int_0^R \mathcal{K}(r, \omega) \left[\frac{\delta_x \Gamma_1}{\Gamma_1} + \frac{\delta_x p}{p} \right] dr. \quad (3.4.3)$$

Celle-ci n’apparaît cependant pas de façon isolée; elle est accompagnée de la perturbation relative de la pression, $\delta_x p/p$, dont la modélisation ne découle pas immédiatement de celle de $\delta_x \Gamma_1/\Gamma_1$. L’argument avancé pour l’éliminer est le suivant. En exprimant le laplacien de la pression par exemple :

$$\begin{aligned} \Delta p &= -g \frac{d\rho}{dr} - \rho \underbrace{\left[\frac{dg}{dr} + \frac{2g}{r} \right]}_{\Delta\Phi=4\pi G\rho} \\ &= \rho^2 \left[\frac{g^2}{\gamma p} - 4\pi G \right] \end{aligned} \quad (3.4.4)$$

et puisque γ est assimilable à Γ_1 dans la région d’ionisation, on voit que le premier exposant adiabatique n’est pas directement lié à la pression, ni même à sa dérivée mais seulement à sa dérivée seconde. Puisque l’objectif est de modéliser l’oscillation en fréquence issue d’une variation abrupte dans la structure, il paraît alors vraisemblable que la perturbation de pression n’a que peu d’impact sur cette oscillation. En effet, si l’on a vu que le premier exposant adiabatique manifeste bien cette variation abrupte dans la région d’ionisation (cf. [FIGURE 3.2](#)), celle-ci doit être bien moins nette dans le profil de pression, lequel résulte de l’intégration par deux fois de Γ_1 . Ainsi, une approximation usuelle pour le glitch se trouve plutôt être (voir par exemple [GOUGH,](#)

2002; HOUDEK et GOUGH, 2007) :

$$\frac{\delta\omega}{\omega} \simeq \int_0^R \frac{\delta_x \Gamma_1}{\Gamma_1}(r) \mathcal{K}(r, \omega) dr. \quad (3.4.5)$$

Ayant supposé une modélisation de $\delta_x \Gamma_1 / \Gamma_1$ (typiquement avec un des profils proposés en FIGURE 3.4), il ne reste qu'à trouver une approximation de \mathcal{K} en fonction de la fréquence ω pour dériver une forme explicite du glitch. Une telle expression peut être obtenue par une analyse JWKB, comme présentée en section 2.4.11. En reprenant l'approximation à l'ordre 0 (2.4.112), et en réexprimant Ξ en fonction de $c\Lambda$ (avec Λ la divergence du champ de déplacement) à l'aide de l'équation (B.3.26) dans la limite des hautes fréquences :

$$\Xi^2 \simeq -\frac{4\pi r^2 \rho c}{\omega^2} c^2 \Lambda^2, \quad (3.4.6)$$

il vient l'approximation :

$$\rho r^2 c^2 \Lambda^2 \propto \frac{\omega^2}{c} \sin^2(\omega\tau + \psi) \quad (3.4.7)$$

avec ψ une constante à l'ordre 0 (cf. approximation 2.4.124). À la vue de l'équation (3.4.1), il ne reste qu'à trouver une approximation de l'inertie du mode :

$$\mathcal{I} \equiv \int_0^R \rho r^2 (\xi_r^2 + \ell(\ell+1)\xi_h^2) dr \quad (3.4.8)$$

comme fonction de la fréquence. Toujours dans la limite des hautes fréquences, l'équation (B.3.6) nous donne :

$$\frac{d\xi_r}{dr} + \frac{2}{r}\xi_r \simeq \Lambda, \quad (3.4.9)$$

ce qui implique $\xi_h \simeq 0$ en plus de lier ξ_r et Λ par la relation suivante :

$$\frac{dr^2 \xi_r}{dr} \simeq r^2 \Lambda \propto \frac{\omega}{c\sqrt{\rho c}} \sin(\omega\tau + \psi). \quad (3.4.10)$$

On peut réécrire cette relation de proportionnalité comme :

$$\frac{dr^2 \xi_r}{d\tau} \propto \frac{1}{\sqrt{\rho c}} \frac{d}{d\tau} \cos(\omega\tau + \psi). \quad (3.4.11)$$

Maintenant, dans la limite où $\omega \rightarrow \infty$, il est possible de négliger les variations de $1/\sqrt{\rho c}$ devant celles de $\cos(\omega\tau + \psi)$, permettant d'intégrer immédiatement l'équation (3.4.11). En remplaçant alors l'équivalent obtenu dans la définition de l'inertie, on obtient l'approximation :

$$\begin{aligned} \mathcal{I} &\simeq \int_0^R \rho r^2 \xi_r^2 dr \\ &\propto \int_0^R \frac{1}{c} \cos^2(\omega\tau + \psi) dr \\ &\propto \int_0^{\mathcal{T}} \cos^2(\omega\tau + \psi) d\tau \simeq \frac{\mathcal{T}}{2}, \end{aligned} \quad (3.4.12)$$

montrant finalement que celle-ci est indépendante de la fréquence à l'ordre 0 et où \mathcal{T} désigne le rayon acoustique total de l'étoile. L'approximation finale pour \mathcal{K} peut donc être directement déduite de (3.4.7) :

$$\mathcal{K} \propto \frac{1}{c\mathcal{T}} \sin^2(\omega\tau + \psi), \quad (3.4.13)$$

ce qui donne pour l'expression du glitch :

$$\delta\omega \propto \frac{\omega}{\mathcal{T}} \int_0^{\mathcal{T}} \frac{\delta_x \Gamma_1}{\Gamma_1}(\tau) \cos(2\omega\tau + 2\psi) d\tau, \quad (3.4.14)$$

où l'on a conservé ici que la partie oscillatoire de l'identité $\sin^2(x) = 1/2 - 1/2 \cos(2x)$. Pour donner une forme plus explicite à cette équation il est nécessaire de définir analytiquement la fonction $\delta_x \Gamma_1 / \Gamma_1(\tau)$ et donc un choix de modèle de perturbation de structure. Le modèle le plus utilisé est le modèle gaussien (cf. FIGURE 3.4) :

$$\frac{\delta_x \Gamma_1}{\Gamma_1}(\tau; \Gamma_{\text{II}}, \tau_{\text{II}}, \Delta_{\text{II}}) = -\frac{1}{\sqrt{2\pi}} \frac{\Gamma_{\text{II}}}{\Delta_{\text{II}}} e^{-(\tau - \tau_{\text{II}})^2 / 2\Delta_{\text{II}}^2} \quad (3.4.15)$$

introduit par [GOUGH, 2002](#) ; [HOUDEK et GOUGH, 2007](#). L'intégrale prend alors la forme :

$$\int_0^{\mathcal{T}} \frac{\delta_x \Gamma_1}{\Gamma_1}(\tau) \cos(2\omega\tau + 2\psi) d\tau = -\frac{\Gamma_{\text{II}}}{\sqrt{2\pi}} \Re \left\{ \int_{-\tau_{\text{II}}/\Delta_{\text{II}}}^{(\mathcal{T} - \tau_{\text{II}})/\Delta_{\text{II}}} \exp(-u^2/2 + 2i\omega\Delta_{\text{II}}u + 2i\omega\tau_{\text{II}} + i\psi) \right\} \quad (3.4.16)$$

où l'on a effectué le changement de variable $\tau \rightarrow u = (\tau - \tau_{\text{II}}) / \Delta_{\text{II}}$. Puisque la perturbation de structure amenant le glitch est, par définition, très localisée par rapport à l'ensemble de la structure, on pourra supposer qu'à la fois $\tau_{\text{II}} / \Delta_{\text{II}} \rightarrow \infty$ et $(\mathcal{T} - \tau_{\text{II}}) / \Delta_{\text{II}} \rightarrow \infty$. À l'aide de la propriété :

$$\int_{-\infty}^{\infty} e^{-(ax^2 + bx + c)} dx = \sqrt{\pi/a} e^{b^2/4a - c}, \quad (3.4.17)$$

il vient :

$$\begin{aligned} \int_0^{\mathcal{T}} \frac{\delta_x \Gamma_1}{\Gamma_1}(\tau) \cos(2\omega\tau + 2\psi) d\tau &= -\Gamma_{\text{II}} \Re \left\{ \exp(-2\Delta_{\text{II}}^2 \omega^2 + 2i\omega\tau_{\text{II}} + i\psi) \right\} \\ &= -\Gamma_{\text{II}} e^{-2\Delta_{\text{II}}^2 \omega^2} \cos(2\omega\tau_{\text{II}} + \psi). \end{aligned} \quad (3.4.18)$$

En incorporant ce développement dans l'équation (3.4.14), on écrira finalement

$$\delta\omega \simeq \mathcal{G}(\omega) \cos(2\omega\tau_{\text{II}} + \psi) \quad (3.4.19)$$

avec \mathcal{G} l'enveloppe du glitch qui dépend de la fréquence (ainsi que du choix de modèle de perturbation) :

$$\mathcal{G}(\omega; \Gamma_{\text{II}}, \Delta_{\text{II}}) \equiv \alpha \Gamma_{\text{II}} \mathcal{T}^{-1} \omega e^{-2\Delta_{\text{II}}^2 \omega^2}, \quad (3.4.20)$$

expression conforme aux équations (15) et (16) de [HOUDEK et GOUGH, 2007](#) et où α désigne l'opposé de la constante de proportionnalité sans dimension de l'équation (3.4.14). Si cette forme correspond à l'expression générale donnée en équation (2.5.46), elle fournit davantage de détails sur les propriétés auxquelles le glitch donne accès. Tout d'abord, on remarque l'amplitude du glitch \mathcal{G} est une fonction évanescence de la fréquence ω ⁹. Si l'on a vu dans le paragraphe 2.5.6 que la période du glitch était liée à la profondeur acoustique de la perturbation τ_{II} , cette amplitude nous montre que son taux d'amortissement, $\frac{1}{\Delta_{\text{II}}\sqrt{2}}$, est inversement proportionnel à l'étendue de la perturbation. La dualité entre les domaines structurel et fréquentiel est alors très nette : une perturbation de structure locale (Δ_{II} faible) aboutira à un glitch étendu en fréquence, là où une perturbation étendue mènera à un glitch s'atténuant très vite. Une illustration de l'impact des paramètres Δ_{II} et τ_{II} sur la forme du glitch est ainsi disponible en FIGURE 3.5. À côté de cela, le paramètre Γ_{II} est également en mesure de nous renseigner sur la structure interne. Puisque :

$$\frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} \frac{\delta_x \Gamma_1}{\Gamma_1}(\tau) d\tau = \Gamma_{\text{II}} \mathcal{T}^{-1}, \quad (3.4.21)$$

9. Notons que le facteur ω produit une phase ascendante dans l'amplitude à basses fréquences, là où $e^{-2\Delta_{\text{II}}^2 \omega^2} \sim 1$. Cependant, celle-ci est à peine visible dans la gamme de fréquences observables (cf. FIGURE 3.5).

on voit en effet que l'amplitude du glitch est directement proportionnelle à la valeur moyenne de la perturbation de structure sur l'ensemble de l'étoile.

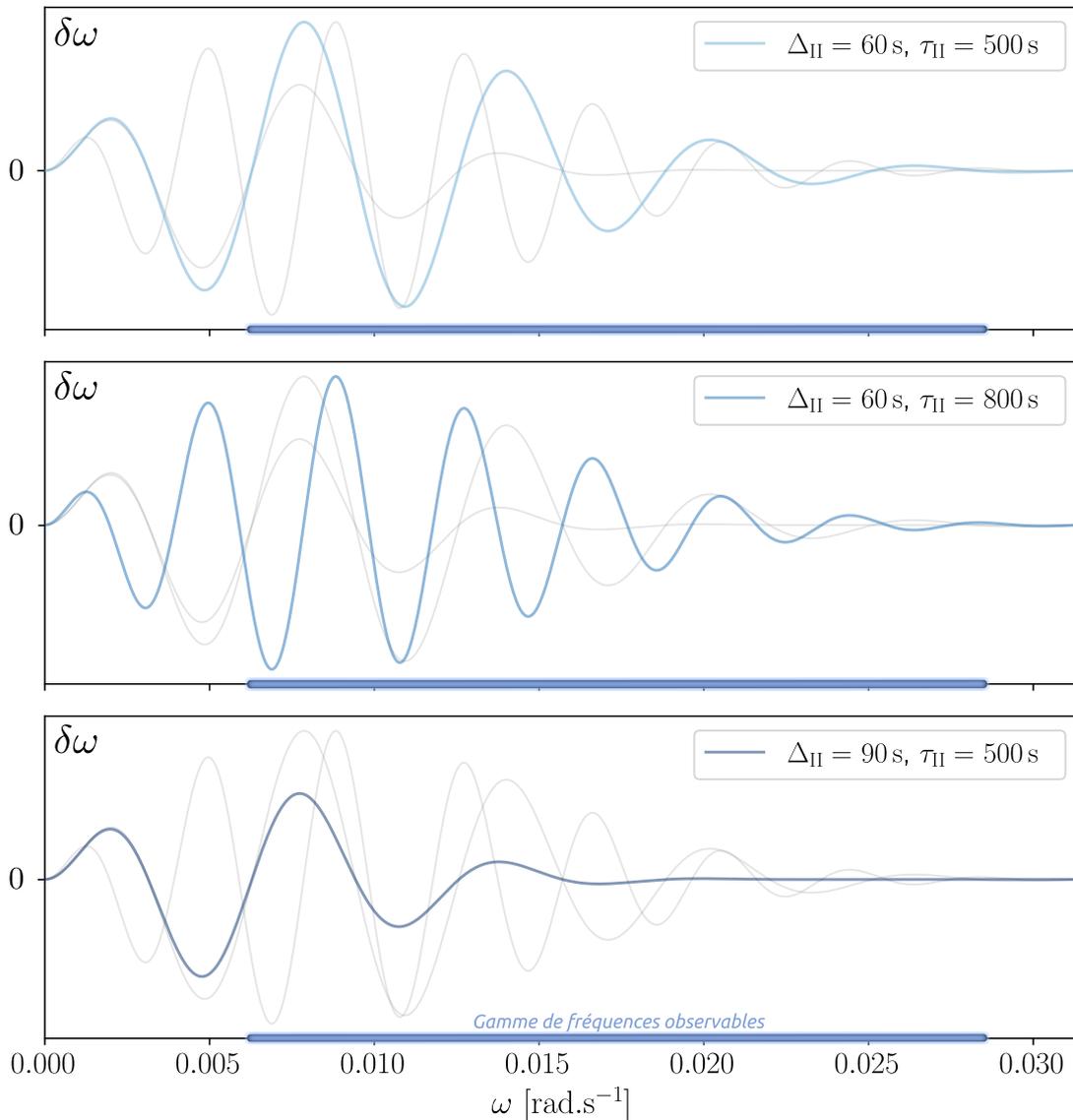


FIGURE 3.5 – Illustration de l'impact des deux paramètres Δ_{II} et τ_{II} sur la forme du glitch $\delta\omega$ donnée par les équations (3.4.19) & (3.4.20). L'amplitude Γ_{II} est la même dans les trois cas considérés, d'où l'échelle normalisée pour l'axe vertical. La gamme de fréquence typiquement observable pour un pulsateur de type solaire est indiquée sur chaque panneau par la barre horizontale et permet notamment d'estimer le nombre d'oscillations visibles. Les courbes estompées permettent une comparaison entre les différentes formes de glitch d'un panneau à l'autre.

3.5 Modèle de référence et diagnostic sismique

L'expression du glitch obtenue, il reste à trouver comment l'exploiter en pratique. Par définition, l'équation (3.4.19) fournit une approximation de la différence de fréquences entre celles provenant d'observations et celles issues d'un modèle de référence ne possédant pas le puits de la seconde région d'ionisation. La prochaine étape semble donc être la conception de ce modèle

de référence et le calcul de ses fréquences propres afin d'en déterminer la différence avec les observations $\delta\omega = \omega_{\text{obs}} - \omega_{\text{ref}}$. Cette différence peut ensuite être ajustée par l'équation (3.4.19) afin d'estimer les paramètres Γ_{II} , τ_{II} et Δ_{II} , lesquels nous renseignent sur la structure de l'étoile observée dans la région d'ionisation.

Néanmoins, la plupart des travaux récents (GOUGH, 2002; BASU et al., 2004; HOUDEK et GOUGH, 2007; VERMA et al., 2014b; VERMA et al., 2014a; MAZUMDAR et al., 2014; BROOMHALL et al., 2014; VERMA et al., 2017; VERMA et al., 2019; VERMA et SILVA AGUIRRE, 2019; VERMA et al., 2022) utilisent une astuce pour éviter de faire appel à un modèle de référence et définir le glitch sismique comme un propriété intrinsèque à la seule étoile observée. L'argument est le suivant. Puisque les fréquences de référence sont issues d'un modèle duquel la perturbation de structure a été retirée, celles-ci ne devraient pas présenter la composante oscillatoire caractéristique du glitch. Ses fréquences pourront donc être qualifiées de "lisses", dans le sens où elles varient sur des échelles bien plus grandes que le glitch lui-même :

$$\omega_{\text{obs}} = \delta\omega + \omega_{\text{lisse}}. \quad (3.5.1)$$

Cet argument peut être rendu d'autant plus fort en considérant les *secondes différences*, qui s'apparentent à une "dérivée" seconde vis-à-vis de l'ordre radial n :

$$d^2\omega(\omega_n) = \omega_{n+1} - 2\omega_n + \omega_{n-1}. \quad (3.5.2)$$

De manière générale, on réfère par *diagnostic sismique* au fait d'exploiter une fonction $f((\omega_n)_n)$ des fréquences plutôt que les fréquences ω_n directement. En l'occurrence, l'intérêt des secondes différences comme diagnostics sismiques se déduit de leurs proximité avec la dérivée seconde : les composantes possédant des fluctuations rapides se retrouvent amplifiées par rapport aux composantes plus lisses. il vient alors :

$$d^2\omega_{\text{obs}} = d^2\delta\omega + d^2\omega_{\text{lisse}}. \quad (3.5.3)$$

où la composante $d^2\delta\omega$ se retrouve amplifiée vis-à-vis de $d^2\omega_{\text{lisse}}$. La composante du glitch, $d^2\delta\omega$, peut directement être déduite de l'équation (3.4.19). HOUDEK et GOUGH, 2007 montrent que celle-ci peut être approximée par :

$$d^2\delta\omega \simeq 2\sqrt{\mathcal{D}_1(\omega)^2 + \mathcal{D}_2(\omega)^2} \mathcal{G}(\omega) \cos \left[2\omega\tau_{\text{II}} + \psi - \text{atan} \left(\frac{\mathcal{D}_1(\omega)}{\mathcal{D}_2(\omega)} \right) \right] \quad (3.5.4)$$

où les grandeurs \mathcal{D}_1 et \mathcal{D}_2 sont deux fonctions correctives¹⁰, variant relativement lentement avec la fréquence ω et dépendant des caractéristiques de la perturbation de structure :

$$\mathcal{D}_1(\omega; \tau_{\text{II}}, \Delta_{\text{II}}) = \frac{\pi(1 - 4\Delta_{\text{II}}^2\omega^2)}{\omega\mathcal{T}} \sin \left(\frac{2\pi\tau_{\text{II}}}{\mathcal{T}} \right), \quad (3.5.5)$$

$$\mathcal{D}_2(\omega; \tau_{\text{II}}, \Delta_{\text{II}}) = 1 - \left[1 - \frac{2\pi^2\Delta_{\text{II}}^2(3 - 4\Delta_{\text{II}}^2\omega^2)}{\mathcal{T}^2} \right] \cos \left(\frac{2\pi\tau_{\text{II}}}{\mathcal{T}} \right). \quad (3.5.6)$$

Il ne reste alors qu'à proposer un modèle pour la composante $d^2\omega_{\text{lisse}}(\omega)$. Divers modélisations ont été proposées, les plus courantes étant polynomiales (de puissances positives par praticité (VERMA et al., 2022) ou négatives en partant d'un développement asymptotiques des fréquences propres à des ordres élevés (HOUDEK et GOUGH, 2011; HOUDAYER et al., 2021)) :

$$d^2\omega_{\text{lisse}}(\omega; a_{i \leq 3}) \simeq \sum_{i=0}^3 a_i \omega^{\pm i}. \quad (3.5.7)$$

10. Les équations (3.5.4), (3.5.5) et (3.5.6) reprennent les équations (C.5), (C.6), (C.7), (C.8) et (C.9) de HOUDEK et GOUGH, 2007. Les notations sont quelque peu modifiées pour exploiter les symétries de ces relations.

Finalement, cette approche permet d'estimer les paramètres du glitch (en plus des paramètres supplémentaires introduits dans le développement) directement à partir des secondes différences des fréquences observées et en y ajustant la fonction qui résulte de la somme des expressions (3.5.4) et (3.5.7).

3.6 Limites et discussion

Moyennant toutes les remarques effectuées dans le paragraphe précédent, on voit que l'ajustement des paramètres du glitch (ici Γ_{II} , τ_{II} et Δ_{II}) permet de recueillir de précieuses informations sur la perturbation de structure à son origine. Néanmoins, on soulignera ici quelques limites à l'approche exposée dans les paragraphes précédents, lesquelles nous permettront de définir des objectifs afin d'améliorer l'exploitation du glitch de l'ionisation. Comme toujours lorsqu'il est question du glitch sismique, ces limites peuvent être classées en deux catégories : celles touchant à la modélisation dans le domaine fréquentiel, et celles tenant de la modélisation dans le domaine structurel. On commencera ainsi par mentionner succinctement les problèmes propres à l'approximation donnée plus haut, avant d'aborder les difficultés plus profondes liées à la modélisation de la perturbation de structure.

L'exploitation du glitch de l'ionisation par l'approche décrite dans les paragraphes 3.4 et 3.5 a été formellement introduite dans HOUDEK et GOUGH, 2007. À cette occasion elle a été testée dans un cadre idéal : les fréquences utilisées étaient celles calculées numériquement et l'incertitude supposée nulle. Connaissant les paramètres Γ_{II} , τ_{II} et Δ_{II} les plus à même de décrire le puits HeII dans un modèle, il était question de voir à quel point l'ajustement des expressions (3.5.4) et (3.5.7) permettait de les retrouver. Cet ajustement effectué, on peut y lire au sujet de la récupération des paramètres :

The values of the amplitudes of the depression are overestimated by some 50 per cent, and the widths underestimated by a similar amount, yielding an approximately correct magnitude of the glitch integral, Γ_{II} . [...] The variation of the acoustic depth τ_{II} of the centre of ionization is more-or-less reproduced. However, the actual magnitude appears to be discrepant; that is due in part to our crude representation of acoustic phase, and perhaps in part to an uncertainty in the origin of τ .

Ces biais de méthode, estimés dans un cadre idéal, laissent perplexe quant à l'application de la procédure à de véritables données sans précautions supplémentaires. Il est vraisemblable que l'origine de ces différences soit à attribuer en partie aux approximations effectuées dans le paragraphe 3.4. On y a notamment négligé la plupart des termes du kernel \mathcal{K}_ρ et seulement utilisé l'approximation JWKB à l'ordre 0. Pour arriver à une expression explicite du glitch, il a également fallu se placer dans la limite non pas des hautes fréquences mais des très hautes fréquences. En effet, l'approximation effectuée sort du cadre de la limite asymptotique telle que définie en annexe B.3 (c'est-à-dire par $\mathcal{L} \rightarrow 0$). Il a été nécessaire de négliger la longueur d'onde de l'oscillation devant l'échelle typique de variation de toutes les quantités structurelles de l'étoile. Cette hypothèse n'est cependant pas vérifiée dans la région de l'ionisation (c'est pour cela que l'on s'est penché sur l'exploitation du glitch sismique en premier lieu) ni dans les couches les plus superficielles de l'étoile, où de nombreuses quantités (p , ρ , T , ...) tendent rapidement vers 0. Ces difficultés étaient connues de HOUDEK et GOUGH, 2007, qui ont proposé plusieurs corrections pour y remédier. On y trouve l'ajout d'une phase pour prendre en compte la divergence de la fréquence de coupure en surface, ou l'emploi de fonctions d'Airy (GOUGH, 2007; HOUDEK et GOUGH, 2011) dans l'approximation JWKB pour approcher la divergence Λ dans l'ensemble de l'étoile. Ces ajouts sont néanmoins complexes et, à ma connaissance, peu utilisés en pratique. En réalité, les fonctions correctrices apparaissant dans l'équation (3.5.4) ne sont déjà plus d'usage et l'ajustement des secondes différences se fait généralement à l'aide de l'expression (3.4.19), bien qu'elle ait été dérivée pour approximer les fréquences.

Par delà ces considérations, qui tiennent plus de la façon de mener le développement que de la méthode elle-même, on notera un certain nombre de points, plus problématiques, à la base de la procédure présentée. À cet égard, il est utile de revenir à l'objectif affichée en début de chapitre, à savoir la détermination de l'abondance en hélium Y . En l'état, Y ne fait pas parti des paramètres du glitch et leur obtention ne permet donc pas d'estimer directement cette abondance. Cependant, il est clair que cette dernière n'est pas complètement indépendante de l'amplitude du puits HeII (voir schéma 3.3), et qu'il doit donc exister un lien entre Y et l'amplitude du glitch $A_{\text{II}} \equiv \alpha \Gamma_{\text{II}} \mathcal{T}^{-1}$ ¹¹ (cf. équation (3.4.20)). Cette relation $A_{\text{II}} = f(Y)$ est déterminée par calibration sur des modèles réalistes. Il s'agit simplement d'observer comment varie l'amplitude du puits, $-\Gamma_{\text{II}}/(\Delta_{\text{II}}\sqrt{2\pi})$ (voir la FIGURE 3.4 et l'équation 3.4.15), dans ces modèles en faisant varier l'abondance d'hélium Y . La loi $A_{\text{II}} = f(Y)$ peut alors être ajustée empiriquement. Derrière cette apparente simplicité se cache cependant un certain nombre d'hypothèses, et il peut être utile de formaliser la procédure afin de les faire apparaître. Notons \mathcal{M} le modèle réaliste utilisé pour la calibration et $\Gamma_1^{\mathcal{M}}(\tau)$ le profil de son premier exposant adiabatique vis-à-vis de la variable acoustique τ . Ce modèle, à un instant donné, possède une certaine abondance d'hélium Y , en plus d'être caractérisé par un grand nombre de paramètres comme sa masse M , son rayon R , sa température de surface T_{eff} , etc... Notons ces paramètres θ . Puisque l'amplitude du puits est nulle en l'absence d'hélium :

$$-\frac{\Gamma_{\text{II}}}{\Delta_{\text{II}}\sqrt{2\pi}}(Y) = \Gamma_1^{\mathcal{M}}(\tau_{\text{II}}; Y, \theta) - \Gamma_1^{\mathcal{M}}(\tau_{\text{II}}; 0, \theta) \simeq \left. \frac{\partial \Gamma_1^{\mathcal{M}}}{\partial Y} \right|_{\tau_{\text{II}}, Y=0, \theta} Y \quad (3.6.1)$$

On obtient alors une simple relation linéaire¹² entre l'amplitude du glitch et l'abondance en hélium : $A_{\text{II}}(Y) = aY$ avec la constante de proportionnalité :

$$a \equiv -\frac{\alpha \Delta_{\text{II}} \sqrt{2\pi}}{\mathcal{T}} \left. \frac{\partial \Gamma_1^{\mathcal{M}}}{\partial Y} \right|_{\tau_{\text{II}}, Y=0, \theta} . \quad (3.6.2)$$

Outre le fait qu'il a été nécessaire de supposer que la linéarité de $\Gamma_1^{\mathcal{M}}(\tau_{\text{II}}; Y, \theta)$ sur l'intervalle $[0, Y]$, ce qui est assez bien vérifié, une nouvelle difficulté apparaît : la constante de proportionnalité dépend explicitement du choix de θ . En pratique, il a vite été remarqué que cette loi de calibration dépendait en effet grandement de la masse de l'étoile, M (VERMA et al., 2014a). La solution apportée a alors été de définir une loi de calibration par étoile, en ayant préalablement obtenu une estimation suffisamment précise de sa masse M^* (VERMA et al., 2019). La constante de proportionnalité peut alors s'écrire :

$$a \equiv -\frac{\alpha \Delta_{\text{II}} \sqrt{2\pi}}{\mathcal{T}} \left. \frac{\partial \Gamma_1^{\mathcal{M}}}{\partial Y} \right|_{\tau_{\text{II}}, Y=0, M=M^*, \theta} . \quad (3.6.3)$$

On touche là à une limite profonde de l'approche par calibration. Il est nécessaire d'obtenir une estimation précise de la masse pour pouvoir déterminer l'abondance en hélium. Or, à cause de la dégénérescence masse-hélium (cf. paragraphe 1.3), c'est précisément pour obtenir une estimation plus précise de la masse que l'on cherche à déterminer l'abondance en hélium. Le problème, soulevé ici pour la masse, peut être transposé aux autres composantes de θ et même à la physique interne au modèle puisqu'il est nécessaire de faire un choix vis-à-vis du modèle réaliste \mathcal{M} .

11. Il existe plusieurs prescriptions pour définir un paramètre caractérisant l'amplitude du glitch. Si celui donné ici est le plus trivial, on trouve parfois comme définition $A_{\text{II}} \equiv \langle \mathcal{G} \rangle_{\omega_0}$, la valeur moyenne de $\mathcal{G}(\omega)$ autour d'une certaine fréquence ω_0 à définir (VERMA et al., 2014b; VERMA et al., 2019). Cela n'impacte aucunement la discussion qui suit.

12. Notons que la relation ajustée en pratique est plutôt affine (VERMA et al., 2019). Cela pose le problème de l'interprétation physique de l'ordonnée à l'origine (non nulle), qui serait alors l'amplitude du glitch causée par l'ionisation de l'hélium en l'absence d'hélium.

Concernant la modélisation de la perturbation de structure, une hypothèse forte a également été de supposer que cette dernière portait uniquement sur le puits HeII. En réalité, une variation de l'abondance en hélium δY impactera nécessairement le puits HeI et même la région d'ionisation de l'hydrogène. Une généralisation de l'expression (3.5.4) aux deux puits de l'hélium :

$$d^2\delta\omega \simeq \sum_{i \in \{I, II\}} 2\sqrt{\mathcal{D}_1^i(\omega)^2 + \mathcal{D}_2^i(\omega)^2} \mathcal{G}_i(\omega) \cos \left[2\omega\tau_i + \psi - \text{atan} \left(\frac{\mathcal{D}_1^i(\omega)}{\mathcal{D}_2^i(\omega)} \right) \right] \quad (3.6.4)$$

a été proposée par **HOUDEK et GOUGH, 2007**, mais est encore une fois peu utilisée en pratique car introduisant de nouveaux paramètres (Γ_I , τ_I et Δ_I) ou nécessitant l'hypothèse de ratio fixes pour Γ_I/Γ_{II} , τ_I/τ_{II} et Δ_I/Δ_{II} . À propos de la paramétrisation, on peut remarquer l'ajout de nombreux paramètres, $a_{i \leq 3}$ causé par l'apparition de la composante lisse (cf. (3.5.7)), lesquels ne possèdent pas de véritables significations physiques. Ainsi, il est rare de voir l'emploi de moins de 8 paramètres libres pour l'ajustement de la seule composante HeII du glitch (et 11 si la composante de la BCZ est incluse), ce qui semble important pour l'exploitation de seulement 3 (respectivement 4) de ces paramètres. L'information issue des données tend alors à se répartir entre ces nombreux paramètres, ce qui implique de plus faibles contraintes pour les quantités d'intérêt.

Plusieurs remarques sont également à faire sur la façon de modéliser la perturbation de structure elle-même. Je me suis longuement attardé, dans le paragraphe 2.5.3, sur l'importance du choix de la variable à fixer pour définir la variation de structure entre deux modèles. Ici, que ce soit au niveau des inspections visuelles du profil du premier exposant adiabatique Γ_1 (cf. FIGURES 3.2 & 3.3) ou la définition de la loi de calibration (3.6.1), ce choix s'est porté sur la variable τ . Outre le fait qu'il ne s'agit pas d'une variable normalisée, le problème principal est que les kernels à approximer proviennent eux d'une perturbation de structure à $x = r/R$ fixé. Comme illustré en FIGURE 2.8, ce choix est loin d'être anodin et suffirait à expliquer des différences de 50% dans la mesure de l'amplitude du puits HeII à lui seul.

Une approximation majeure et non évoquée jusqu'ici réside dans la suppression de la perturbation de structure $\delta_x p/p$ devant $\delta_x \Gamma_1/\Gamma_1$ dans l'équation (3.4.3). On peut chercher à comprendre ce que cela implique implicitement vis-à-vis du modèle de référence. Supposons que les quantités p et Γ_1 désignent les profils de pression et du premier exposant adiabatique dans une étoile observée. Les profils du modèle de référence sont alors donnés par $p^{\text{ref}} = p - \delta_x p$ et $\Gamma_1^{\text{ref}} = \Gamma_1 - \delta_x \Gamma_1$, par définition. Si la perturbation de la pression p est négligée devant celle de Γ_1 , cette structure de référence possède les profils p et $\Gamma_1 - \delta_x \Gamma_1$. Or la pression n'est pas indépendante de Γ_1 comme montrée par l'équation (3.4.4) et, puisque cette relation se déduit seulement de l'équilibre hydrostatique et de l'équation de Poisson, on en conclut que le modèle de référence (ou alors l'étoile observée) brise une de ces lois. Ce point est particulièrement problématique. Si j'ai pu dire, dans le paragraphe 3.1, que le modèle de référence n'avait pas besoin d'être plausible pour être utile, une prérogative reste néanmoins sa cohérence vis-à-vis des équations de l'hydrostatique sur lesquelles reposent l'ensemble des analyses des sections 2.4 et 2.5. Typiquement, dans le cas où le modèle de référence est supposé être lisse (cf. paragraphe 3.5) et où donc ce modèle ne possède pas de région HeII (si ce n'est pas de région d'ionisation tout court), la perturbation relative de pression doit en réalité être très prononcée¹³. On peut alors également se demander si le régime de perturbation linéaire persiste dans ce cas. En effet, si le choix du modèle lisse se porte sur un modèle polytropique où $\Gamma_1 = \text{cste} = 5/3$ par exemple, les changements de structures entre une étoile observée et ce modèle de référence risque d'aboutir à des $\delta_x \Gamma_1/\Gamma_1 \sim 1$, auquel cas il serait nécessaire d'impliquer des perturbations d'ordre supérieur (en $(\delta_x \Gamma_1/\Gamma_1)^2$ par exemple) dans le développement. À l'inverse, le choix d'une référence plus proche, comme celui d'un modèle réaliste où $Y = 0$, aboutit à des amplitudes de perturbation plus faibles mais il est difficile de dire si les fréquences de référence peuvent être réellement considérées comme lisses¹⁴. Une illustration

13. Notamment parce que celle-ci varie de plus d'un facteur 2 au cours de l'ionisation, cf. paragraphe 2.3.7.

14. À cause de la région d'ionisation de l'hydrogène notamment, prépondérante en FIGURE 3.2 et qui constitue elle aussi une variation de structure abrupte dans la région d'ionisation.

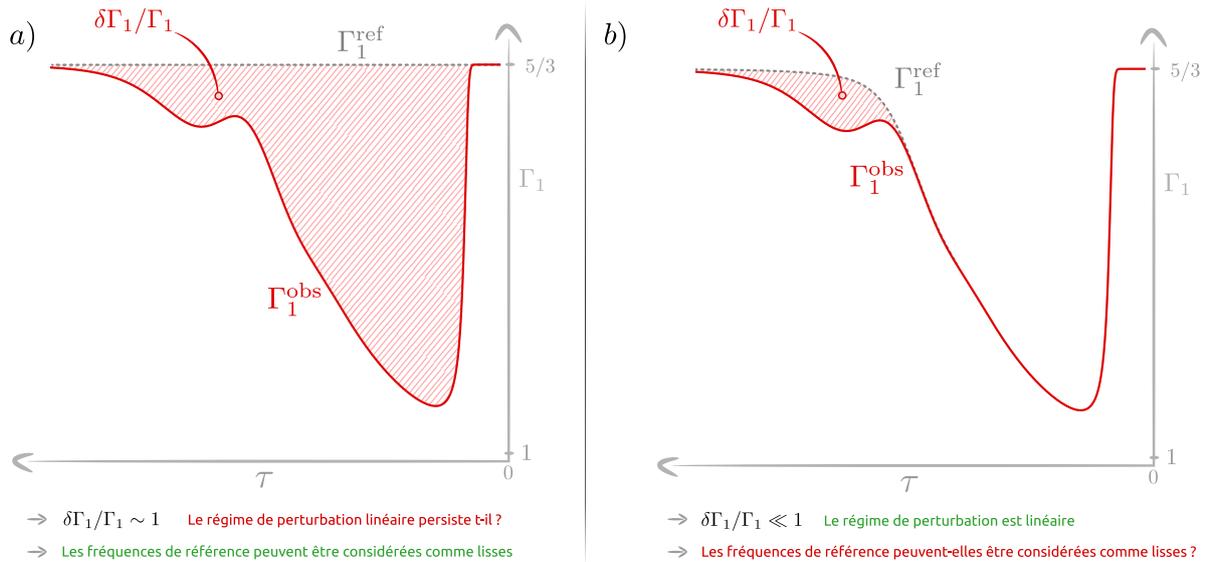


FIGURE 3.6 – Illustration du compromis à trouver pour le choix d’un modèle de référence “lisse”. **a)** Le choix de la référence s’est porté sur un modèle polytropique en l’absence d’ionisation et $\Gamma_1^{\text{ref}} = \text{cste} = 5/3$. Cette référence peut alors être considérée comme lisse mais l’écart structurel avec une véritable étoile (d’exposant adiabatique Γ_1^{obs}) se fait important, aboutissant à des $\delta\Gamma_1/\Gamma_1 \sim 1$. **b)** Le choix de la référence s’est porté sur un modèle plus proche (structurellement) de l’étoile observée. Celui-ci possède des zones d’ionisation H et HeI mais la zone HeII lui a été retirée. Si l’on a bien $\delta\Gamma_1/\Gamma_1 \ll 1$ cette fois-ci, il est revanche difficile de considérer les fréquences d’oscillation de référence comme lisses puisqu’elles contiennent les signatures des zones d’ionisation H et HeI.

du compromis à trouver est donnée en FIGURE 3.6.

Les points qui ont été soulevés ici sur les méthodes d’exploitation actuelles du glitch sismique peuvent être retrouvés dans la TABLE 3.1. Ces thématiques tiennent de la modélisation et j’aimerais clore ce passage par une remarque sortant légèrement de ce cadre. Comme observé plus haut, les risques de biais (tout comme les facteurs d’incertitude) lors de l’estimation de l’abondance en hélium par les méthodes actuelles sont multiples. En l’état, une attention particulière doit être réservée à la quantification de ces biais et incertitudes ou, lorsque celle-ci ne peut se faire, à une certaine précaution vis-à-vis des résultats. Ces estimations de l’abondance en hélium peuvent en effet avoir de très larges retombées, non seulement en astérosismologie mais également en modélisation stellaire ou dans l’étude des populations stellaires. Les valeurs de référence obtenues à l’aide des méthodes actuelles sur un échantillon d’étoiles très bien observées¹⁵ ont été obtenues par VERMA et al., 2019. Nombre d’entre elles font état d’un intervalle d’incertitude $< 3\%$ sur Y , lequel résulte directement de la propagation des incertitudes observationnelles sur les paramètres de l’ajustement. Au regard de la discussion menée dans ce paragraphe, de nombreuses sources d’imprécision potentielles (en particulier celles recensées dans la TABLE 3.1) sont dès lors écartées et ces quantifications de l’incertitude paraissent donc grandement sous-évaluées.

Chacun des aspects évoqués plus haut est à même de causer des biais dans l’estimation de l’abondance en hélium Y et leur nombre souligne la complexité sous-jacente à l’exploitation du glitch de l’ionisation. Certains aspects du problème ressortent du tableau dressé plus haut et

15. À savoir, un sous-échantillon du LEGACY sample de Kepler (GILLILAND et al., 2010; LUND et al., 2017).

TABLE 3.1 – Quelques hypothèses sur lesquelles reposent l’exploitation du glitch d’ionisation, discutées dans le texte.

	Hypothèse / Limite	Domaine de modélisation concerné
1	Développement asymptotique à l’ordre 0	Fréquentiel
2	Variations suffisamment lentes dans la structure interne	Fréquentiel
3	Relation $A_{II} = f(Y)$ établie par calibration	Structurel (paramétrisation)
4	Impact des autres régions d’ionisation (H, HeI) négligées	Structurel (paramétrisation)
5	Nombre de paramètres non physiques	Structurel & Fréquentiel (paramétrisation)
6	Choix de variable à fixer différent entre perturbation de structure et kernel	Structurel
7	Perturbation de la pression $\delta_x p / p$ négligée devant $\delta_x \Gamma_1 / \Gamma_1$	Structurel
8	Modèle de référence non établi	Structurel
9	Persistance du régime de perturbation linéaire	Structurel

constituent autant d’opportunités d’améliorer la méthode présentée. Sur le plan de la modélisation structurelle, on pourra ainsi mettre en avant plusieurs objectifs. Un point essentiel me semble résider dans l’obtention d’une expression faisant apparaître explicitement l’abondance en hélium Y pour éviter l’usage de lois de calibration. Notons, qu’il sera dès lors nécessaire de présenter une description physique de la structure interne et non plus visuelle. Celle-ci devrait, en outre, permettre une modélisation cohérente de la perturbation de structure et réduire les ambiguïtés autour du modèle de référence. Concernant le domaine fréquentiel, un aspect important me semble être l’obtention d’une approximation du glitch valable à toutes fréquences et non plus dans la limite où les variations de la structure se font lentes, incompatible avec l’étude des régions d’ionisation et de surface. Enfin, une perspective utile (et transversale aux deux domaines) réside dans la réduction du nombre de paramètres et, en particulier, des paramètres non physiques.

Répondre aux objectifs sus-cités est une tâche complexe, à laquelle on s’est essayé dans les articles [HOUDAYER et al., 2021](#) et [HOUDAYER, REESE et GOUPIL, 2022](#), disponibles en sections 4.1 et 5.1 et brièvement discutés dans les chapitres les précédant.

3.7 Limits and discussion [en]

Based on the remarks made in the previous paragraph, we can see that fitting the glitch parameters (here Γ_{II} , τ_{II} and Δ_{II}) reveals precious information on the structural perturbation at its origin. Nevertheless, we shall underline here some limits to the approach described in the previous paragraphs, which will lead us to define objectives in order to improve the exploitation of ionisation glitches. As always when dealing with seismic glitches, these limitations can be grouped into two categories: those related to the modelling in the frequency domain, and those related to the modelling in the structural domain. We will therefore start by briefly mentioning problems specific to the approximation given above, before discussing the more fundamental difficulties associated with the modelling of the structural perturbation.

The exploitation of the ionisation glitch as described in paragraphs 3.4 and 3.5 was formally introduced in [HOUDEK et GOUGH, 2007](#). On that occasion, it was tested in an ideal framework: the oscillation frequencies used as data were calculated numerically and the uncertainties assumed to be zero. Knowing the parameters Γ_{II} , τ_{II} and Δ_{II} which best reproduced the HeII well in a realistic model, the aim was to determine to what extent fitting the expressions (3.5.4) and (3.5.7) would enable one to recover them. Once the adjustment was made, we can read about the recovery of the parameters :

The values of the amplitudes of the depression are overestimated by some 50 per cent, and the widths underestimated by a similar amount, yielding an approximately correct magnitude of the glitch integral, Γ_{II} . [...] The variation of the acoustic depth τ_{II} of the centre of ionisation is more-or-less reproduced. However, the actual magnitude appears to be discrepant; that is due in part to our crude representation of acoustic phase, and perhaps in part to an uncertainty in the origin of τ .

These biases in the method, estimated in an ideal framework, raise concerns regarding the application of the procedure to real data without additional precautions. It is likely that part of the differences can be attributed to the approximations made in paragraph 3.4. In particular, most of the terms of the \mathcal{K}_ρ kernel were neglected and the JWKB approximation used at the zeroth order. In order to derive an explicit glitch expression, it was also necessary to consider the limit not of high but very high frequencies. In fact, the approximation carried out goes beyond the asymptotic limit as defined in appendix B.3 (i.e. by $\mathcal{L} \rightarrow 0$). It was necessary to neglect the wavelength of the oscillation in front of the typical scale of variation of any structural quantities within the star. However, this assumption does not hold in the ionisation region (which is the reason why the exploitation of the seismic glitch is considered in the first place) nor in the most superficial layers of the star, where many quantities (p, ρ, T, \dots) rapidly approach 0. These difficulties were known to [HOUDEK et GOUGH, 2007](#), who proposed several corrections to address them. These include the addition of a phase to account for the divergence of the cut-off frequency at the surface, or the use of Airy functions in the JWKB approximation to approximate the divergence Λ in the whole star. These additions are nevertheless complex and, to my knowledge, little used in practice. In fact, even the correction functions in equation (3.5.4) are no longer used, and the fit of the second differences is usually done using expression (3.4.19), although it has been derived to approximate the frequencies.

Beyond these considerations, which have more to do with the way the development was carried out than with the method itself, a certain number of more problematic issues at the basis of the procedure presented can be pointed out. In this respect, it is helpful to recall the objective stated at the beginning of the chapter, namely the determination of the helium abundance Y . As it stands, Y is not part of the glitch parameters and therefore obtaining them does not allow us to estimate this abundance directly. However, it is clear that the latter is not entirely independent of the amplitude of the HeII well (see diagram 3.3), and that there must therefore be a link between Y and the glitch amplitude $A_{\text{II}} \equiv \alpha \Gamma_{\text{II}} \mathcal{T}^{-1}$ ¹⁶ (cf. equation (3.4.20)). This relation $A_{\text{II}} = f(Y)$ is established by calibration on realistic models. One merely has to observe how the well amplitude varies, $-\Gamma_{\text{II}}/(\Delta_{\text{II}}\sqrt{2\pi})$ (see the [FIGURE 3.4](#) and the equation 3.4.15), in these models by varying the helium abundance Y . The law $A_{\text{II}} = f(Y)$ can then be adjusted empirically. Behind this apparent simplicity, however, lies a certain number of assumptions, and it may be useful to formalise the procedure in order to make them apparent. Let us denote \mathcal{M} the realistic model used for the calibration and $\Gamma_1^{\mathcal{M}}(\tau)$ its first adiabatic exponent profile with respect to the acoustic variable τ . This model, at a given time, has a certain abundance of helium Y , in addition to being characterised by a large number of parameters such as its mass M , radius R , surface temperature T_{eff} , etc... Let us denote these parameters θ . Since the amplitude of the well is zero in the absence of helium :

$$-\frac{\Gamma_{\text{II}}}{\Delta_{\text{II}}\sqrt{2\pi}}(Y) = \Gamma_1^{\mathcal{M}}(\tau_{\text{II}}; Y, \theta) - \Gamma_1^{\mathcal{M}}(\tau_{\text{II}}; 0, \theta) \simeq \left. \frac{\partial \Gamma_1^{\mathcal{M}}}{\partial Y} \right|_{\tau_{\text{II}}, Y=0, \theta} Y \quad (3.7.1)$$

16. There are several prescriptions for defining a parameter that describes the glitch amplitude. The one given here is the most trivial, but we sometimes find the definition $A_{\text{II}} \equiv \langle \mathcal{G} \rangle_{\omega_0}$, the average value of $\mathcal{G}(\omega)$ around some given frequency ω_0 ([VERMA et al., 2014b](#); [VERMA et al., 2019](#)). This does not affect the following discussion.

We then obtain a simple linear relation¹⁷ between the amplitude of the glitch and the helium abundance : $A_{\text{II}}(Y) = aY$ with the following proportionality constant:

$$a \equiv - \frac{\alpha \Delta_{\text{II}} \sqrt{2\pi}}{\mathcal{T}} \left. \frac{\partial \Gamma_1^{\mathcal{M}}}{\partial Y} \right|_{\tau_{\text{II}}, Y=0, \theta} . \quad (3.7.2)$$

Apart from the fact that it was necessary to suppose the linearity of $\Gamma_1^{\mathcal{M}}(\tau_{\text{II}}; Y, \theta)$ on the $[0, Y]$ interval, which is fairly well verified, a new difficulty arises : the proportionality constant depends explicitly on the choice of θ . In practice, it was quickly noticed that the calibration law was indeed highly dependent on the star's mass, M (VERMA et al., 2014a). The solution was then to define a calibration law for each star, having previously obtained a sufficiently precise estimate of its mass M^* (VERMA et al., 2019). The proportionality constant can then be written :

$$a \equiv - \frac{\alpha \Delta_{\text{II}} \sqrt{2\pi}}{\mathcal{T}} \left. \frac{\partial \Gamma_1^{\mathcal{M}}}{\partial Y} \right|_{\tau_{\text{II}}, Y=0, M=M^*, \theta} . \quad (3.7.3)$$

This raises a profound limitation of the calibration-based approach. It is necessary to first obtain a precise estimate of the mass in order to determine the helium abundance. However, because of mass-helium degeneracy (cf. paragraph 1.3), it is precisely in order to obtain a more precise estimate of the mass that we seek to determine the helium abundance. The problem, mentioned here for the mass, can be transposed to the other components of θ and even to the internal physics of the model since it is necessary to make a choice regarding the realistic model \mathcal{M} .

Regarding the modelling of the structural perturbation, a further strong assumption is that the perturbation only affects the HeII well. In practice, however, a variation in the helium abundance δY will necessarily impact the HeI well and even the hydrogen ionisation zone. A generalisation of the expression (3.5.4) to include both helium wells:

$$d^2\delta\omega \simeq \sum_{i \in \{\text{I,II}\}} 2\sqrt{\mathcal{D}_1^i(\omega)^2 + \mathcal{D}_2^i(\omega)^2} \mathcal{G}_i(\omega) \cos \left[2\omega\tau_i + \psi - \text{atan} \left(\frac{\mathcal{D}_1^i(\omega)}{\mathcal{D}_2^i(\omega)} \right) \right] \quad (3.7.4)$$

was proposed by HOUDEK et GOUGH, 2007, but is once again rarely used in practice because it introduces new parameters (Γ_I , τ_I and Δ_I) or requires the assumption of fixed ratios for $\Gamma_I/\Gamma_{\text{II}}$, τ_I/τ_{II} and $\Delta_I/\Delta_{\text{II}}$. Concerning the parametrisation, one may notice the addition of numerous parameters, $a_{i \leq 3}$ caused by the smooth component (cf. (3.5.7)), while they do not have concrete physical meanings. Hence, it is rare to see the use of less than 8 free parameters when fitting the HeII glitch component alone (and 11 if the BCZ component is included), which seems rather significant for the exploitation of only 3 (respectively 4) of these parameters. The information from the data then tends to be distributed among these many parameters, which implies weaker constraints on the quantities of interest.

Several comments should also be made about the modelling of the structural perturbation. I stressed, in the paragraph 2.5.3, the importance of choosing the variable to be fixed in order to define the variation of structure between two models. Here, whether it be when visually inspecting the first adiabatic exponent profile Γ_1 (cf. FIGURES 3.2 & 3.3) or defining the calibration law (3.7.1), this choice went to the τ variable. Besides the fact that it is not a normalised variable, the main problem is that the kernels that are subsequently approximated come from a perturbation of the structure at fixed $x = r/R$. As illustrated in FIGURE 2.8, this choice is far from benign and could alone explain 50% of the differences in the HeII well amplitude measurement.

17. We note that the fitted relation is affine rather than linear in practice (VERMA et al., 2019). This raises the problem of the physical interpretation of the (non-zero) intercept, which would then be the amplitude of the glitch caused by the ionisation of helium in the absence of helium.

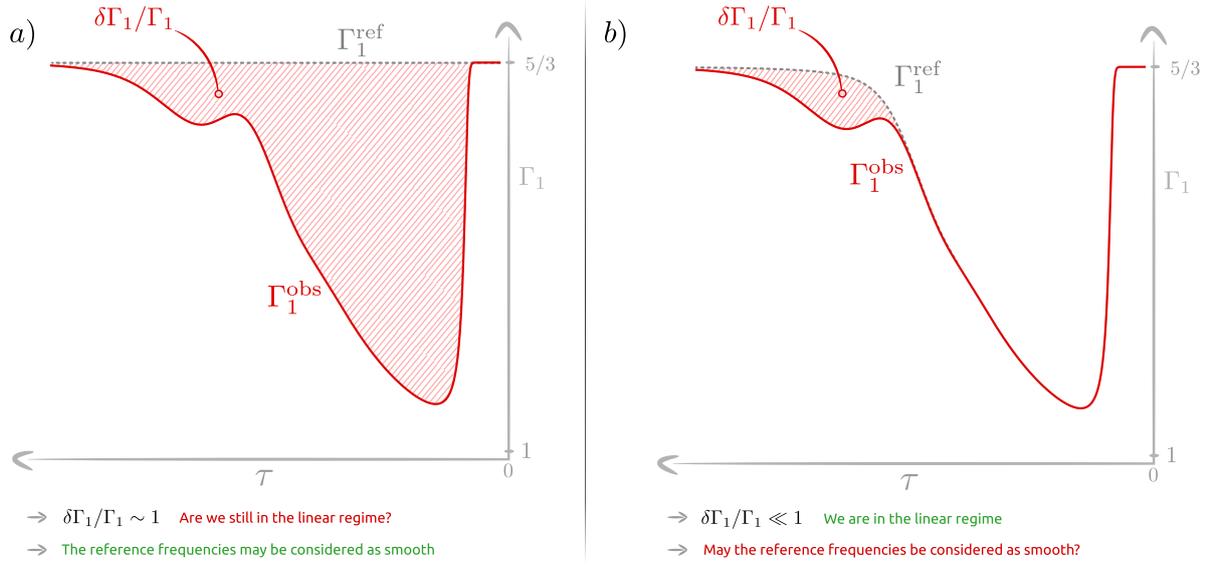


FIGURE 3.7 – Illustration of the trade-off for the choice of a "smooth" reference model. **a)** The reference was chosen to be a polytropic model in the absence of ionisation and $\Gamma_1^{\text{ref}} = \text{cste} = 5/3$. This reference can therefore be considered as smooth, but the structural difference with an actual star (of adiabatic exponent Γ_1^{obs}) becomes important, leading to $\delta\Gamma_1/\Gamma_1 \sim 1$. **b)** The reference was chosen to be a closer model (structurally) to the observed star. This model has both H and HeI ionisation zones but the HeII zone has been removed. Although we indeed have $\delta\Gamma_1/\Gamma_1 \ll 1$ this time, it is however difficult to consider the reference oscillation frequencies as smooth since they contain the signatures of both the H and HeI ionisation zones.

A major approximation not mentioned so far consists in neglecting the structural perturbation $\delta_x p/p$ compared to $\delta_x \Gamma_1/\Gamma_1$ in equation (3.4.3). We can try to understand what this implicitly implies with regard to the reference model. Let us assume that the quantities p and Γ_1 denote the pressure and the first adiabatic exponent profiles in an observed star. The profiles of the reference model are then given by $p^{\text{ref}} = p - \delta_x p$ and $\Gamma_1^{\text{ref}} = \Gamma_1 - \delta_x \Gamma_1$, by definition. If the pressure perturbation of p is neglected in front of that of Γ_1 , this reference structure then has the profiles p and $\Gamma_1 - \delta_x \Gamma_1$. However, the pressure is not independent of Γ_1 as shown by equation (3.4.4) and, since this relation can be deduced from the hydrostatic equilibrium and Poisson's equation alone, we conclude that the reference model (or the observed star) must break one of these laws. This point is highly problematic. Although I may have said, in the paragraph 3.1, that the reference model does not need to be plausible to be useful, a prerequisite remains its coherence with respect to the hydrostatic equations, on which all the analyses of the sections 2.4 and 2.5 are based. Typically, in the case where the reference model is assumed to be smooth (cf. paragraph 3.5) and where, as a result, it does not have a HeII region (or not any ionisation region for that matter), the relative pressure perturbation must in fact be very pronounced¹⁸. One can then also wonder if the linear perturbation regime persists in this case. Indeed, if one chooses a polytropic model where $\Gamma_1 = \text{cste} = 5/3$ as the smooth model for instance, the structural changes between an observed star and this reference model may lead to $\delta_x \Gamma_1/\Gamma_1 \sim 1$, in which case it would be necessary to involve higher order perturbations (varying as $(\delta_x \Gamma_1/\Gamma_1)^2$ for instance) in the derivation. Alternatively, the choice of a closer reference, such as a realistic model where $Y = 0$, leads to smaller perturbation amplitudes but it is difficult to say whether the reference frequencies can really be

18. Mostly because it varies by more than a factor of 2 during the ionisation, cf. paragraph 2.3.7.

considered as smooth¹⁹. An illustration of the trade-off to be addressed is given in FIGURE 3.7.

The points raised here about current methods for exploiting the seismic glitch are summarised in TABLE 3.2. These issues focus on modelling, and I would like to close this section with a slightly broader remark. As observed above, the risks of bias (as well as the factors leading to uncertainties) when estimating the helium abundance by current methods are diverse. As it stands, particular attention should be devoted to quantifying of these biases and uncertainties or, where this cannot be done, a certain degree of precaution should be taken regarding the results. These estimates of helium abundance can have very broad implications, not only for asteroseismology but also for stellar modelling and the study of stellar populations. Reference values obtained using current methods on a sample of very well observed stars²⁰ have been obtained by VERMA et al., 2019. Many of them show an uncertainty interval $< 3\%$ on Y , which is a direct consequence of the propagation of observational uncertainties on the fitting parameters. In the light of the above discussion, many potential sources of imprecision (in particular those identified in TABLE 3.2) are thus discarded and these uncertainty estimates appear to be greatly underestimated.

TABLE 3.2 – Some of the assumptions on which the exploitation of the ionisation glitch is based, as discussed in the text.

	Assumptions / Limits	Relevant modelling domain
1	0^{th} order asymptotic expansion	Frequential
2	Sufficiently slow variations in the internal structure	Frequential
3	$A_{\text{II}} = f(Y)$ relation established through calibration	Structural (parameterisation)
4	Impact of other ionisation regions (H, HeI) neglected	Structural (parameterisation)
5	Number of non-physical parameters	Structural & Frequential (parameterisation)
6	Inconsistent choice of the variable to fix between the structure perturbation and kernel	Structural
7	Pressure perturbation $\delta_x p/p$ neglected compared to $\delta_x \Gamma_1/\Gamma_1$	Structural
8	reference model non-established	Structural
9	Persistence of the linear perturbation regime	Structural

Each of the aspects mentioned above is likely to cause biases when estimating the helium abundance Y , and their number highlights the underlying complexity of exploiting the ionisation glitch. Some aspects of the problem stand out in the above table and thus provide opportunities to improve the method introduced above. In terms of structural modelling, several objectives can be outlined. A point that seems essential to me is obtaining an expression in which the helium abundance Y intervenes explicitly to prevent the use of calibration laws. We note that it will then be necessary to present a physical description of the internal structure and no longer a visual one. This should, moreover, allow a coherent modelling of the structural perturbation and reduce the ambiguities surrounding the reference model. Concerning the frequency domain, an aspect that seems important to me is the derivation of an approximation of the glitch valid at all frequencies and no longer in the limit of slow variations of the structure, which is incompatible with the study of ionisation and surface regions. Finally, a perspective beneficial in both domains lies in reducing the number of parameters and, in particular, the non-physical parameters.

Addressing the above-mentioned objectives is a complex task, which was attempted in the articles HOUDAYER et al., 2021 and HOUDAYER, REESE et GOUPIL, 2022, available in sections 4.1 and 5.1 and briefly discussed in the preceding chapters.

19. Mostly because of the hydrogen ionisation, which is dominant in FIGURE 3.2 and also causes an abrupt structural variation in the ionisation region

20. Namely a sub-sample of *Kepler's LEGACY sample*

Chapitre 4

Modéliser la région d'ionisation

Contenu du chapitre

- 4.1 Properties of the ionisation glitch. I. Modelling the ionisation region **116**
-

Properties of the ionisation glitch

I. Modelling the ionisation region

Pierre S. Houdayer¹, Daniel R. Reese¹, Marie-Jo Goupil¹, and Yveline Lebreton^{1,2}

¹ LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92195 Meudon, France
e-mail: pierre.houdayer@obspm.fr

² Univ Rennes, CNRS, IPR (Institut de Physique de Rennes) – UMR 6251, 35000 Rennes, France

Received 5 July 2021 / Accepted 30 September 2021

ABSTRACT

Context. Determining the properties of solar-like oscillating stars can be subject to many biases. A particularly important example is the helium-mass degeneracy, where the uncertainties regarding the internal physics can cause a poor determination of both the mass and surface helium content. Accordingly, an independent helium estimate is needed to overcome this degeneracy. A promising way to obtain such an estimate is to exploit the so-called ionisation glitch, that is, the deviation from the asymptotic oscillation frequency pattern caused by the rapid structural variation in the He ionisation zones.

Aims. Although it is progressively becoming more sophisticated, the glitch-based approach faces problems inherent to its current modelling such as the need for calibration using realistic stellar models. This requires a physical model of the ionisation region that explicitly involves the parameters of interest, such as the surface helium abundance, Y_s .

Methods. Through a thermodynamic treatment of the ionisation region, an analytical approximation for the first adiabatic exponent Γ_1 is presented.

Results. The induced stellar structure is found to depend on only three parameters, including the surface helium abundance Y_s and the electron degeneracy ψ_{CZ} in the convective region. The model thus defined allows a wide variety of structures to be described, and it is in particular able to approximate a realistic model in the ionisation region. The modelling work we conducted enables us to study the structural perturbations causing the glitch. More elaborate forms of perturbations than those that are usually assumed are found. It is also suggested that there might be a stronger dependence of the structure on the electron degeneracy in the convection zone and on the position of the ionisation region rather than on the amount of helium itself.

Conclusions. When analysing the ionisation glitch signature, we emphasise the importance of having a relation that can take these additional dependences into account.

Key words. asteroseismology – stars: abundances – stars: fundamental parameters – stars: interiors – stars: solar-type – stars: oscillations

1. Introduction

Asteroseismology, which is the study of resonant modes in stars, reveals information on the physical properties of the layers that a wave passes through on its way to the surface. Coupled with a physical model of the star, it thus allows us to constrain the various internal processes involved better than any other method. Undoubtedly, the constraints thus obtained depend on what is chosen to be modelled and what is assumed to be known. This choice is highly complex and largely dependent on the quality of the information provided (i.e., the available precision on the observables). Although high-precision photometric data provided by CoRoT¹ (Baglin et al. 2006), *Kepler* (Gilliland et al. 2010; Lund et al. 2017), and now TESS² (Ricker et al. 2015; Stassun et al. 2019) allow ever more precise estimates of oscillation frequencies, the evaluation of certain stellar parameters remains uncertain, however. This illustrates that accuracy rather than precision becomes the limiting factor in this case. Although the value of seismic inference in the determination of stellar parameters is undeniable, it also opens the door to potential

biases on these very parameters (some of them are illustrated in Fig. 1).

In this respect, the determination of abundances using realistic stellar models constitutes a textbook case that is potentially susceptible to a wide range of biases. Identified sources of difficulties undoubtedly reside in the complexity of the physics that must be considered. In particular, the equation of state, opacities (Kosovichev et al. 1992), or transport processes such as diffusion, turbulent mixing (Christensen-Dalsgaard et al. 1993), or radiative acceleration (Deal et al. 2018) can be evoked, whose consideration or exclusion may result in physical biases. Recent studies using model grids also suggest a strong anti-correlation between the mass and initial helium abundance estimates (Lebreton & Goupil 2014; Noll et al. 2021). This degeneracy makes the inference more complex by resulting in a high volatility of both parameter estimates. Moreover, when a large number of frequencies are available, as for the Sun, inversion techniques highlight significant discrepancies, and solar models are forced to choose between inconsistent abundances, densities, or convective zone (CZ) depths (Basu & Antia 2004; Asplund et al. 2009; Serenelli et al. 2009). In addition to these processes that directly involve the composition, additional

¹ Convection, Rotation et Transits planétaires.

² Transiting Exoplanet Survey Satellite.

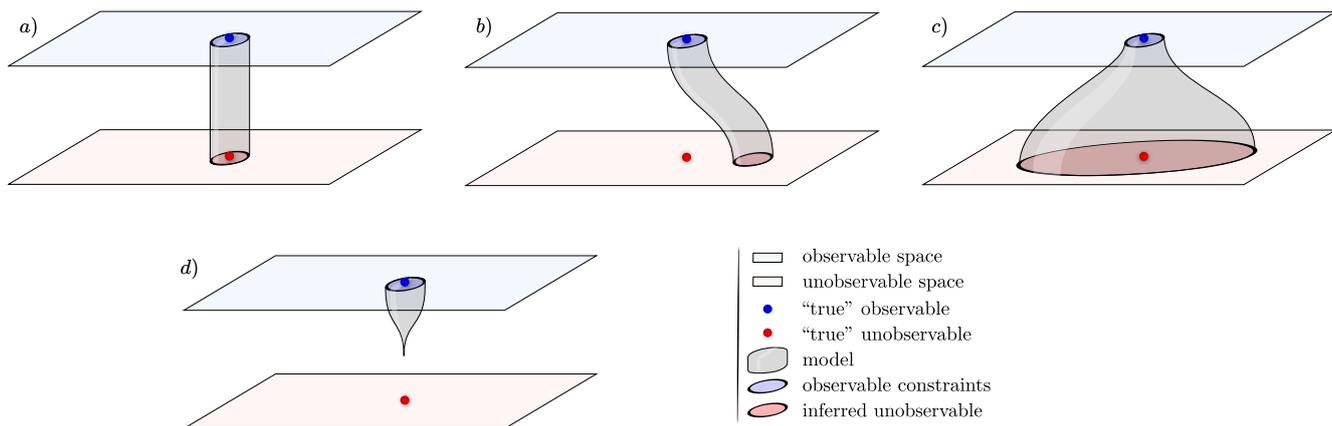


Fig. 1. Schematic illustration of several potential biases involved in the inference problem. Applied to the abundance determination problem, the observable space represents (notably) the oscillation frequency sets, while the unobservable space would represent the model abundances. *Case a* shows an ideal scenario, where for given constraints (blue circle) our modelling (in grey) provides unbiased abundances (red circle). *Panel b*: possible expression of physical processes that are missing or are poorly taken into account in the model. It thus results in physical biases although it has ideal observables. *Case c* provides an example of degeneracy, that is, parameters that can vary widely because they are too weakly constrained by the observables. The last *panel d* is a typical illustration of the solar problem, where no mixture seems to satisfy the constraints. In this case, inconsistencies in the model physics or much weaker constraints than actually provided by the oscillation frequencies must be considered in order to provide a solution.

uncertainties surrounding the near-surface region are to be faced, namely surface effects (Christensen-Dalsgaard et al. 1988). As a result of theoretical developments (Canuto 1997; Belkacem et al. 2021) and 3D hydrodynamical simulations (Belkacem et al. 2019; Schou & Birch 2020), our understanding of the mechanisms involved in these effects (as well as their contributions) has definitely improved. However, the more costly procedures (Mosumgaard et al. 2020) are generally avoided in favour of ad hoc corrections of the oscillation frequencies (Kjeldsen et al. 2008; Ball & Gizon 2014; Sonoi et al. 2015), which may affect stellar parameter estimates (Nsamba et al. 2018).

Methods allowing independent measurements of abundances (in particular helium) are therefore the subject of much investigation. The challenge is in fact twofold because a reliable estimate of abundances in turn would provide a constraint on the internal physics that depends on it.

The specific connection between composition and seismic properties of a star has long been known and can easily be understood through the structural change caused by ionisation if ever the abundances were to vary. The fact that this change is localised (in this case in the near-surface region) causes what is called a seismic glitch and manifests itself through an oscillatory deviation of the observed frequencies from a chosen reference (Gough 1990). Although this effect is not specific to the ionisation regions, these regions have benefited from numerous treatments, being both the most pronounced glitch and a potential marker of the helium abundance (Perez Hernandez & Christensen-Dalsgaard 1994; Lopes et al. 1997). From this point on, many studies (Monteiro et al. 1994; Basu et al. 1994; Monteiro & Thompson 1998, 2005; Gough 2002; Houdek & Gough 2007) considered various shapes of structural perturbations to analytically derive the expected frequency shifts. Most procedures exploit inversion formulae (Dziembowski et al. 1990; Gough & Thompson 1991; Antia & Basu 1994) and rely on an analytical modelling of the glitch to give a parameterised form of the oscillation. The expression thus obtained only needs to be fitted to data in order to provide information on the introduced parameters. In order to overcome dependences on irrelevant disturbances in the frequen-

cies (e.g., surface effects in our case), it has also been proposed that the second differences, $d^2\nu$, be studied, rather than the frequencies directly (Gough 1990),

$$d^2\nu_{n,\ell} = \nu_{n+1,\ell} - 2\nu_{n,\ell} + \nu_{n-1,\ell}, \quad (1)$$

where n and ℓ are the associated oscillation radial order and degree, respectively. The seismic diagnostic defined in this way is less sensitive to surface effects and core perturbations while highlighting signal that would come from an intermediate acoustic depth (Ballot et al. 2004).

The main benefit of this (now usual) approach is to avoid complex modelling and therefore part of the issues mentioned above. Being based on a model with far fewer parameters, which reflects a local rather than global structure, the method seeks to make the best use of the information residing only in the low-degree frequencies. Additionally, this procedure is lighter than a direct minimisation or inversion of the frequency differences, which makes it convenient for application to large samples of stars (Verma et al. 2019). On the other hand, it would be a mistake to think that the procedure is “model independent” as it depends on the form of the perturbation that is considered. This procedure may therefore well lead to nonphysical or inaccurate frequency shifts if it relies on a nonphysical or inaccurate glitch modelling. Moreover, if the parameter to be estimated does not appear directly in the model, there is no alternative to calibration on stellar models (Houdek & Gough 2007, 2011; Verma et al. 2014a, 2019), which adds its own uncertainty on the internal physics. This method is therefore largely dependent on the work done beforehand, and this first paper in the series is intended to cover this perspective. A second paper will focus on the seismic effects of the structural perturbation and on the information these effects provide about the abundances.

In the present paper, we propose a physical model of the ionisation region that allows us to derive a semi-analytic description of the structural perturbation caused by a change in abundances. In this respect, we first briefly describe previous modelling in Sect. 2 and provide more detailed picture of the most commonly used model of the ionisation glitch, which is that of Houdek & Gough (2007), hereafter HG07. The formalism of our

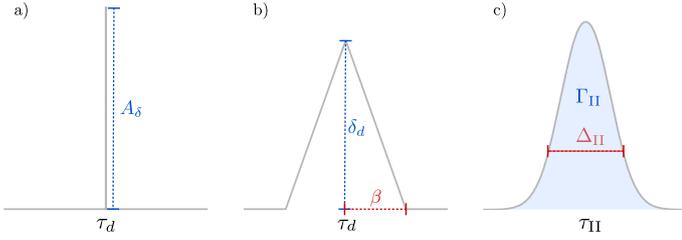


Fig. 2. Various shapes along with their associated parameters that are used to describe a structural perturbation. (a) Dirac function used in Monteiro et al. (1994) to model the variations in the acoustic potential (in the case of an overshoot) passing from a convective to a radiative region. For this modelling, it is only necessary to specify the acoustic depth τ_d and amplitude A_s of the perturbation. (b) Triangular shape of the Γ_1 perturbation in the second helium ionisation region as used in Monteiro & Thompson (2005). An additional parameter controls the width (β) of the perturbation. (c) Gaussian shape of the Γ_1 perturbation in the second helium ionisation region as used in Houdek & Gough (2007). Δ_{II}^2 and Γ_{II} describe the variance and the area of the distribution, respectively.

own model of the ionisation region is introduced in Sect. 3 and its structural behaviour is studied in Sect. 4. In Sect. 5 we propose an analysis of the structural perturbation induced by the model and determine how it is related to analytical models of the ionisation glitch. The last part of the article is dedicated to our conclusions.

2. Previous seismic glitch frameworks

As we briefly mentioned, a twofold approach in the modelling is required to study seismic glitches. First, an expression for the structural perturbation must be considered, and then the shape of the associated frequency shift must be inferred.

2.1. Structural perturbation modelling

The modelling of the perturbation usually boils down to choosing a localised and parametrisable shape for the induced structural perturbation. Depending on the region and the phenomenon under study, various forms can be considered. For the transition between radiative and convective regions, very sharp shapes such as a Dirac or a step function are generally used to reflect a discontinuity in the density derivatives (Monteiro et al. 1994; Houdek & Gough 2007). However, because the ionisation region is more spread out, the shapes that are used to model the structural perturbations usually involve an additional dispersion parameter (Monteiro & Thompson 2005; Houdek & Gough 2007). In Fig. 2 we represent shapes introduced in previous papers with their associated parameterisation. In what follows, we focus on the most commonly used model (i.e. the one presented in HG07) for estimating helium abundances and on the derived expressions for the frequencies.

In HG07, the structural change caused by a change in helium abundance is first modelled as a Gaussian perturbation of the first adiabatic exponent $\Gamma_1 = (\partial \ln P / \partial \ln \rho)_S$,

$$\frac{\delta \Gamma_1}{\Gamma_1} = -\frac{1}{\sqrt{2\pi}} \frac{\Gamma_{\text{HeII}}}{\Delta_{\text{HeII}}} e^{-(\tau - \tau_{\text{HeII}})^2 / 2\Delta_{\text{HeII}}^2}. \quad (2)$$

The idea is as follows: The transition from a pure-hydrogen model to one partially composed of helium causes the appearance of a well in the Γ_1 profile (cf. Fig 3). This results from the second helium ionisation to which the index HeII refers. While

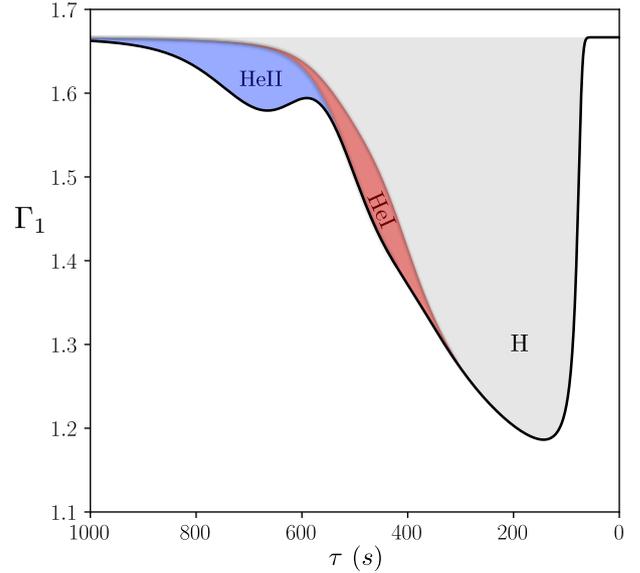


Fig. 3. Representation of a typical Γ_1 profile in the ionisation region as a function of the acoustic depth τ (the surface corresponds to $\tau = 0$). The helium abundance in this plot is $Y \sim 0.25$. The contributions of the three main ionisation zones have been distinguished, i.e., the hydrogen (H), the first (HeI), and second (HeII) helium ionisation zones. Each of them causes a deviation from the Γ_1 reference value of $5/3$.

various analytical functions can be used to model a well, the well appears to be closely reproduced by a Gaussian expressed as a function of the acoustic depth $\tau = \int_r^R dr'/c$; $c(r)$, designating the sound speed at a given position. An important point is that if the depth, position, or width of the well depends on the helium abundance Y or on the thermodynamic conditions of the CZ, then a connection (though implicit) can be made between physical properties of the CZ and the parameters Γ_{HeII} , Δ_{HeII} , and τ_{HeII} .

Before further discussion, it may be useful here to clarify an ambiguity concerning the notation δ that symbolises the perturbation. Because it denotes a difference between two profiles (e.g., between profiles Γ_1^A and Γ_1^B), the variable over which they are expressed has to be chosen carefully. For instance, for arbitrary variables x and y ,

$$\delta_x \Gamma_1 \equiv \Gamma_1^B(x) - \Gamma_1^A(x) \neq \Gamma_1^B(y) - \Gamma_1^A(y) \equiv \delta_y \Gamma_1. \quad (3)$$

As shown above, the notation δ_x allows us to overcome this ambiguity. This notation was introduced by Christensen-Dalsgaard & Thompson (1997), which developed this point extensively. Because the quantity x should vary in the same range for the two profiles A and B , it is natural to choose it as a normalised variable, for example, r/R , m/M , $t = \tau/\tau_0$, \dots (R , M , and $\tau_0 = \int_0^R dr/c$ are the total radius, mass, and acoustic radius of the model). Moreover, the variable on which the perturbation is expressed can differ from the variable that was used to calculate the difference; both $\delta_x \Gamma_1(x)$ and $\delta_x \Gamma_1(y)$ have a meaning.

This point clarified, it appears that the perturbation used in HG07 for Eq. (2) is $\delta_\tau \Gamma_1$ because the Gaussian shape has been established after an expansion at fixed τ (cf. Eq (31) of HG07). As mentioned, δ_τ is ambiguous, however, if $\tau_0^A \neq \tau_0^B$. It can be replaced by δ_t following

$$\frac{\delta_t \Gamma_1}{\Gamma_1} = \frac{\delta_\tau \Gamma_1}{\Gamma_1} + \frac{\delta_t \tau}{\tau} \frac{d \ln \Gamma_1}{d \ln \tau} = \frac{\delta_\tau \Gamma_1}{\Gamma_1} + \frac{\delta \tau_0}{\tau_0} \frac{d \ln \Gamma_1}{d \ln \tau}, \quad (4)$$

with $\delta\tau_0 = \tau_0^B - \tau_0^A$ simply being the perturbation of a constant.

2.2. Expected frequency shift

The resulting signature in the frequencies can be divided in two distinct parts: the ionisation component, and the smooth component. We describe them in turn below.

Ionisation component. The connection between Eq. (2) and the frequency shift $\delta\nu_{n,\ell} = \nu_{n,\ell}^B - \nu_{n,\ell}^A$ is made by considering the following asymptotic form (Gough 1990):

$$\frac{\delta\nu_{n,\ell}}{\nu_{n,\ell}} = \int_0^1 \left[\mathcal{K}_{\rho c^2}^{n,\ell} \frac{\delta x \rho}{\rho} + \mathcal{K}_{c^2 \rho}^{n,\ell} \frac{\delta x c^2}{c^2} \right] dx \sim \int_0^1 \mathcal{K}_{c^2 \rho}^{n,\ell} \frac{\delta x \Gamma_1}{\Gamma_1} dx, \quad (5)$$

where $\mathcal{K}_{\rho c^2}^{n,\ell}$ and $\mathcal{K}_{c^2 \rho}^{n,\ell}$ denote the usual structural kernels that can be found in Gough & Thompson (1991). This derivation, as well as the majority of the frequency shift treatment, will be fully discussed in a second paper. We emphasise that the perturbation used in Eq. (5) must be applied with caution. It can be shown that

$$\frac{\delta x \Gamma_1}{\Gamma_1} = \frac{\delta_r \Gamma_1}{\Gamma_1} - \left[\frac{\delta\tau_0}{\tau_0} - \frac{\delta R}{R} + \frac{1}{x} \int_0^x \frac{\delta_r c}{c} dx' \right] \frac{d \ln \Gamma_1}{d \ln r}. \quad (6)$$

This nuance cannot be appreciated without the notation introduced here, however.

The frequency shift derived in HG07 is written as a continuous function of frequency,

$$\frac{\delta\nu}{\nu} = A_{\text{HeII}} \nu e^{-8\pi^2 \Delta_{\text{HeII}}^2 \nu^2} \cos 2(2\pi\tau_{\text{HeII}}\nu + \epsilon_{\text{HeII}}), \quad (7)$$

as are the second differences,

$$d^2\nu_{\Gamma_1} = F_{\text{HeII}} A_{\text{HeII}} \nu e^{-8\pi^2 \Delta_{\text{HeII}}^2 \nu^2} \cos [2(2\pi\tau_{\text{HeII}}\nu + \epsilon_{\text{HeII}}) - \delta_{\text{HeII}}], \quad (8)$$

with $A_{\text{HeII}} = \pi\Gamma_{\text{HeII}}/\tau_0$. F_{HeII} and δ_{HeII} are functions of the other parameters and the frequency (although it is assumed that they only fluctuate slowly with it). The total number of parameters is therefore only four.

Smooth component. The helium component is not the only perturbation expected in the second differences. In addition to the signature of the transition between radiative and convective regions, a smooth component $d^2\nu_s$ can also be considered and modelled as

$$d^2\nu_s = \sum_{k=0}^3 a_k \nu^{-k}. \quad (9)$$

The idea is to first consider a star that would not contain any glitch. Without any structural perturbation (thus simplifying the problem), it can be assumed that its frequencies follow the asymptotic expansion provided by Tassoul (1980), that is,

$$\nu_{n,\ell} = \left[\frac{1}{2}(2n + \ell + \varepsilon) + \frac{2V_\ell}{2n + \ell + \varepsilon} \right] \Delta\nu + O\left(\frac{1}{\nu_{n,\ell}^2}\right), \quad (10)$$

where ε and V_ℓ are two dimensionless parameters that are independent of the radial order n , and $\Delta\nu \equiv 1/2\tau_0$ is the asymptotic large frequency separation. The last term of Eq. (9), which is proportional to ν^{-3} , can then be interpreted as the leading term of the second differences applied to Eq. (10),

$$d^2\nu_{\text{asym}} \sim \frac{2V_\ell}{\nu^3} (\Delta\nu)^4. \quad (11)$$

However, it is difficult to give as convincing a justification for the other terms, which hide a much more complex deviation to the asymptotic expansion. Houdek & Gough (2011) in particular listed hydrogen ionisation, a sharp stratification of the upper layers of the convection zone, non-adiabatic processes, and turbulent perturbations caused by the oscillations. In the end, this component adds four free parameters to those identified so far. It should be noted that the combination of positive powers of ν introduced in Verma et al. (2014a) to describe the smooth component is incompatible with the asymptotic expansion (10).

2.3. Improvements and limits of the approach

Although it is not apparent from this short summary, the derivation of Eqs. (5)–(8) is as complex as it is intricate, and an analytical derivation of second differences therefore requires making use of many approximations. The first consequence is that a fit of the second differences with Eq. (8) only allows a very approximate retrieval of the parameters introduced in Eq. (2). HG07 revealed typical discrepancies of 50, 33, and 25% on parameters Γ_{HeII} , Δ_{HeII} , and τ_{HeII} , respectively. These differences can be reduced to 5, 15, and 15% by introducing a frequency dependence in ϵ_{HeII} that is induced by the cut-off frequency and a more complete model by adding the first helium ionisation contribution in Eq. (2) (see Fig. 3). Having now two Gaussians (indexed by HeI and HeII), the expected second differences become

$$d^2\nu_{\Gamma_1} = \sum_{i = \{\text{HeI}, \text{HeII}\}} F_i A_i \nu e^{-8\pi^2 \Delta_i^2 \nu^2} \cos [2(2\pi\tau_i \nu + \epsilon_i(\nu)) - \delta_i] \quad (12)$$

instead of Eq. (8). Because the parameters of the two Gaussians are not independent, three of them are fixed in HG07 by the empirical relations $A_{\text{HeI}}/A_{\text{HeII}} = 0.5$, $\Delta_{\text{HeI}}/\Delta_{\text{HeII}} = 0.9$, and $\tau_{\text{HeI}}/\tau_{\text{HeII}} = 0.7$. The number of independent parameters in the expression is still four (only one is needed to determine $\epsilon_{\text{HeI}}(\nu)$ and $\epsilon_{\text{HeII}}(\nu)$). However, despite the substantial improvement in the parameter retrieval, Eq. (8) remains the currently most frequently used equation for studying ionisation glitches (Verma et al. 2017, 2019; Farnir et al. 2019).

More importantly, the abundances, and in particular, the helium abundance Y , do not appear directly as parameters in Eq. (12). As mentioned above, without further theoretical work to establish a dependence between these parameters and Y , calibration based on realistic models therefore seems to be necessary in order to make it appear. This is probably the greatest criticism that can be made of the method. To illustrate this point, we place ourselves in a broader framework and consider two models \mathcal{P} and \mathcal{M}_\star that represent the structural perturbation caused by a change in helium δY and the structure of a realistic stellar model, respectively. The first is parameterised by a set of values θ_p and the second by the set θ_\star from which we explicitly distinguish Y for its particular role in this context. As an example, if the model \mathcal{P} is chosen to be a Gaussian, as done by HG07 (Eq. (2)), it would then involve three parameters $\theta_p = (\Gamma_{\text{HeII}}, \Delta_{\text{HeII}}, \tau_{\text{HeII}})$. In contrast, θ_\star instead contains quantities needed to calculate a realistic model, such as fundamental parameters (mass M_\star , radius R_\star , age A_\star , etc.), but also all the quantities required to model the physical processes involved in the star (overshoot, mixing length, parameters implied in diffusion or rotation, etc.; cf. Lebreton & Goupil 2014, which provides a broad idea of the various possible prescriptions). As a result, θ_\star generally contains many more components than θ_p . With this in mind, we would like to be able to relate a fitted set of parameters, θ_p , to a differ-

ence in helium δY from a chosen reference Y . For this purpose, the solution proposed by calibration is to assimilate the perturbation model \mathcal{P} to a model perturbation $\delta\mathcal{M}$, and therefore to assume

$$\mathcal{P}(\theta_p) \simeq \mathcal{M}_*(Y + \delta Y; \theta_*) - \mathcal{M}_*(Y; \theta_*) \simeq \left. \frac{\partial \mathcal{M}_*}{\partial Y} \right|_{Y; \theta_*} \delta Y \quad (13)$$

in order to numerically derive a relation $\theta_p(\delta Y)$. We note that Eq. (13) is equivalent to Eq. (33) of HG07 when \mathcal{P} is chosen to be the sum of two Gaussians and $Y = 0$ as a reference (although the choice of θ_* is not specified in HG07). In order to be able to associate a set of parameters θ_p with a unique δY , the calibration method therefore makes the assumption that $\partial \mathcal{M}_*/\partial Y$ does not depend on the reference point $(Y; \theta_*)$. However, despite some arguments that are given in favour of a relative independence regarding the choice of Y , we show in Sect. 5 that this assumption about θ_* is inadequate. These dependences cannot be reflected in model \mathcal{P} , which depends on too few parameters, namely θ_p (we note that this is part of the reasons for introducing \mathcal{P} in the first place). To minimise the error that is introduced, it is then necessary to determine as many $\theta_p(\delta Y)$ relations as stars that are studied (given independent constraints on θ_* , see Verma et al. 2019), which can be quite costly. Moreover, by using such a relation anyway, we reintroduce the biases we wished to avoid when studying the ionisation glitch. The relation that is obtained will largely depend on the physics considered in \mathcal{M}_* , in particular, on mechanisms that are not necessarily relevant for modelling the ionisation region.

To summarise, even though it is based on a more sophisticated model than previous methods, the method described above faces challenges that are mainly due to its empirical introduction. In particular, the lack of an explicit dependence on Y greatly reduces its applicability. We would like to reverse the approach. The idea is firstly, to introduce a physics-based parameterisation allowing inferences concerning ionisation regions, and secondly, to give a mathematical description regarding the Γ_1 profile. The challenge is to provide such a modelling while keeping the number of parameters as low as possible. This problem is addressed in the two following sections.

3. First adiabatic exponent in an ionisation region

In order to model the structure of the ionisation zone, we first try to provide an analytical expression of Γ_1 in this specific region. As previously mentioned, we avoid introducing it through similarities with mathematical functions, but rather derive it from thermodynamic relations. Expressions like this have already been obtained (e.g. Cox & Giuli 1968; Kippenhahn et al. 2012), but are generally functions of the occupation number rather than state variables (such as T and V). Furthermore, because the chemical equilibrium resulting from Saha's equations is complex, these equations are generally solved numerically, and analytical expressions are limited to the ionisation of single species. We try to present in the following the simplest model that can nevertheless explicitly involve chemical composition.

3.1. Free energy

Our starting point is the free energy of a multi-component ideal gas, that is,

$$F(T, V, N_\alpha) = \sum_{\alpha} N_{\alpha}(\mu_{\alpha} - kT), \quad (14)$$

where V designates the volume, kT is the temperature expressed in energy units, and N_{α} and $\mu_{\alpha}(T, V, N_{\alpha})$ are the number and chemical potential of particles of type α . In the context of partial ionisation, α refers to (i, r) , $1 \leq i \leq N_{\text{sp}}$ being an index on the chemical species and $0 \leq r \leq z_i$ the ionisation state (where z_i is the atomic number), or it may also correspond to e for an electron. The key point of Eq. (14) lies in its ability to provide us with the pressure P and entropy S through its first derivative. Moreover, its second derivative gives access to virtually any other thermodynamic quantity ranging from the speed of sound through the adiabatic exponents including Γ_1 to various compressibilities. Before the calculations, we would like to define the particular meaning of a derivative in stellar physics compared with pure thermodynamics in more detail. The first adiabatic exponent is often defined as follows:

$$\Gamma_1 \equiv \left(\frac{\partial \ln P}{\partial \ln \rho} \right)_S = - \left(\frac{\partial \ln P}{\partial \ln V} \right)_S, \quad (15)$$

where the subscript S denotes a partial derivative taken at constant entropy. The above expression can easily be rewritten using the more common variables T and V ,

$$\Gamma_1 = \frac{V}{P} \left(\frac{\partial S}{\partial T} \right)_V^{-1} \left[\left(\frac{\partial P}{\partial T} \right)_V \left(\frac{\partial S}{\partial V} \right)_T - \left(\frac{\partial P}{\partial V} \right)_T \left(\frac{\partial S}{\partial T} \right)_V \right]. \quad (16)$$

Although this equation is generally valid, its use in the present case requires some clarification because it only mentions the two state variables T and V and does not explicitly show the dependence on the various number of particles N_{α} . Assuming that species are indexed so that $z_i = i$ ($N_i = \sum_r N_i^r$ potentially being null), the free energy introduced in Eq. (14) is a function of $N_{\text{sp}}(N_{\text{sp}} + 3)/2 + 3$ variables. As a result, the partial derivatives of Eqs. (15) and (16) are generally ambiguous unless some indication is given concerning the $N_{\text{sp}}(N_{\text{sp}} + 3)/2 + 1$ implicit conservation equations. Partial derivatives are commonly considered at constant values of state variables, for example,

$$S = - \left(\frac{\partial F}{\partial T} \right)_{V, N_{\alpha}}, \quad P = - \left(\frac{\partial F}{\partial V} \right)_{T, N_{\alpha}}, \quad \forall \alpha \quad \mu_{\alpha} = \left(\frac{\partial F}{\partial N_{\alpha}} \right)_{T, V, N_{\beta \neq \alpha}}. \quad (17)$$

However, it is clear that the derivatives appearing in Eq. (16) (and second derivatives of F considering Eq. (17)) are not of this kind. As mentioned in the introductory section, the Γ_1 profile used in stellar physics shows clear signs of ionisation that derivatives taken at constant N_{α} do not. This choice imposes a particular ionisation state in the whole region (described by the constant N_{α} values), which instead should fluctuate widely with the thermodynamic conditions in the CZ. The solution is to consider the following conservation equations instead:

$$\forall i, \forall r > 0, \quad \mu_i^r + \mu_e - \mu_i^{r-1} = 0 \quad \left[\frac{N_{\text{sp}}(N_{\text{sp}} + 1)}{2} \right] \quad (18)$$

$$\forall i, \quad \sum_{r=0}^{z_i} N_i^r = N_i = \text{cnst} \quad [N_{\text{sp}}] \quad (19)$$

$$\sum_{i=1}^{N_{\text{sp}}} \sum_{r=1}^{z_i} r N_i^r = N_e \quad [1]. \quad (20)$$

The first equation corresponds to the chemical equilibrium of the ionisation reaction, $A_i^{r-1} \rightleftharpoons A_i^r + e^-$. The second equation is merely the conservation of each atom number and can also be

considered as a chemical equilibrium in the absence of reactions and homogeneous mixing of the CZ. The last expression corresponds to the overall charge equilibrium, that is, electroneutrality. This set of relations can be interpreted as a local equilibrium for given temperature and volume conditions that should be verified at each point of the CZ, and partial derivatives taken with respect to these constraints are hereafter be denoted with the subscript EQ. The resulting number of constraints is indicated in square brackets, and these constraints add up to $N_{\text{sp}}(N_{\text{sp}}+3)/2+1$ as required.

Hereafter, we use the notation $\partial_{\alpha\beta}^2 F$ to designate a second derivative taken with respect to β (subject to N_α being constant) and then α (subject to EQ). For example, Eq. (16) becomes

$$\Gamma_1 \equiv - \left(\frac{\partial \ln P}{\partial \ln V} \right)_{s, \text{EQ}} = V \frac{(\partial_{VT}^2 F)(\partial_{TV}^2 F) - (\partial_{VV}^2 F)(\partial_{TT}^2 F)}{(\partial_V F)(\partial_{TT}^2 F)} \quad (21)$$

by using relations (17). Considering Eq. (14), the evaluation of the Eq. (21) now requires finding $N_i^r(T, V)$ at a given EQ.

3.2. Approximate local equilibrium

To solve the system (18)–(20), the expression of the chemical potential of each particle must first be explicitly be written out. We have

$$\mu_\alpha(T, V, N_\alpha) = kT \ln \left(\frac{N_\alpha \lambda_\alpha^3}{\mathcal{Z}_\alpha V} \right), \quad (22)$$

with $\lambda_\alpha(T)$ and $\mathcal{Z}_\alpha(T)$ the thermal De Broglie wavelength and the canonical partition function of a particle α , respectively. Conditions (18) thus become

$$\forall i, \forall r > 0, \quad \frac{x_i^r}{x_i^{r-1}} = \frac{\mathcal{Z}_i^r \mathcal{Z}_e}{\mathcal{Z}_i^{r-1} N_e \lambda_e^3} V, \quad (23)$$

with $x_i^r = N_i^r/N_i$ the occupation number of state (i, r) . In our effort to obtain the simplest model that can be used to handle chemical composition, the ideal gas assumption remains relevant. We note, however, that in order to take the ionisation process in a more accurate way into account, it would be necessary to consider the Coulomb effects (Rogers 1981; Rogers et al. 1996). In this ideal case, \mathcal{Z}_i^r and \mathcal{Z}_e are often approximated by (e.g., Kippenhahn et al. 2012)

$$\mathcal{Z}_i^r(T) \simeq u_i^r(T) \exp \left(\sum_{s=1}^r \chi_i^s / kT \right) \simeq g_i^r \exp \left(\sum_{s=1}^r \chi_i^s / kT \right) \quad (24)$$

$$\mathcal{Z}_e \simeq 2, \quad (25)$$

with u_i^r describing the fine structure of state (i, r) approached by the degeneracy of the ground state g_i^r , and χ_i^r the ionisation energy separating state (i, r) from $(i, r-1)$ ($\sum_s \chi_i^s / kT$ is then the energy from the reference state $(i, 0)$, which is assumed here to have a null internal energy). The second approximation (25) directly results from neglecting the electron spin energy contribution.

Thus, conditions (23) simply become Saha's equations,

$$\forall i, \forall r > 0, \quad \frac{x_i^r}{x_i^{r-1}} = \frac{2g_i^r}{g_i^{r-1}} \frac{V}{N_e \lambda_e^3} e^{-\chi_i^r / kT}. \quad (26)$$

These equations are highly coupled, especially because of the term N_e , which can be expressed from (20) as

$$N_e = N \sum_{i=1}^{N_{\text{sp}}} \sum_{r=1}^{z_i} r x_i x_i^r \quad (27)$$

with $N = \sum_i N_i$ and $x_i = N_i/N$ the total number of atoms and the number abundance of element i , respectively, which are two constants from (19). Because of this, Saha's equations are generally solved numerically by iterating over estimates of the x_i^r . In the following, we try to give analytical approximations by considering a few simplifying assumptions. First, we introduce the cumulative occupation number $y_i^r = \sum_{s \geq r} x_i^s$, which is more suited to the study of Eqs. (26) (Baker & Kippenhahn 1962). We also use from this point the convention

$$\sum_{ir} \equiv \sum_{i=1}^{N_{\text{sp}}} \sum_{r=1}^{z_i}$$

In this way, we can write $N_e = N \bar{z}$ with

$$\bar{z} = \sum_{ir} x_i y_i^r, \quad (28)$$

the mean number of electrons per atom. Saha's equations can be rewritten as

$$\forall i, \forall r > 0, \quad \frac{y_i^r - y_i^{r+1}}{y_i^{r-1} - y_i^r} = \frac{2g_i^r}{g_i^{r-1}} \frac{V}{N \bar{z} \lambda_e^3} e^{-\chi_i^r / kT}, \quad (29)$$

with the convention $y_i^{z_i+1} = 0$ (and by definition $y_i^0 = 1$). We start by noting that it is rare to find more than two ionisation states simultaneously for a given element. It is thus advantageous to make the following approximation (Baker & Kippenhahn 1962):

$$\forall i, \forall r > 0, \text{ if } y_i^r \neq 0, \text{ then, } \quad \begin{aligned} \forall s < r, y_i^s &= 1 \\ \forall s > r, y_i^s &= 0 \end{aligned} \quad (30)$$

Then, in the single domain of interest ($y_i^r \neq 0$), (29) is approached by

$$\forall i, \forall r > 0, \quad \frac{y_i^r}{1 - y_i^r} = \frac{2g_i^r}{g_i^{r-1}} \frac{V}{N \bar{z} \lambda_e^3} e^{-\chi_i^r / kT}. \quad (31)$$

A final assumption is needed to express \bar{z} as a function of y_i^r only and to completely decouple the system. One idea may be to exploit the fact that stars are predominantly composed of hydrogen in number: For typical Y values (i.e. $1/4 < Y < 1/3$), $x_1 \sim 0.89-0.92$, and it can be seen using (28) that after a complete ionisation of the hydrogen, $\bar{z} > x_1$. Another bound can then be obtained on \bar{z} because the latter tends asymptotically towards $\sum_i z_i x_i = 2 - x_1 + \sum_{i>2} (z_i - 2) x_i$. Because for heavy elements ($i > 2$), $m_i \simeq 2z_i m_u$ with m_i the mass of atom i and m_u the atomic mass unit, and because the mean mass $m_0 < 2m_u$ in a stellar mixture, we have $m_0/m_i < 1/z_i$. Finally, the number and mass abundances are related by $x_i = m_0 X_i / m_i$, meaning that \bar{z} should remain below the upper boundary $2 - x_1 + Z \sim 1.1$ with $Z = \sum_{i>2} X_i$ the metal abundance in mass. This simple reasoning allows us to convince ourselves that \bar{z} must remain close to 1 when the ionisation of hydrogen is completed, and in a way relatively independent of the mixture under consideration.

The pragmatic assumption can therefore be made that the ionisation of hydrogen occurs first and is completed before any other ionisation begins. Thus we can use the fact that hydrogen is the dominant component to approximate \bar{z} by the mean electron number in a pure-hydrogen model. Although this is not entirely true (particularly not for a low ionisation state), this assumption greatly simplifies the problem. For all elements $i > 1$, $\bar{z} \sim 1$, which, as discussed above, is a reasonable approximation. For the ionisation equation corresponding to hydrogen ($i = 1$), we

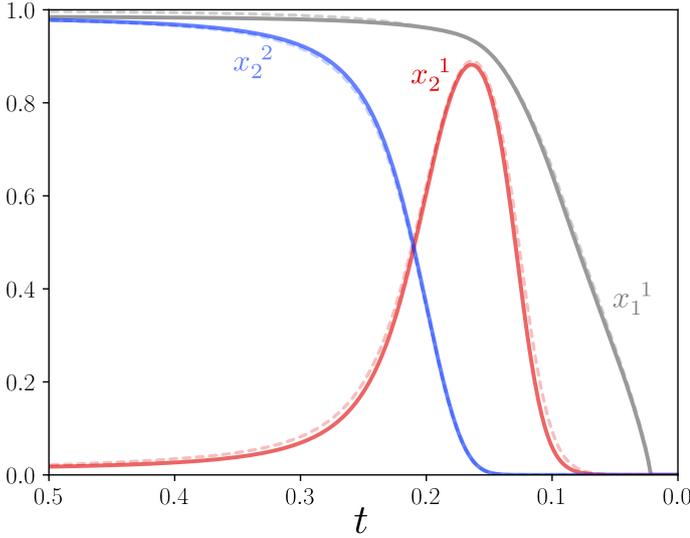


Fig. 4. Comparison of approximations (33) and (34) in the case $i = 1, 2$ (opaque curves) with the numerical solution of (26) (dashed light curves) for solar near-surface conditions of $(T, V, X_{i>1})$. These quantities are represented as a function of the normalised acoustic depth $t = \tau/\tau_0$ (the surface is then located at $t = 0$), along which all future plots are represented.

instead approximate $\bar{z} \sim y_1^1$. Our version of Saha's equations can thus finally be written as

$$\forall i, \forall r > 0, \quad \frac{(y_i^r)^n}{1 - y_i^r} = \frac{2g_i^r}{g_i^{r-1}} \frac{V}{N\lambda_e^3} e^{-X_i^r/kT}, \quad (32)$$

with $n = 2$ if $i = 1$ and $n = 1$ in any other case. We define $K_i^r(T, V)$ as the term on the right-hand side (RHS) of (32). We then obtain the desired relations.

$$y_1^1(T, V) = \frac{1}{2} \left[\sqrt{K_1^1(K_1^1 + 4)} - K_1^1 \right] \quad (33)$$

$$\forall i > 1, \forall r > 0, \quad y_i^r(T, V) = \frac{K_i^r}{1 + K_i^r} \quad (34)$$

It is easy to retrieve $x_i^r(V, T)$ from there using $x_i^r = y_i^r - y_i^{r+1}$ (as well as $N_i^r(V, T)$ using $N_i^r = N_i x_i^r$). Figure 4 shows a comparison of hydrogen and helium ionisation fractions x_1^1, x_2^1 , and x_2^2 obtained analytically from (33) and (34) for given T and V values with the numerical solution of Saha's Eqs. (26). The result is shown as a function of the normalised acoustic depth t so that it can be reduced to a comparison of profiles rather than a (V, T) mapping. The solution presented above appears to be satisfying considering the number of assumptions made. However, x_2^1 is slightly underestimated because \bar{z} is slightly lower than 1 in the first helium ionisation region. The opposite reasoning holds for the second helium ionisation region.

3.3. First adiabatic exponent

Approximations (33) and (34) allows us to derive all derivatives appearing in (21). First of all, we have

$$\partial_V F = -P = \frac{-kT}{V} \sum_{\alpha} N_{\alpha} = \frac{-NkT}{V} \left(1 + \sum_{ir} x_i y_i^r \right), \quad (35)$$

where we explicitly made the y_i^r apparent to fully exploit relations (33) and (34). It is straightforward to show that we simply

retrieve the classical expression for the pressure in an ideal gas. Second derivatives taken at EQ ensue,

$$\partial_{VV}^2 F = \frac{NkT}{V^2} \left[1 + \sum_{ir} x_i y_i^r \left(1 - \frac{\partial \ln y_i^r}{\partial \ln V} \right) \right] \quad (36)$$

$$\partial_{TV}^2 F = \frac{-Nk}{V} \left[1 + \sum_{ir} x_i y_i^r \left(1 + \frac{\partial \ln y_i^r}{\partial \ln T} \right) \right]. \quad (37)$$

More explicit expressions can be deduced by writing $\frac{\partial \ln y_i^r}{\partial \ln \alpha} = \frac{d \ln y_i^r}{d \ln K_i^r} \frac{\partial \ln K_i^r}{\partial \ln \alpha}$ and exploiting Eq. (32),

$$\forall i, \forall r > 0, \quad \frac{d \ln y_i^r}{d \ln K_i^r} = \frac{1 - y_i^r}{1 + \delta_i^1 (1 - y_i^r)} \quad (38)$$

$$\frac{\partial \ln K_i^r}{\partial \ln V} = 1 \quad (39)$$

$$\frac{\partial \ln K_i^r}{\partial \ln T} = \frac{3}{2} + \frac{X_i^r}{kT} \equiv \phi_i^r \quad (40)$$

δ_i^j standing for Kronecker's symbol. Resulting expressions of the derivatives are

$$\partial_{VV}^2 F = \frac{NkT}{V^2} \left[1 + \sum_{ir} x_i y_i^r \left(1 - \frac{1 - y_i^r}{1 + \delta_i^1 (1 - y_i^r)} \right) \right] \quad (41)$$

$$\partial_{TV}^2 F = \frac{-Nk}{V} \left[1 + \sum_{ir} x_i y_i^r \left(1 + \frac{(1 - y_i^r) \phi_i^r}{1 + \delta_i^1 (1 - y_i^r)} \right) \right]. \quad (42)$$

The calculation of the two remaining derivatives is more subtle, however. We propose here to start with a determination of the energy E ,

$$E = -T^2 \left(\frac{\partial}{\partial T} \left(\frac{F}{T} \right) \right)_{V, N_{\alpha}} = NkT \left(\frac{3}{2} + \sum_{ir} x_i y_i^r \phi_i^r \right). \quad (43)$$

Because $\partial_{VT}^2 F = -(\partial S / \partial V)_{T, EQ}$ and $\partial_{TT}^2 F = -(\partial S / \partial T)_{V, EQ}$, it might be tempting to take directly the derivatives of (43) using $dE = TdS - PdV$. However, in the present framework,

$$\begin{aligned} dE &= TdS - PdV + \sum_{\alpha} \mu_{\alpha} dN_{\alpha} \\ &= TdS - PdV + N \sum_{ir} x_i (\mu_i^r + r\mu_e) dx_i^r + N \sum_i x_i \mu_i^0 dx_i^0 \\ &= TdS - PdV + N \sum_{ir} x_i (\mu_i^r + \mu_e - \mu_i^{r-1}) dy_i^r \\ &\quad + N \sum_{ir} x_i \mu_i^0 dx_i^r + N \sum_i x_i \mu_i^0 dx_i^0. \end{aligned} \quad (44)$$

Clearly, the last two sums cancel each other because $\forall i, \sum_{r>0} dx_i^r = d(1 - x_i^0) = -dx_i^0$, which is verified in our derivation as long as we define $x_i^0 \equiv 1 - y_i^1$. Electroneutrality, used in line 2 of Eq. (44), is also verified by imposing $N_e \equiv N \sum_{ir} x_i y_i^r$. According to Eq. (18), $\mu_i^r + \mu_e - \mu_i^{r-1} = 0$, and the third term must also cancel exactly. In practice, Eq. (18) is too complex to be perfectly solved, and numerical calculations as well as analytical approximations can lead to some small departures from exact cancellation. These may lead to slight thermodynamic inconsistencies (which could be corrected through departures from equality of partial mixed second derivatives of state functions such as the free energy F). However, these residuals remain small, and we considered that chemical equilibrium is perfectly

satisfied. The only part that therefore remains from (44) is the following well-known identity:

$$dE = TdS - PdV. \quad (45)$$

This means that $T\partial_{VT}^2 F = -(\partial E/\partial V)_{T,EQ} - P$ and $T\partial_{TT}^2 F = -(\partial E/\partial T)_{V,EQ}$. It follows from Eq. (43) that

$$\partial_{VT}^2 F = \frac{-Nk}{V} \left[1 + \sum_{ir} x_i y_i^r \left(1 + \frac{(1-y_i^r)\phi_i^r}{1+\delta_i^1(1-y_i^r)} \right) \right] \quad (46)$$

$$\partial_{TT}^2 F = \frac{-Nk}{T} \left[\frac{3}{2} + \sum_{ir} x_i y_i^r \left(\frac{3}{2} + \frac{(1-y_i^r)(\phi_i^r)^2}{1+\delta_i^1(1-y_i^r)} \right) \right]. \quad (47)$$

Denoting the bracket part of $\partial_{\alpha\beta}^2 F$ as $\partial_{\alpha\beta}^2 f$, we obtain

$$\Gamma_1 = \frac{(\partial_{VT}^2 f)(\partial_{TV}^2 f) + (\partial_{VV}^2 f)(\partial_{TT}^2 f)}{(\partial_V f)(\partial_{TT}^2 f)}, \quad (48)$$

where

$$\partial_V f = 1 + \sum_{ir} x_i y_i^r, \quad (49)$$

$$\partial_{VV}^2 f = 1 + \sum_{ir} x_i y_i^r \left(1 - \frac{1-y_i^r}{1+\delta_i^1(1-y_i^r)} \right), \quad (50)$$

$$\partial_{TV}^2 f = \partial_{VT}^2 f = 1 + \sum_{ir} x_i y_i^r \left(1 + \frac{(1-y_i^r)\phi_i^r}{1+\delta_i^1(1-y_i^r)} \right), \quad (51)$$

$$\partial_{TT}^2 f = \frac{3}{2} + \sum_{ir} x_i y_i^r \left(\frac{3}{2} + \frac{(1-y_i^r)(\phi_i^r)^2}{1+\delta_i^1(1-y_i^r)} \right). \quad (52)$$

The various symmetries present in these equations may give the impression that expression (48) can be further simplified. We show in Appendix A that Γ_1 can finally be written as

$$\Gamma_1 = \frac{5}{3} - \frac{2}{3}\gamma_1, \quad (53)$$

with $0 \leq \gamma_1 < 1$ simply expressed as

$$\gamma_1 = \frac{1}{\partial_{TT}^2 f} \sum_{ir} x_i y_i^r (1-y_i^r) \frac{(\chi_i^r/kT)^2}{1+\delta_i^1(1-y_i^r)}. \quad (54)$$

Together with Eqs. (32)–(34), Eqs. (52)–(54) provide an analytical approximation of $\Gamma_1(T, V, x_i)$ in the CZ. Even though this expression is applicable to any kind of mixture (but is assumed to be homogeneous in the CZ!), the following examples are based on a hydrogen-helium mixture of mass fraction $1 - Y$ and Y ($Y = 4x_2/(1 + 3x_2)$) to facilitate the study. We recall that $N/V = \rho/m_0$, and represent the resulting $\Gamma_1(\rho, T, Y)$ in Fig. 5 for $Y = 0.25$ as well as for extreme values, $Y = 0$ and $Y = 1$. Each panel also shows the relation $\rho(T)$ extracted from a CESTAM solar model (Morel & Lebreton 2008; Marques et al. 2013). The latter is obviously only indicative as it depends on the composition, but it still provides an idea of what part of the map is visible in a Γ_1 profile. The figure clearly reflects the Γ_1 depressions caused by (from left to right) the hydrogen and first and second helium ionisations. Each element contribution is enhanced in the left panels. In the top right panel, the superposition of the hydrogen and helium first ionisation zones in typical solar conditions is presented, whereas the distinction becomes apparent at lower densities. In contrast, this signature appears to be more diffuse at the highest densities. This behaviour appears in each panel and

is further discussed in Sect. 4. The last frame shows the variation $\delta_{\rho,T}\Gamma_1$ caused by a change in helium abundance, $\delta Y = 0.1$, from a reference value $Y = 0.25$. Although the variables are not normalised, the notation $\delta_{\rho,T}$ is still meaningful because Γ_1 does not refer to a profile here, but to a function defined for all (ρ, T) at a given Y . As expected, this variation in abundance causes the two helium wells to become lower while elevating the hydrogen well. However, the variations appear to be fairly disproportionate between the two elements. The change in the hydrogen ionisation region would be barely noticeable considering the typical solar relation $\rho(T)$, shown in black. Appendix A provides clues that help to understand this difference in behaviour as well as the particular shape of the perturbation in the hydrogen region compared with that of helium. In particular, the amplitude of the variation seems more related to the relative change $\delta x_i/x_i$ than δY . While we have $\delta x_1 = -\delta x_2$ for the hydrogen-helium case, we see that $\delta x_1/x_1 = -\delta x_2/x_1 = -(x_2/x_1)\delta x_2/x_2$ with $(x_2/x_1) = 1/12$ for $Y = 0.25$. This explains why the change in hydrogen is expected to cause a variation that is an order of magnitude lower than the change in helium.

Although Eqs. (53) and (54) appear to be promising, they are merely functional forms and do not allow us at this point to model a parameterised structure of the ionisation region. However, Γ_1 , ρ and T are closely related in stellar interiors.

4. Model structure and properties

In this section we obtain the stellar structure associated with the Γ_1 expression derived in the previous section (see Eqs. (53) and (54)). To do this, we assume the conditions of a CZ that correspond to the assumptions for deriving the first adiabatic exponent. These are (1) an isentropic region, (2) chemical equilibrium at any position, (3) uniform abundances (due to homogeneous mixing), and (4) electroneutrality at any position. The last three conditions correspond to applying EQ conditions at any point in the region. Moreover, to avoid getting lost in the many possibilities offered by these relations in terms of the mixture, we again consider a relation of the form $\Gamma_1(\rho, T, Y)$ in this part. It is possible to generalise all the observations made here to any (reasonable) mixture by replacing Y by $X_{i>1}$, however.

4.1. Ionisation region structure

As mentioned above, the quantities Γ_1 , ρ and T are not independent in stellar interiors, and although it may be interesting to study Γ_1 for any condition of temperature and density, relevant profiles in these environments are much more constrained. The appropriate equations to express these constraints are the well-known hydrostatic equilibrium and Poisson's equation assuming spherical symmetry,

$$\frac{dP}{dr} = -\rho g \quad (55)$$

$$\frac{dg}{dr} + \frac{2g}{r} = 4\pi G\rho, \quad (56)$$

where we have introduced the additional intermediate variable g , the norm of the gravity field. At this point, it is important to recall that we try to account for the variation of T and ρ in the region rather than that of P . This can be achieved by noting

that for any quantity α , $\frac{d\alpha}{dr} = \frac{-\rho g \alpha}{P} \left(\frac{d \ln P}{d \ln \alpha} \right)^{-1}$. Then, crucially,

we note that derivatives taken with respect to the radius r such as $d \ln P/d \ln \alpha = (d \ln P/dr)(d \ln \alpha/dr)^{-1}$ exactly correspond to

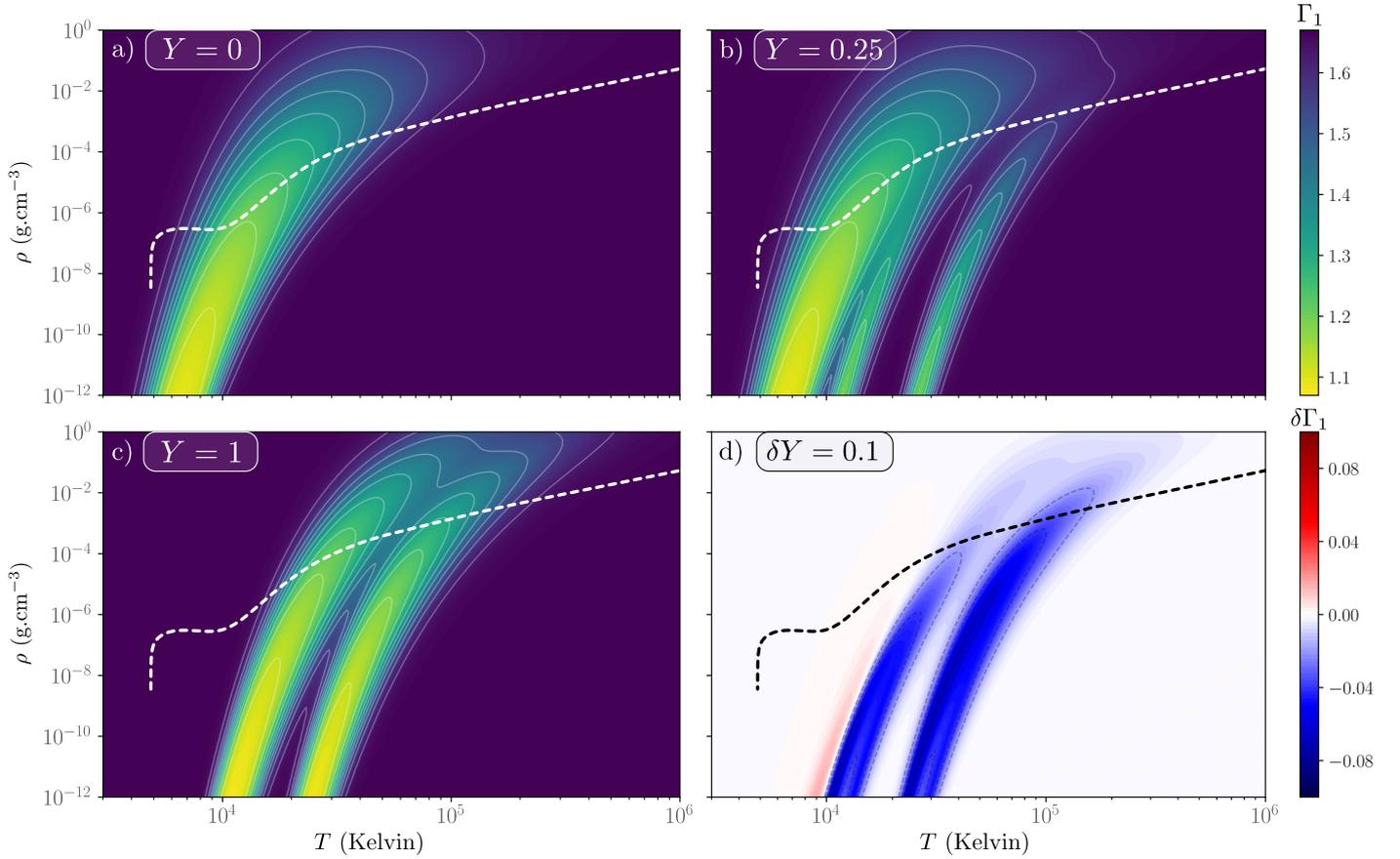


Fig. 5. 2D maps of the first adiabatic exponent obtained from Eq. (53). *Panels a–c:* $\Gamma_1(\rho, T, Y)$ as a function of ρ and T for different values of Y . *Panel d:* variation $\delta_{\rho, T}\Gamma_1$ caused by a perturbation $\delta Y = 0.1$ from a reference value $Y = 0.25$. The dashed line in each panel shows the relation $\rho(T)$ extracted from a CESTAM solar model.

derivatives taken at constant entropy and EQ (cf. assumptions 1–4 at the beginning of the section). Using this for $\alpha = T, \rho$, Eqs. (55) and (56) can be changed into a differential system of three equations, the solution of which is (T, ρ, g)

$$\frac{dT}{dr} = -\frac{\rho T g}{P} \frac{\Gamma_2 - 1}{\Gamma_2} \quad (57)$$

$$\frac{d\rho}{dr} = -\frac{\rho^2 g}{P} \frac{1}{\Gamma_1} \quad (58)$$

$$\frac{dg}{dr} = 4\pi G \rho - \frac{2g}{r}, \quad (59)$$

where we introduced the second adiabatic exponent Γ_2 such that $\frac{\Gamma_2}{\Gamma_2 - 1} \equiv \left(\frac{\partial \ln P}{\partial \ln T} \right)_{S, EQ}$. Its expression can easily be deduced from Γ_1 by noting that their definitions are symmetrical with respect to a change $V \leftrightarrow T$,

$$\frac{\Gamma_2}{\Gamma_2 - 1} = \frac{T}{P} \left(\frac{\partial S}{\partial V} \right)_T^{-1} \left[\left(\frac{\partial P}{\partial T} \right)_V \left(\frac{\partial S}{\partial V} \right)_T - \left(\frac{\partial P}{\partial V} \right)_T \left(\frac{\partial S}{\partial T} \right)_V \right]. \quad (60)$$

An analogous expression to Eq. (53) can be provided for Γ_2 ,

$$\frac{\Gamma_2}{\Gamma_2 - 1} = \frac{5}{2} + \gamma_2, \quad (61)$$

by introducing

$$\gamma_2 = \frac{1}{\partial_{VT}^2 f} \sum_{ir} x_i y_i^r (1 - y_i^r) \frac{(\chi_i^r / kT)(5/2 + \chi_i^r / kT)}{1 + \delta_i^1 (1 - y_i^r)}. \quad (62)$$

The pressure can also be removed using the ideal gas law (35), which leads to the following system:

$$\frac{dT}{dr} = -\frac{m_0 g}{k} \frac{1}{\partial_{Vf}} \frac{\Gamma_2 - 1}{\Gamma_2} \quad (63)$$

$$\frac{d\rho}{dr} = -\frac{\rho m_0 g}{kT} \frac{1}{\partial_{Vf}} \frac{1}{\Gamma_1} \quad (64)$$

$$\frac{dg}{dr} = 4\pi G \rho - \frac{2g}{r}. \quad (65)$$

All quantities on the RHS are functions of ρ , T , and g only (for a given value of Y). By imposing the central values T_c and ρ_c (g_c must be null in any case), we now have a complete system,

the solution of which is $\left\{ \begin{array}{l} T(r; Y, T_c, \rho_c) \\ \rho(r; Y, T_c, \rho_c) \\ g(r; Y, T_c, \rho_c) \end{array} \right\}$. The resulting model is

fully convective and contains an ionisation region whose properties are controlled by the triplet (Y, T_c, ρ_c) . Although this triplet seems the most intuitive, it is possible to parameterise the model by considering other triplets derived from these three quantities. Hereafter, we instead consider the following dimensionless parameterisation:

$$Y_s \equiv Y, \quad (66)$$

$$\psi_{CZ} \equiv \ln \left[\frac{\rho_c h^3}{2m_u (2\pi m_e k T_c)^{3/2}} \right], \quad (67)$$

$$\varepsilon \equiv \frac{\chi_H}{k T_c} \quad (68)$$

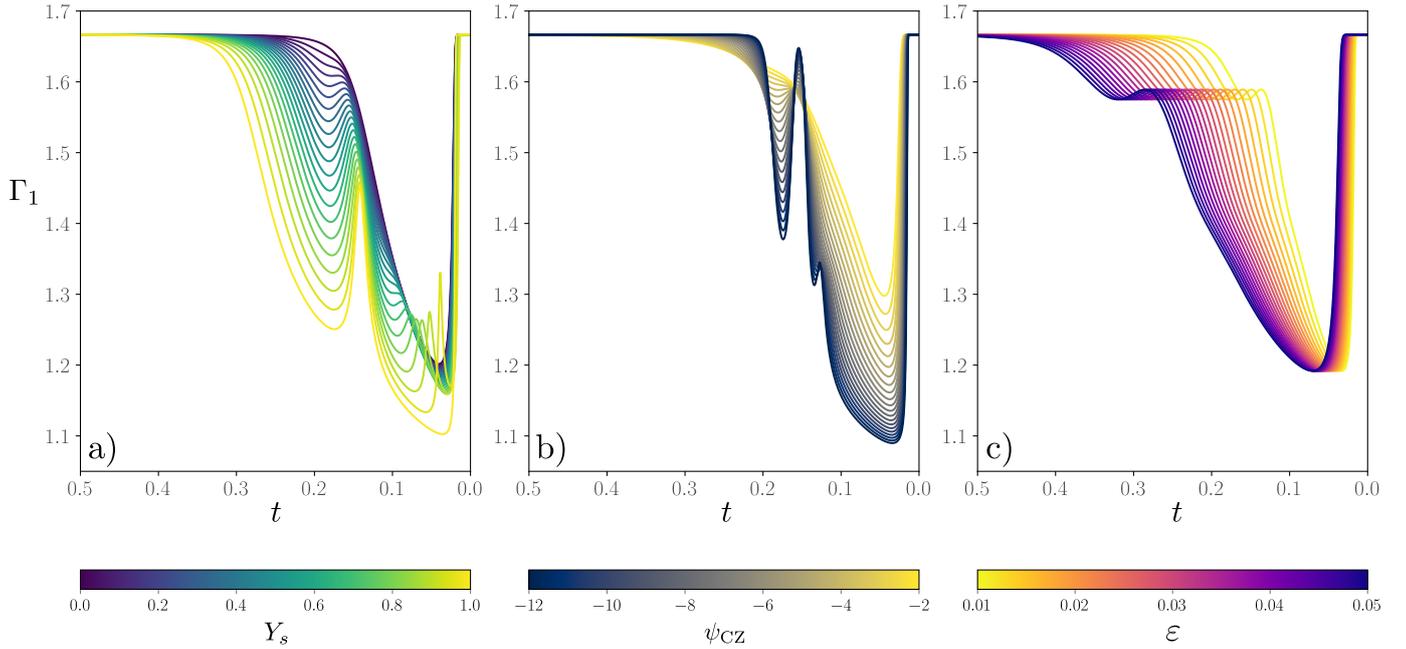


Fig. 6. $\Gamma_1(t)$ profiles obtained by considering various sets $(Y_s, \psi_{CZ}, \varepsilon)$ around a reference $(Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ) = (0.255, -5, 0.015)$ that corresponds to a Γ_1 profile approaching that of the Sun (see Sect. 4.2). (a) Dependence of the profile on $0 \leq Y_s \leq 1$. (b) Dependence of the profile on $-12 \leq \psi_{CZ} \leq -2$. (c) Dependence of the profile on $0.01 \leq \varepsilon \leq 0.05$.

where we introduced the Planck constant h , the electron mass m_e , and the ionisation potential of hydrogen χ_H . This particular choice and the associated notations will become clearer later on. At this point, this choice entirely constrains the Γ_1 value at the centre, the latter being only a function of Y , $2m_u/\rho\lambda_e^3$ (we recall that $\lambda_e^{-2} = 2\pi m_e kT/h^2$), and χ_H/kT (cf. Eqs. (32)–(34) and Eqs. (52)–(54)). However, their impact on our model is much more significant than this.

Physical interpretation of the parameters. Because $Y(r)$ is constant (see assumption 3), the helium amount Y_s is clearly imposed at any location of the model up to the surface symbolised by the index s .

The second parameter derives its notation from its close link to the electron degeneracy parameter $\psi = \mu_e/kT = \ln(n_e\lambda_e^3/2)$ in the ideal gas limit. We can show that

$$\psi_{\text{atcentre}} = \psi_{CZ} - \frac{Y}{2} + O(Y^2). \quad (69)$$

However, as long as $\Gamma_1 = 5/3$ (i.e., outside the ionisation region), $\rho \propto T^{3/2}$, which means that ψ is preserved. Therefore ψ_{CZ} approximates the electron degeneracy parameter from the centre up to the ionisation region.

The last parameter ε results from the ratio of two energy scales in the model. The first is the ionisation energy for the hydrogen (13.6 eV), and the second is the order of magnitude of the thermal energy at the centre of the model (\sim keV), making this ratio generally rather small (hence the notation ε). Through conservations, we just saw that Y_s and ψ_{CZ} reflect properties of the ionisation region; this also holds for its helium abundance and electron degeneracy. It is complicated to apply this to ε because χ_H/kT varies by several orders of magnitude in the model. We try to give an intuitive understanding of its impact on the model structure, however. We first note that if Γ_1 did not depend on χ_H/kT , the latter would be equal to $5/3$ in the whole model. Because Γ_1 would depend only on Y (which is constant) and $2m_u/\rho\lambda_e^3$, the value of which is fixed as long as $\rho \propto T^{3/2}$, the

first adiabatic exponent would keep its central value, that is, $5/3$. The value of χ_H/kT at a given position then alone determines whether Γ_1 deviates from the value $5/3$, that is, if the ionisation region is entered starting from the model centre. It can be shown that the ionisation of the element i in state r takes place when $\chi_i^r/kT \sim -\psi$. The role of ε can now be clearly identified: When ε is to be large (e.g., of the order of $-\psi$), the ionisation will be complete close to the centre. In contrast, if ε is small before $-\psi$, it will take until $T \ll T_c$ for the ionisation to be completed and the region will be shifted closer to the surface. This shows that although it is not explicitly linked to any quantity of the ionisation zone, ε controls the extent of the region.

After this physical interpretation for the parameters, we now try to provide a geometrical interpretation of the impact of each parameter regarding the Γ_1 profile. We rely on Fig. 6, which shows the $\Gamma_1(t)$ (with $t = \tau/\tau_0$, the normalised acoustic depth) profiles resulting from multiple parameter sets $(Y_s, \psi_{CZ}, \varepsilon)$ by varying them one by one around a reference value.

Geometrical interpretation of the parameters regarding the Γ_1 profile. We gave a representation of $\Gamma_1(t)$ for all possible values of Y_s in panel a of Fig. 6, including such extreme cases as pure-hydrogen or helium stars. As expected, the profile changes from one that only shows the hydrogen well ($Y_s = 0$) to one that only shows the two helium wells ($Y_s = 1$), passing through profiles that account for the three components ($0 < Y_s < 1$). Clearly, increasing Y_s has the effect of extending both helium wells while reducing the hydrogen well.

The effects of modifying ψ_{CZ} are shown in panel b of Fig. 6. The range under consideration corresponds to typical solar degeneracy values ranging from the surface (-12 : very low degeneracy) to the centre (-2 : high degeneracy). The impact is again clearly visible: A high degeneracy tends to spread out the ionisation wells, making them indistinguishable when $\psi_{CZ} \rightarrow -2$. In the opposite situation, low degeneracies in the ionisation region result in much deeper and localised wells. It is even possible to distinguish the contributions of the hydrogen and the first

helium ionisations when ψ_{CZ} approaches -12 . This last point is consistent with what we showed in Fig. 5 : high degeneracies correspond to relations of the form $\ln \rho(T) = C + (3/2) \ln T$ with high C values. Thus, structures with high ψ_{CZ} correspond to $\rho(T)$ that are located in the upper part of Fig. 5, where the ionisation regions tend to overlap. The opposite reasoning holds just as well.

The impact of ε is illustrated in panel c of Fig. 6. As expected from the previous paragraph, varying its value from 0.01 to 0.05 results in a shift of the ionisation region. However, this change differs between each well, the HeII well is more strongly affected than the hydrogen well. In order of magnitude, this effect can be explained by the dependence of the normalised acoustic depth on temperature close to the surface: $t \propto \sqrt{T/T_c} \propto \sqrt{\varepsilon T}$. Thus, for a given ionisation temperature T_{ion} , the impact of ε becomes more noticeable on the associated depth t_{ion} as T_{ion} is high. This reasoning also explains why the shift does not appear to be linear with respect to ε for a given well (this is particularly visible for the HeII well); this shift varies as $\sqrt{\varepsilon}$.

Figure 6 shows an interesting point of the model defined with $(Y_s, \psi_{CZ}, \varepsilon)$: the relative independence of the impact of each parameter. When each well is treated as a distribution, all parameters appear to impact a specific moment. For instance, Y_s appears to be directly related to the distribution area and does not impact the shapes of the wells or the position much (for reasonable values of Y_s). ψ_{CZ} is clearly related to the dispersion of the distribution and ε to its position. These distinct impacts on the structure are encouraging for subsequent parameter retrieval as we can expect to avoid degeneracy biases.

4.2. Comparison with a realistic stellar model

The first and most important point to address is to verify the extent to which the model described in this article can approximate the ionisation region structure of a realistic stellar model. To do this, we compare in panel a of Fig. 7 $\Gamma_1(t)$ profiles extracted from a CESTAM solar model and from our model for a particular parameter set $(Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ)$. These values were chosen so that

$$\int_0^1 \left(\frac{\delta_t \Gamma_1}{\Gamma_1} \right)^2 dt = \int_0^1 \frac{(\Gamma_1^{\text{CESTAM}}(t) - \Gamma_1(t; Y_s, \psi_{CZ}, \varepsilon))^2}{\Gamma_1^2} dt \quad (70)$$

is minimal. The behaviour of the realistic model can clearly be very well reproduced, the curves differ mainly in only two regions. A first difference is located in the first helium ionisation region ($0.12 < t < 0.18$), where the depth of the well is slightly overestimated. This is mainly caused by the assumption made in Sect. 3.2, where we assumed that $\bar{z} \sim 1$ in the helium ionisation region and thus underestimated the number of electrons in the HeI region. This causes the ionisation to start slightly deeper in the model and therefore to be more localised. By reducing the dispersion of the HeI well, the assumption slightly enlarged its depth. As expected, another difference is visible in the atmospheric region ($t < 0.03$). Because the ionisation has not yet started, our Γ_1 still has the value $5/3$, which is obviously not the case in a realistic model containing an atmospheric modelling. The overall profile shape is nevertheless well reproduced in the ionisation region, especially considering that only three parameters need to be adjusted.

A second point to be clarified is the exact meaning of the adjusted values $(Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ)$ regarding the realistic model. When the minimal differences $(\delta_t \Gamma_1 / \Gamma_1)^2$ are assumed, the model parameterised by $(Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ)$ is interpreted as being the

closest to the CESTAM model from the perspective of a particular structural aspect. This aspect will be discussed in much more detail in a forthcoming paper, but we can already relate it to the “visible” part of the structure through the frequency shift considering Eq. (5). In this sense, if our modelling and its interpretation are correct, we can expect our parameters to approach the quantities they refer to in the realistic model. However, in previous paragraphs we gave an understanding of these parameters in *our* model, but a realistic model only locally verifies the assumptions used to derive our structure. In particular, the relation between P and ρ is described by $\gamma \equiv \frac{d \ln P}{d \ln \rho}$ rather than Γ_1 , the two differing outside the isentropic region. Accordingly, we expect the adjusted values to approximate the CESTAM model quantities where assumptions 1 to 4 are verified, that is, in the CZ. For the helium abundance Y_s , because $Y_s = Y(t)$ at any position of our model, the value Y_s° should approximate $Y^{\text{CESTAM}}(t)$ in the whole CZ up to the atmosphere. In the case of ψ_{CZ} , we saw that the approximation remains valid as long as $\Gamma_1 = 5/3$, therefore we expect ψ_{CZ}° to approach $\psi^{\text{CESTAM}}(t)$ in the CZ part that is located below the ionisation region. As mentioned, because ε does not relate to any particular quantity of the CZ, it is difficult to deduce more than the position of the ionisation region. These approximations can be summarised as follows:

$$Y_s^\circ \simeq Y^{\text{CESTAM}}(t) \quad \text{for } \text{BCZ} < t < \text{ATM} \quad (71)$$

$$\psi_{CZ}^\circ \simeq \psi^{\text{CESTAM}}(t) \quad \text{for } \text{BCZ} < t < \text{ION}, \quad (72)$$

with BCZ, ATM, and ION designating the base of the convective zone, the base of the atmosphere, and the end of the ionisation region, respectively. We illustrate this point in panels b and c of Fig. 7 by representing both $Y^{\text{CESTAM}}(t)$ and $\psi^{\text{CESTAM}}(t)$, as well as the estimates Y_s° and ψ_{CZ}° obtained through the minimisation of Eq. (70). We also highlight the regions in which the quantities are expected to correspond to each other based on Eqs. (71) and (72). The values thus obtained $(Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ) = (0.255, -5, 0.015)$ are consistent with the CESTAM quantities in the associated regions (see panels b and c of Fig. 7 and Table 1).

It is risky to be quantitative about this result because there are no obvious values with which the difference can be compared; in particular the relative difference does not appear to be relevant. Nevertheless, this example strongly suggests that the modelling defined here allows us to define structures that are close to realistic models (in the sense of criteria (70)) and to also obtain similar helium abundances and electron degeneracies in the regions of interest.

5. Analysis of first adiabatic exponent perturbations

So far, we focused on introducing a physical model of the ionisation region that depends on only a few parameters in order to study its properties. However, Eq. (5) shows that the analysis of frequency shifts relies more on the modelling of structural perturbations than on the structure itself. We therefore investigate what our model predicts as a perturbation caused by a change in surface helium abundance compared with the ad hoc profiles used in previous papers (see Fig. 2). The advantage of having modelled the structure is that we can easily reduce it to the study of a perturbation of helium abundance by considering the following difference between profiles:

$$\frac{\delta_t \Gamma_1}{\Gamma_1} = \frac{\Gamma_1(t; Y_s + \delta Y_s, \psi_{CZ}, \varepsilon) - \Gamma_1(t; Y_s, \psi_{CZ}, \varepsilon)}{\Gamma_1} \quad (73)$$

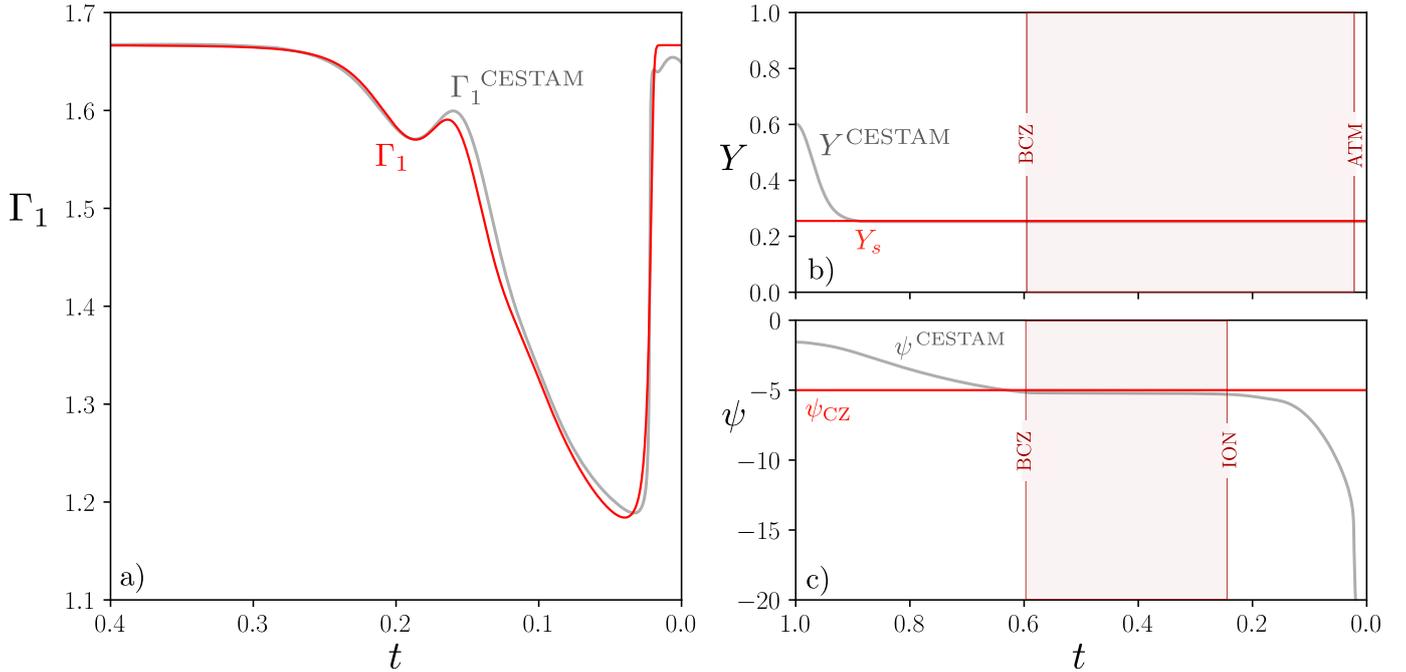


Fig. 7. Correspondence between the CESTAM profiles and their respective analogues in our model. (a) Comparison of $\Gamma_1^{\text{CESTAM}}(t)$ extracted from a CESTAM solar model (in grey) and $\Gamma_1(t; Y_s^\circ, \psi_{\text{CZ}}^\circ, \varepsilon^\circ)$ obtained from our model (in red). (b) Comparison of $Y^{\text{CESTAM}}(t)$ extracted from the same model with Y_s° . (c) Comparison of $\psi^{\text{CESTAM}}(t)$ with ψ_{CZ}° . In the last two panels, the regions in which the curves are expected to correspond are shaded in red.

Table 1. Inferred electron degeneracy parameter and helium abundance and actual CESTAM quantities.

Quantity	Inferred value	CESTAM
Y	0.255	0.253
ψ	$-5.00 - 0.13 = -5.13$ (*)	-5.19

Notes. (*) Our ψ estimate results from the combination $\psi_{\text{CZ}} - Y_s/2$ in accordance with Eq. (69).

The analysis of this perturbation is carried out in Fig. 8. An important aspect to bear in mind is the number of variables on which the above function depends. The perturbation naturally depends on the helium difference δY_s but also on the point $(Y_s, \psi_{\text{CZ}}, \varepsilon)$ around which the differences are calculated, thus leading to four dimensions to explore. Our analysis of the 3D space required for the structure was obviously already incomplete in Fig. 6, and it follows that the perturbation analysis is necessarily even more superficial. Figure 8 provides plots for various reference points, $(Y_s, \psi_{\text{CZ}}, \varepsilon)$, but for a fixed difference $\delta Y_s = 0.05$. Nevertheless, it is straightforward to see by linearising Eq. (73) that a difference caused by $\delta Y_s' = \alpha \times \delta Y_s$ would be relatively similar to the profile $\alpha \times \delta_t \Gamma_1 / \Gamma_1$. In this sense, Fig. 8 still provides a good representation of what the change in helium abundance might cause. Panel a shows that all the perturbed profiles roughly overlap, which reflects a relative independence of the helium amount chosen as a reference (the range considered, $0.2 < Y_s < 0.35$, is representative of most realistic stellar helium abundances). The shape is quite similar to what is expected from a helium variation: Each helium ionisation region contributes to a Gaussian-like well ($t \sim 0.13$ & $t \sim 0.19$), and the second well is more pronounced. The hydrogen contribution is less intuitive, resulting in a very localised peak ($t \sim 0.03$, present in every panel) at the beginning of ionisation. This peak reflects more of

a shift of the hydrogen ionisation region (visible in panel a of Fig. 6) caused by the helium change than a real drop in the well. Nevertheless, this component is clearly noticeable in terms of amplitude.

The situation in panel b is less intuitive. In order to provide realistic values for this rather uncommon parameter, we considered $-9 < \psi_{\text{CZ}} < -3$ as the electron degeneracy range, which corresponds to typical values inside solar-like oscillators (a justification of this point is provided later on). This time, the shape of the perturbation caused by a fixed change in the helium amount is very sensitive to the electron degeneracy value taken as a reference, and this highlights how counter-intuitive differences of profiles can be. At higher electron degeneracy levels, the perturbation looks like a more spread-out version of the perturbation visible in panel a, but its behaviour becomes more complex as the degeneracy decreases. Both helium wells lose their Gaussian aspect: the first well approaches a “heartbeat” shape, that is, it becomes clearly positive (probably under the influence of the hydrogen well), while the second well becomes progressively bimodal. In contrast, the results in panel c could have been guessed from Fig. 6. The ε value shifts both reference and perturbed profiles, thus resulting in a scaled version of the perturbation shown in panel a.

As a consequence, Fig. 8 shows the issues inherent to calibration methods that we mentioned in Sect. 2. Even if it can be established that the choice of the helium amount Y_s taken as reference does not ultimately matter greatly, Fig. 8 illustrates the diversity of $\delta_t \Gamma_1 / \Gamma_1$ profiles corresponding to the same helium difference by simply considering different values of ψ_{CZ} and ε . Reusing the notations introduced in Sect. 2, we clearly find that any change in a component of θ_\star that may impact the best-fitting pair $(\psi_{\text{CZ}}, \varepsilon) \subset \theta_p$ will lead to substantially different estimates of δY . In this respect, a parameter with a particularly strong impact is given in the forthcoming paragraphs, thus further distancing us from a unique relation of the form $\theta_p(\delta Y)$ as assumed in Eq. (13).

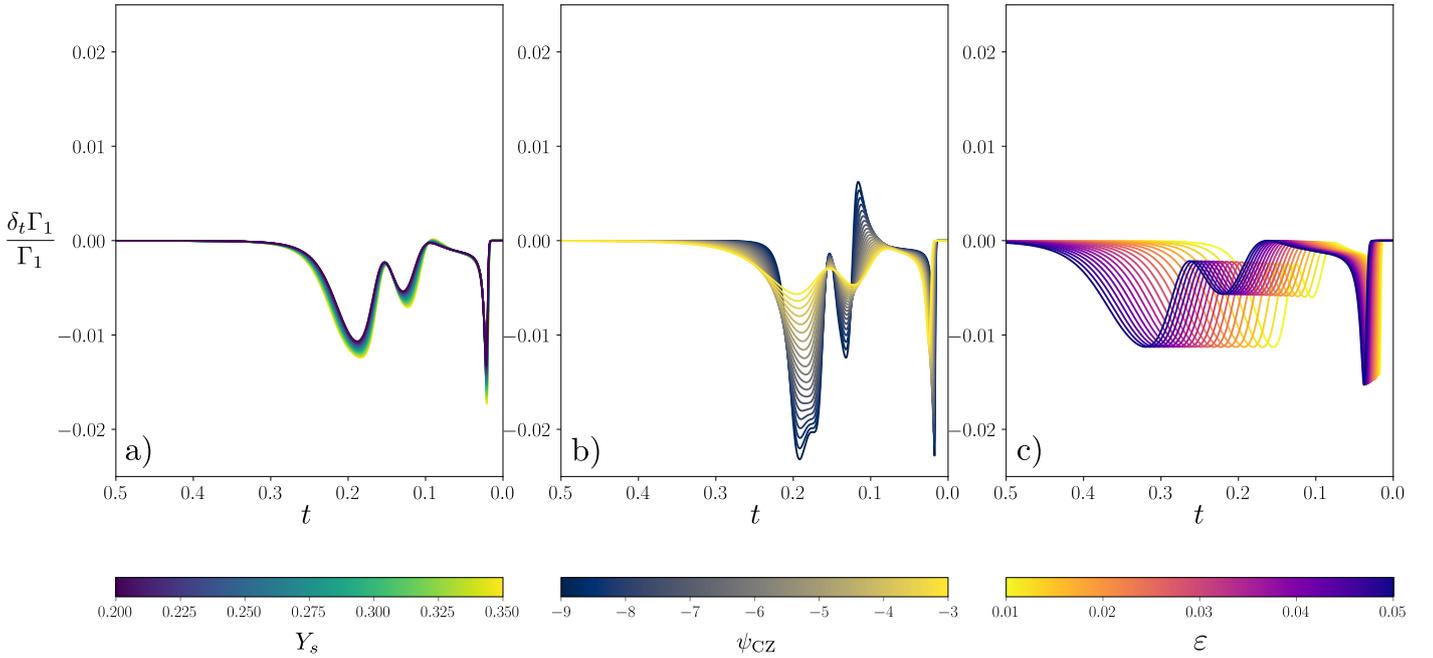


Fig. 8. Profile differences $\delta_t \Gamma_1 / \Gamma_1 = [\Gamma_1(t; Y_s + \delta Y_s, \psi_{CZ}, \varepsilon) - \Gamma_1(t; Y_s, \psi_{CZ}, \varepsilon)] / \Gamma_1$ obtained for a fixed value of $\delta Y_s = 0.05$ and various sets $(Y_s, \psi_{CZ}, \varepsilon)$ (the values used here are not the same as those in Fig. 6) around the reference $(Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ) = (0.255, -5, 0.015)$ as a function of the normalised acoustic depth t . (a) Dependence on $0.2 \leq Y_s \leq 0.35$. (b) Dependence on $-9 \leq \psi_{CZ} \leq -3$. (c) Dependence on $0.01 \leq \varepsilon \leq 0.05$.

Furthermore, it should be added that the particular case of a perturbation around the reference value $Y_s = 0$ (as considered in HG07) is no exception. Although we only varied the two remaining parameters ψ_{CZ} and ε , Fig. 9 shows the many possibilities for $\delta_t \Gamma_1 / \Gamma_1$ even when an identical helium difference $\delta Y_s = 0.25$ is considered from a pure hydrogen model $Y_s = 0$. The spread of these randomly drawn curves provides an idea of the potential dispersion of the perturbed profile at any point of the structure even under these “favourable” conditions.

Secondly, Figs. 8 and 9 underline how accustomed we are to conceive the structural perturbations caused by a variation in helium under solar conditions of electron degeneracy. We may note here the added value of having introduced a model physically; it would have been complicated to imagine and then parameterise these types of perturbations by simply observing a Γ_1 profile in a realistic model. Accordingly, we may question the validity of using ad hoc profiles to approximate the perturbation caused by a change in helium abundance, and this particularly at low electron degeneracy (whose domain of relevance is discussed in the following). It appears to be likely that adjusting functions whose form is inappropriate for these complex profiles may result in inconsistencies regarding their parameterisations.

Up to this point, we only considered a given helium difference, in particular, to determine how the model presented in this paper agrees or contrasts with the types of perturbations for ionisation regions from previous studies. However, Eq. (73) clearly extends to the much more general difference

$$\frac{\delta_t \Gamma_1}{\Gamma_1} = \frac{\Gamma_1(t; Y_s + \delta Y_s, \psi_{CZ} + \delta \psi_{CZ}, \varepsilon + \delta \varepsilon) - \Gamma_1(t; Y_s, \psi_{CZ}, \varepsilon)}{\Gamma_1}. \quad (74)$$

When it applied to a frequency shift analysis, it appears to be unlikely that the model differs from its target in its helium abundance alone. We therefore propose to study the structural impact of a $\delta \psi_{CZ}$ or a $\delta \varepsilon$ difference below. It is obvious that the problems mentioned above in terms of the number of dependences will

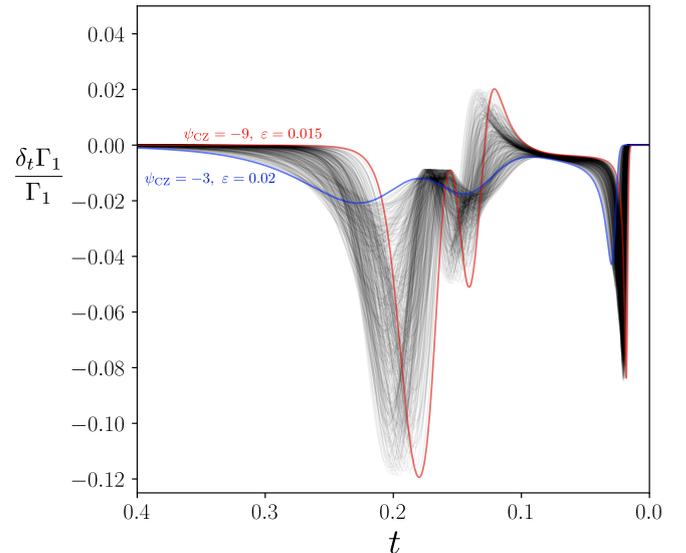


Fig. 9. Profile differences $\delta_t \Gamma_1 / \Gamma_1$ again obtained for a fixed value of $\delta Y_s = 0.25$, but from a pure-hydrogen reference model $Y_s = 0$. The 500 profile differences in this plot were obtained by randomly drawing the remaining two reference quantities (ψ_{CZ}, ε) in the 2D box $[-9, -3] \times [0.015, 0.02]$ (the ε interval has been reduced compared to Fig. 8 for clarity). Two opposite corners of this box are highlighted in red and blue.

only be further amplified. For this reason, we fixed the reference set of parameters by taking for example the values $(Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ)$ and varied the differences δY_s , $\delta \psi_{CZ}$, and $\delta \varepsilon$ one by one. The resulting profiles are shown in Fig. 10.

As expected, the variation observed in panel a is very similar to that in panel a of Fig. 8 scaled by a factor $\alpha = \delta Y_s / 0.05$. The other two perturbations, however, take different shapes and notably make the wells unrecognisable. A change in the electron degeneracy (panel b) has a strong impact on the hydrogen

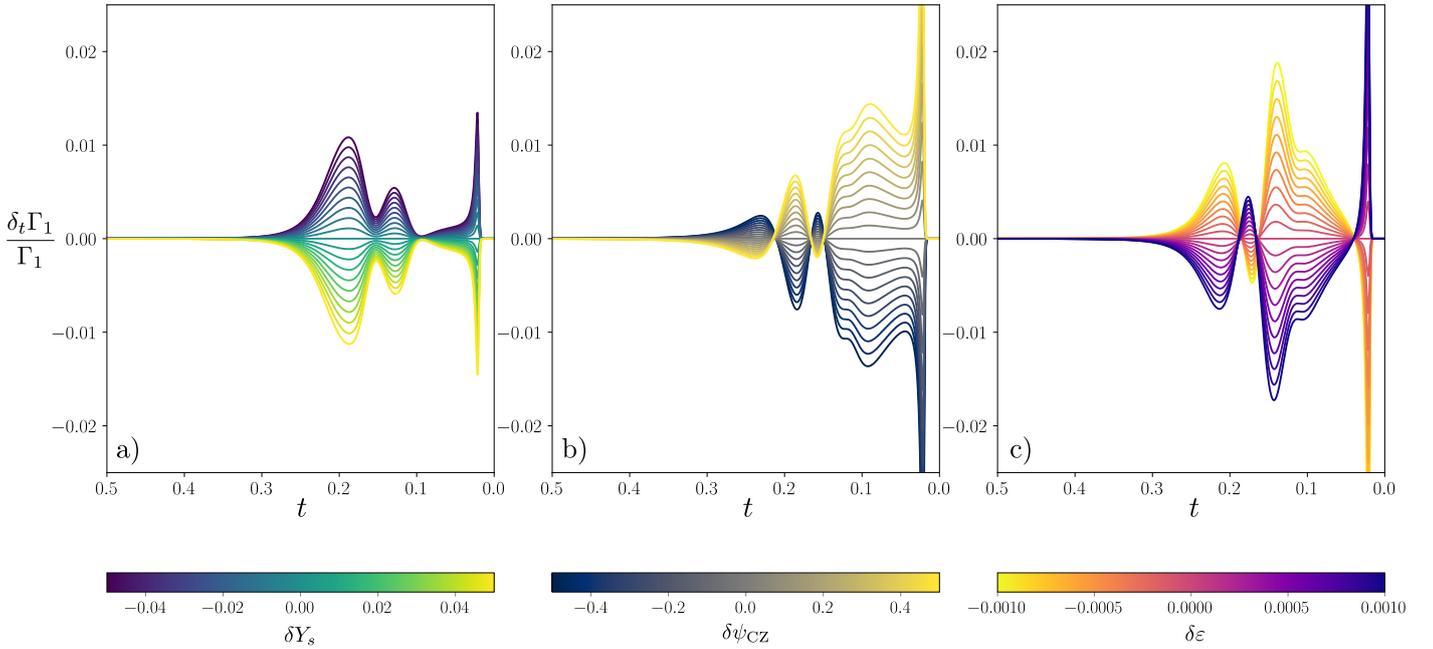


Fig. 10. Profile differences $\delta_t \Gamma_1 / \Gamma_1 = [\Gamma_1(t; Y_s + \delta Y_s, \psi_{CZ} + \delta \psi_{CZ}, \varepsilon + \delta \varepsilon) - \Gamma_1(t; Y_s, \psi_{CZ}, \varepsilon)] / \Gamma_1$ obtained for a fixed reference parameter set $(Y_s, \psi_{CZ}, \varepsilon) = (Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ)$ and by considering various sets $(\delta Y_s, \delta \psi_{CZ}, \delta \varepsilon)$. (a) Dependence on $(\delta Y_s, 0, 0)$ for $-0.05 \leq \delta Y_s \leq 0.05$. (b) Dependence on $(0, \delta \psi_{CZ}, 0)$ for $-0.5 \leq \delta \psi_{CZ} \leq 0.5$. (c) Dependence on $(0, 0, \delta \varepsilon)$ for $-0.001 \leq \delta \varepsilon \leq 0.001$.

ionisation structure, especially at a point ($t \sim 0.09$), where a change in helium has almost no effect. The ionisation region shift is also enhanced, while the second helium ionisation contributes to a weaker and more localised well. The variation in the position $\delta \varepsilon$ of the ionisation region is also fairly counter-intuitive. Panel c illustrates a contribution that reaches its maximum between the two helium wells ($t \sim 0.15$), the one corresponding to the peak in the $\Gamma_1(t)$ profile.

Figure 10 highlights how elaborate the ionisation perturbation profile becomes in the general case. We recall that a structural perturbation from a reference $(Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ)$ can be approximated as a combination of profiles composing Fig. 10. According to Eq. (74),

$$\frac{\delta_t \Gamma_1}{\Gamma_1} \simeq \delta Y_s \left. \frac{\partial \ln \Gamma_1}{\partial Y_s} \right|_{\circ} + \delta \psi_{CZ} \left. \frac{\partial \ln \Gamma_1}{\partial \psi_{CZ}} \right|_{\circ} + \delta \varepsilon \left. \frac{\partial \ln \Gamma_1}{\partial \varepsilon} \right|_{\circ}, \quad (75)$$

with for instance

$$\left. \frac{\partial \ln \Gamma_1}{\partial Y_s} \right|_{\circ} \equiv \lim_{\delta Y_s \rightarrow 0} \frac{\Gamma_1(t; Y_s^\circ + \delta Y_s, \psi_{CZ}^\circ, \varepsilon^\circ) - \Gamma_1(t; Y_s^\circ, \psi_{CZ}^\circ, \varepsilon^\circ)}{\Gamma_1 \delta Y_s} \quad (76)$$

being the structural perturbation studied in panel a within a multiplicative factor. With regard to Fig. 10, the result of the linear combination (75) can become highly complex and is unlikely to have anything in common with the functions generally used to fit $\delta_t \Gamma_1 / \Gamma_1$. Combined with the issues already raised about calibration, this tends to suggest that the study of the ionisation glitch based on ad hoc perturbation profiles in order to calibrate them afterwards might lead to fairly inconsistent results.

Moreover, we have tried to represent structural disturbances in the figure that are similar in order of magnitude. This allows us to qualitatively relate the values of the different parameters. In this sense, perturbations with $\delta Y_s = 0.1$, $\delta \psi_{CZ} = 1$ or $\delta \varepsilon = 0.002$ have a comparable structural impact in amplitude and hence on the frequencies. However, although the first impact nearly

spans the entire range of relevant values (the difference in helium between two stars can hardly exceed 0.1), this is not the case for the two others. In the ε parameter, this change is very small: two consecutive curves in panel c of Fig. 6 are separated by exactly $\delta \varepsilon = 0.002$, which is the entire span for the perturbation represented in Fig. 10. This last point suggests a very high dependence of the perturbation on a shift of the ionisation region. As mentioned above, it is interesting to note that the glitch signature would come from the Γ_1 peak rather than from a well if this contribution were to dominate. This is consistent with observations that have been made in previous studies (Broomhall et al. 2014; Verma et al. 2014b), although it provides an alternative explanation.

To place the significance of a difference in the ψ_{CZ} parameter in perspective, it can be interesting to relate it to more intuitive quantities such as fundamental parameters. Because the electron degeneracy ψ is known to fluctuate with the mass of the star (e.g. Hayashi et al. 1962), we represent in Fig. 11 various $\psi(t)$ profiles (along with the corresponding $\Gamma_1(t)$ profiles) obtained from CESTAM models of different masses. These models cover a wide range of values from $0.7 M_\odot$ to $1.35 M_\odot$ appropriate for solar-like pulsators. All the models also share the same composition (the solar composition) and subsequently the same helium abundance as well as the same evolutionary state ($X_c = 0.6$). Clearly, although their composition is identical, these models show fairly distinct $\Gamma_1(t)$ profiles in the ionisation region. The behaviour of these curves is fairly similar to that of the curves shown in panel b of Fig. 6 (which can be extrapolated to the wider range $0.45 M_\odot < M < 1.5 M_\odot$ presented as dashes as an indication). Thus, at low masses, the wells seem to merge and then split at higher masses, which in particular makes the HeI well visible. In addition to this effect, an expansion and contraction of the profiles is visible depending on the considered mass range ($< 1 M_\odot$ or $> 1 M_\odot$, cf. arrows shown in panels a and b). From examination of the similarities with Fig. 6, we interpret these changes as manifestations of variations in ε and the electron degeneracy just below the ionisation region. Although

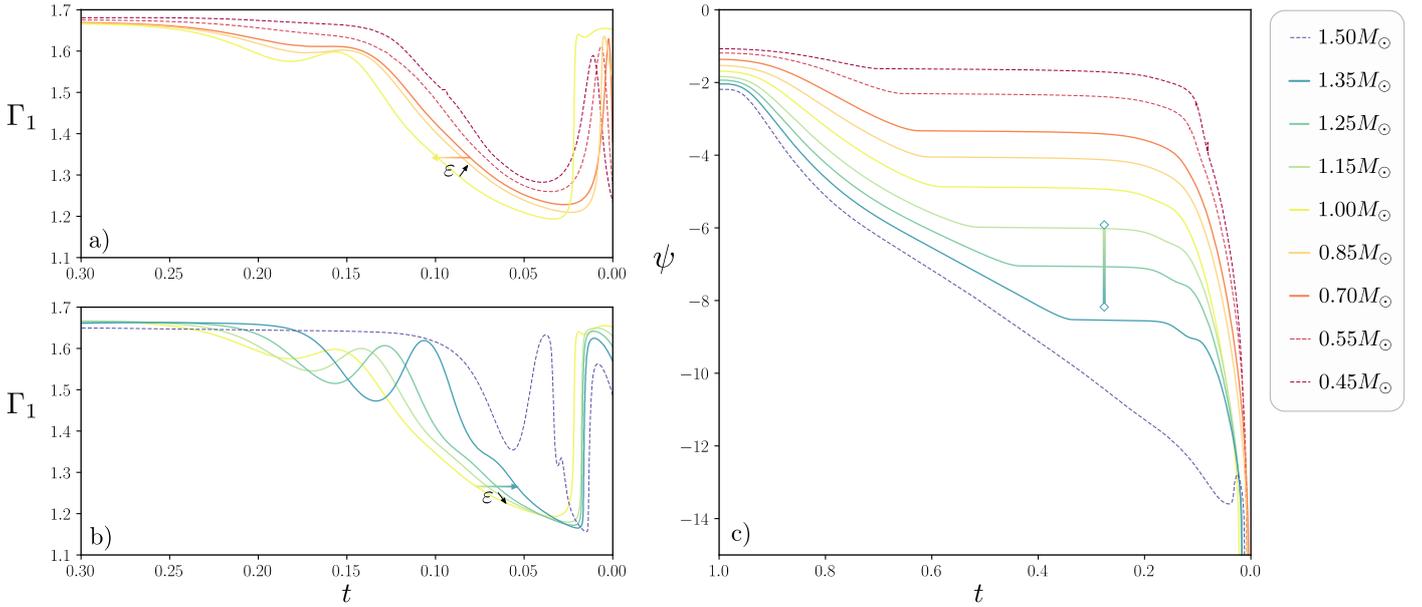


Fig. 11. $\Gamma_1(t)$ (both panels *a* & *b* for clarity) and $\psi(t)$ (panel *c*) profiles extracted from CESTAM early main sequence models with masses ranging from $0.7 M_\odot$ to $1.35 M_\odot$ (and from $0.45 M_\odot$ to $1.5 M_\odot$ as an indication in dashes). All models share the same composition which is the solar one as well as the same evolutionary state: $X_c = 0.6$. The qualitative variations of ε with mass are represented as well as the ψ span corresponding to the mass uncertainty of HD 52265 taken from [Lebreton & Goupil \(2014\)](#).

the variations of ε can only be evaluated qualitatively (we recall that this quantity is only defined in entirely convective models), this is not the case for the electron degeneracy. The $\psi(t)$ profiles typically undergo little to no variation in convection zones, as shown by the plateaus in the curves in panel *c*, thus making it easy to identify ψ_{CZ} . The range of ψ_{CZ} values covered by these plateaus corresponds to the range studied in Fig. 8 (~ -3 for $0.7 M_\odot$ stars and ~ -9 for the $1.35 M_\odot$ model). It is then possible to seek to characterise the significance of the change $\delta\psi_{CZ} = 1$ mentioned above. To provide an order of magnitude, we represent in panel *c* the ψ span corresponding to the mass uncertainty ($1.14 < M_* < 1.32$) of HD 52265 taken from [Lebreton & Goupil \(2014\)](#). This star provides an example of a seismically determined mass uncertainty, taking into account a large number of biases as discussed in Sect. 1. The corresponding extent in terms of electron degeneracy is about $\delta\psi_{CZ} \sim 2$, which is twice larger than the range considered in panel *b* of Fig. 10 and thus larger than any possible variation in helium abundance. This range in ψ_{CZ} depends on the mass; this mass difference would correspond to a lower $\delta\psi_{CZ}$ at $0.7 M_\odot$ and would instead increase at higher masses. In this regard, a mass difference between two models could contribute to a perturbation of the structure of the ionisation region as large as (if not larger than) that of helium. Here, we note that a change in mass in a given evolutionary state has been studied. The impact of an age difference at a given mass has also been examined, which leads to smaller changes in both ψ_{CZ} and ε under the constraints provided by [Lebreton & Goupil \(2014\)](#). The only substantial variation in the Γ_1 profile occurs in the $1.5 M_\odot$ model, with the appearance of a convective surface zone for $X_c \leq 0.4$.

Finally, to further echo the point made in Fig. 8 about calibration, the mass taken as a reference, M_* , provides a representative example of a component of θ_* that impacts the (ψ_{CZ}, ε) pair, as already observed by [Verma et al. \(2014b\)](#). In this article, the authors proposed to use it by considering different calibrations to determine the helium abundance of stars of distinct masses. However, besides the fact that this presupposes a relatively good

knowledge of the star in question before its helium abundance can be estimated, and we also saw that stars whose masses are already seismically constrained can possess significantly different degeneracy levels in the convective region. Using the example of HD 52265, the question then arises whether it is possible to obtain consistent results using calibration and being able to consider both models with $\psi_{CZ} \sim 6$ or $\psi_{CZ} \sim 8$, which implies perturbation profiles that differ in form and magnitude for a fixed difference in helium amount (see panel *b* of Fig. 8). This statement does not apply exclusively to the mass, M_* , but can reasonably be extended to any fundamental parameters or physical processes (i.e., the other components of θ_*) that may affect the average density in the convection zone (related to ψ_{CZ}) or the relative acoustic depth of the ionisation region (linked to ε).

In addition to this last remark, it appears likely by comparing the structural perturbations in Fig. 10 that the component caused by a helium change between two stars does not contribute alone. It might even be dominated by a change in electron degeneracy or in the ionisation region position. From the perspective of frequency shifts, it thus seems incorrect to assimilate the ionisation glitch as a repercussion of a helium difference only in order to infer its value. In this light, it appears essential to study forms of frequency shifts that can take these additional dependences into account, thus relying on physical models as proposed in this paper.

6. Conclusion

Determining abundances can be subject to many biases in its classical approach, which we intend to overcome by exploiting the ionisation glitch. However, although it progressively becomes more sophisticated, the glitch-based approach faces problems inherent to its current modelling, such as the need for calibration by realistic stellar models.

To address these problems, a physical model of the ionisation region was proposed here, explicitly involving the parameters of interest, such as the surface helium abundance Y_s , and which

can be generalised to more elaborate compositions. In the case of a hydrogen-helium mixture, the model involves three parameters and highlights the importance of characterising the electron degeneracy state in the convective zone ψ_{CZ} as well as the position of the ionisation region here controlled by ε . While it is well known that the abundance of surface helium contributes greatly to the appearance of the first adiabatic exponent (Γ_1) profile by shaping the size of the helium wells, the state of degeneracy seems to affect the profile just as much by altering their dispersion. Taking this into consideration, the model is thus able to describe a wide variety of ionisation structures while providing them with a physical meaning. By comparing them to a CES-TAM model, we also verified that it was able to approach realistic Γ_1 profiles for consistent parameter values, which confers a predictive capacity to the model.

The modelling work we conducted allowed us to study the shape of structural perturbations by analysing differences of profile with distinct parameterisations. In particular, the form of the perturbations caused by a helium difference was addressed. Expected shapes were found, such as a Gaussian-like contribution for the helium ionisation regions with an additional sharper component at the surface, caused by a shift in the hydrogen ionisation region. However, we have observed that this form is only valid for a restricted case, namely a helium difference under solar electron degeneracy conditions. This perturbation appears to be highly variable with respect to these conditions, however, and can easily become much more complex at lower electron degeneracies. This point in particular stresses the major dependence of the perturbation on the choice of realistic models chosen for calibration purposes. We then proceeded to study more general parameter differences. More elaborate forms of perturbations than those usually assumed are found. We also suggested that there is a stronger dependence on the electron degeneracy in the CZ or on the position of the ionisation region than on the helium amount itself. Furthermore, the connection between electron degeneracy and stellar mass therefore enables us to clarify the strong dependence of the helium glitch amplitude on the stellar mass already observed by Verma et al. (2014b, 2019). Moreover, the fluctuation in the ionisation region position thus induced seems compatible with a glitch signature coming from the peak of the first adiabatic exponent as reported by Broomhall et al. (2014) or Verma et al. (2014b).

When analysing the oscillation frequencies, we therefore emphasise the importance of having a relation that can take these additional dependences into account. In this sense, a second paper based on the introduced modelling is planned in order to derive more general forms of frequency shifts. The objective should be to interpret the ionisation glitch as a combination of multiple contributions for which a variation of helium abundance is one component among others.

References

- Antia, H. M., & Basu, S. 1994, *A&AS*, 107, 421
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, 47, 481
- Baglin, A., Auvergne, M., Barge, P., et al. 2006, in *The CoRoT Mission Pre-Launch Status - Stellar Seismology and Planet Finding*, eds. M. Fridlund, A. Baglin, J. Lochard, & L. Conroy, *ESA SP*, 1306, 33
- Baker, N., & Kippenhahn, R. 1962, *Z. Astrophys.*, 54, 114
- Ball, W. H., & Gizon, L. 2014, *A&A*, 568, A123
- Ballot, J., Turck-Chièze, S., & García, R. A. 2004, *A&A*, 423, 1051
- Basu, S., & Antia, H. M. 2004, *ApJ*, 606, L85
- Basu, S., Antia, H. M., & Narasimha, D. 1994, *MNRAS*, 267, 209
- Belkacem, K., Kupka, F., Samadi, R., & Grimm-Strele, H. 2019, *A&A*, 625, A20
- Belkacem, K., Kupka, F., Philidet, J., & Samadi, R. 2021, *A&A*, 646, L5
- Broomhall, A. M., Miglio, A., Montalbán, J., et al. 2014, *MNRAS*, 440, 1828
- Canuto, V. M. 1997, *ApJ*, 482, 827
- Christensen-Dalsgaard, J., Dappen, W., & Lebreton, Y. 1988, *Nature*, 336, 634
- Christensen-Dalsgaard, J., Proffitt, C. R., & Thompson, M. J. 1993, *ApJ*, 403, L75
- Christensen-Dalsgaard, J., & Thompson, M. J. 1997, *MNRAS*, 284, 527
- Cox, J. P., & Giuli, R. T. 1968, *Principles of Stellar Structure*
- Deal, M., Alecian, G., Lebreton, Y., et al. 2018, *A&A*, 618, A10
- Dziembowski, W. A., Pamyatnykh, A. A., & Sienkiewicz, R. 1990, *MNRAS*, 244, 542
- Farnir, M., Dupret, M. A., Salmon, S. J. A. J., Noels, A., & Buldgen, G. 2019, *A&A*, 622, A98
- Gilliland, R. L., Brown, T. M., Christensen-Dalsgaard, J., et al. 2010, *PASP*, 122, 131
- Gough, D. O. 1990, in *Progress of Seismology of the Sun and Stars, Proceedings of the Oji International Seminar Held at Hakone, Japan, 11–14 December 1989*, eds. Y. Osaki, & H. Shibahashi (Berlin Heidelberg New York: Springer-Verlag), *Lecture Notes Phys.*, 367, 283
- Gough, D. O. 2002, in *Stellar Structure and Habitable Planet Finding*, eds. B. Battrick, F. Favata, I. W. Roxburgh, & D. Galadi, *ESA SP*, 485, 65
- Gough, D. O., & Thompson, M. J. 1991, *The Inversion Problem* (Tucson: University of Arizona Press), 519
- Hayashi, C., Hōshi, R., & Sugimoto, D. 1962, *Prog. Theor. Phys. Suppl.*, 22, 1
- Houdek, G., & Gough, D. O. 2007, *MNRAS*, 375, 861
- Houdek, G., & Gough, D. O. 2011, *MNRAS*, 418, 1217
- Kippenhahn, R., Weigert, A., & Weiss, A. 2012, *Stellar Structure and Evolution, 2nd Edition* (Berlin Heidelberg: Springer-Verlag)
- Kjeldsen, H., Bedding, T. R., & Christensen-Dalsgaard, J. 2008, *ApJ*, 683, L175
- Kosovichev, A. G., Christensen-Dalsgaard, J., Däppen, W., et al. 1992, *MNRAS*, 259, 536
- Lebreton, Y., & Goupil, M. J. 2014, *A&A*, 569, A21
- Lopes, I., Turck-Chièze, S., Michel, E., & Goupil, M.-J. 1997, *ApJ*, 480, 794
- Lund, M. N., Silva Aguirre, V., Davies, G. R., et al. 2017, *ApJ*, 835, 172
- Marques, J. P., Goupil, M. J., Lebreton, Y., et al. 2013, *A&A*, 549, A74
- Monteiro, M. J. P. F. G., & Thompson, M. J. 1998, in *New Eyes to See Inside the Sun and Stars*, eds. F. L. Deubner, J. Christensen-Dalsgaard, & D. Kurtz, *IAU Symp.*, 185, 317
- Monteiro, M. J. P. F. G., & Thompson, M. J. 2005, *MNRAS*, 361, 1187
- Monteiro, M. J. P. F. G., Christensen-Dalsgaard, J., & Thompson, M. J. 1994, *A&A*, 283, 247
- Morel, P., & Lebreton, Y. 2008, *Ap&SS*, 316, 61
- Mosumgaard, J. R., Jørgensen, A. C. S., Weiss, A., Silva Aguirre, V., & Christensen-Dalsgaard, J. 2020, *MNRAS*, 491, 1160
- Noll, A., Deheuvels, S., & Ballot, J. 2021, *A&A*, 647, A187
- Nsamba, B., Campante, T. L., Monteiro, M. J. P. F. G., et al. 2018, *MNRAS*, 477, 5052
- Perez Hernandez, F., & Christensen-Dalsgaard, J. 1994, *MNRAS*, 269, 475
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *J. Astron. Telescopes Instruments Systems*, 1, 014003
- Rogers, F. J. 1981, *Phys. Rev. A*, 24, 1531
- Rogers, F. J., Swenson, F. J., & Iglesias, C. A. 1996, *ApJ*, 456, 902
- Schou, J., & Birch, A. C. 2020, *A&A*, 638, A51
- Serenelli, A. M., Basu, S., Ferguson, J. W., & Asplund, M. 2009, *ApJ*, 705, L123
- Sonoi, T., Samadi, R., Belkacem, K., et al. 2015, *A&A*, 583, A112
- Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, *AJ*, 158, 138
- Tassoul, M. 1980, *ApJS*, 43, 469
- Verma, K., Faria, J. P., Antia, H. M., et al. 2014a, *ApJ*, 790, 138
- Verma, K., Antia, H. M., Basu, S., & Mazumdar, A. 2014b, *ApJ*, 794, 114
- Verma, K., Raodeo, K., Antia, H. M., et al. 2017, *ApJ*, 837, 47
- Verma, K., Raodeo, K., Basu, S., et al. 2019, *MNRAS*, 483, 4678

Appendix A: Further simplifications and perturbation of the first adiabatic exponent

A.1. Deriving Eq. (53)

This part details the derivation of (53) and (54) from (48)-(52). We first introduce the notation

$$\chi_n \equiv \sum_{ir} x_i y_i^r (1 - y_i^r) \frac{(\chi_i^r/kT)^n}{1 + \delta_i^1(1 - y_i^r)}. \quad (\text{A.1})$$

Equations (50)-(52) can then be rewritten as

$$\partial_{VV}^2 f = \partial_v f - \chi_0 \quad (\text{A.2})$$

$$\partial_{TV}^2 f = \partial_{VT}^2 f = \partial_v f + \chi_1 + \frac{3}{2}\chi_0 \quad (\text{A.3})$$

$$\partial_{TT}^2 f = \frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0. \quad (\text{A.4})$$

Injecting this into Eq. (48) then leads to

$$\Gamma_1 = \frac{(\partial_v f + \chi_1 + \frac{3}{2}\chi_0)^2 + (\partial_v f - \chi_0)(\frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0)}{\partial_v f (\frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0)}, \quad (\text{A.5})$$

and since

$$\begin{aligned} \left(\partial_v f + \chi_1 + \frac{3}{2}\chi_0\right)^2 &= \partial_v f (\partial_v f + 2\chi_1 + 3\chi_0) + \left(\chi_1 + \frac{3}{2}\chi_0\right)^2 \\ &= \frac{2}{3}\partial_v f \left(\frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0\right) \\ &\quad + \partial_v f \left(-\frac{2}{3}\chi_2 + \frac{3}{2}\chi_0\right) + \left(\chi_1 + \frac{3}{2}\chi_0\right)^2 \end{aligned} \quad (\text{A.6})$$

we obtain the following expression, in which 5/3 appears, as expected:

$$\begin{aligned} \Gamma_1 &= \frac{5}{3} - \frac{\partial_v f (\frac{2}{3}\chi_2 - \frac{3}{2}\chi_0) - (\chi_1 + \frac{3}{2}\chi_0)^2}{\partial_v f (\frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0)} \\ &\quad + \frac{\chi_0 (\frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0)}{\partial_v f (\frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0)}. \end{aligned} \quad (\text{A.7})$$

The simplifications in the numerator lead to the following more compact expression:

$$\Gamma_1 = \frac{5}{3} - \frac{\partial_v f (\frac{2}{3}\chi_2) + \chi_2 \chi_0 - (\chi_1)^2}{\partial_v f (\frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0)}. \quad (\text{A.8})$$

We study the term $\chi_2 \chi_0 - (\chi_1)^2$,

$$\begin{aligned} \chi_2 \chi_0 - (\chi_1)^2 &= \sum_{ijrs} x_i x_j y_i^r y_j^s \frac{(1 - y_i^r)(1 - y_j^s) \chi_i^r (\chi_i^r - \chi_j^s)/(kT)^2}{[1 + \delta_i^1(1 - y_i^r)][1 + \delta_j^1(1 - y_j^s)]} \\ &= \sum_{ir < js} x_i x_j y_i^r y_j^s \frac{(1 - y_i^r)(1 - y_j^s) (\chi_i^r - \chi_j^s)^2 / (kT)^2}{[1 + \delta_i^1(1 - y_i^r)][1 + \delta_j^1(1 - y_j^s)]} \\ &< \sum_{ir < js} x_i x_j y_i^r y_j^s (1 - y_i^r)(1 - y_j^s) \left(\frac{\chi_i^r - \chi_j^s}{kT}\right)^2, \end{aligned} \quad (\text{A.9})$$

where we grouped the symmetrical terms under the notation $ir < js$. Most of the terms in this sum are negligible compared with the denominator, which is always greater than 3/2. Several effects contribute to keeping these terms small:

1. In order for the product $x_i x_j$ to contribute to the total, the number abundances of the two elements, i and j , needs to be non-negligible. Accordingly, the only non-negligible terms are those coupled to hydrogen.
2. The ionisation regions (i, r) and (j, s) should overlap as much as possible because otherwise $y_i^r y_j^s (1 - y_i^r)(1 - y_j^s)$ would be null. Even in the most favourable case, we have the upper limit $y_i^r y_j^s (1 - y_i^r)(1 - y_j^s) < 1/16$.
3. At the same time, ionisation regions that overlap tend to have a similar ionisation potential, thus making $(\chi_i^r - \chi_j^s)^2$ small.
4. As discussed in Section 3, ionisation regions are more likely to merge at higher conditions of temperature and density. However, in this case, $1/(kT)^2$ tends rapidly towards 0.

All these points appear to be arguing in favour of a small $\chi_2 \chi_0 - (\chi_1)^2$, and we confirm numerically that this term remains below 2×10^{-3} at any position of the (ρ, T) plane. We thus considered that $\chi_2 \chi_0 - (\chi_1)^2 = 0$ for the remainder of our derivation. Eq. (53) quickly ensues by introducing

$$\gamma_1 \equiv \frac{\chi_2}{\frac{3}{2}\partial_v f + \chi_2 + 3\chi_1 + \frac{9}{4}\chi_0} = \frac{\chi_2}{\partial_{TT}^2 f}. \quad (\text{A.10})$$

Relation (54) is simply obtained by replacing χ_2 by its definition (A.1). Eq. (A.4) shows $\partial_{TT}^2 f > \chi_2$, and this results in the property stated in the main part, namely $0 \leq \gamma_1 < 1$, or equivalently, $1 < \Gamma_1 \leq 5/3$.

A.2. Perturbing Γ_1 at given V, T values

The resulting relation, Eq. (A.10), is so compact that we can perturb it analytically in terms of abundances (which was not an option when considering Eq. (48)). We consider a change δx_i of the x_i at given V and T values (with the implicit condition $\sum_i \delta x_i = 0$). Considering Eqs. (32)-(34), the y_i^r thus defined are functions of T and V , but not of x_i . It can thus be established that $\delta_{V,T}(y_i^r) = 0$. The induced perturbations of χ_2 and $\partial_{TT}^2 f$ is therefore given by

$$\delta_{V,T}(\chi_2) = \sum_{ir} \frac{\delta x_i}{x_i} x_i y_i^r (1 - y_i^r) \frac{(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \quad (\text{A.11})$$

$$\delta_{V,T}(\partial_{TT}^2 f) = \sum_{ir} \frac{\delta x_i}{x_i} x_i y_i^r \left(\frac{3}{2} + (1 - y_i^r) \frac{(3/2 + \chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \right), \quad (\text{A.12})$$

thus leading to the following impact on γ_1 :

$$\delta_{V,T}(\gamma_1) = \frac{(\partial_{TT}^2 f) \delta_{V,T}(\chi_2) - (\chi_2) \delta_{V,T}(\partial_{TT}^2 f)}{(\partial_{TT}^2 f)^2}. \quad (\text{A.13})$$

By calculating the numerator explicitly, four sums appear,

$$\begin{aligned} (\partial_{TT}^2 f)^2 \delta_{V,T}(\gamma_1) &= \frac{3}{2} \sum_{ir} \frac{\delta x_i}{x_i} x_i y_i^r (1 - y_i^r) \frac{(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \\ &\quad + \frac{3}{2} \sum_{ijrs} \left(\frac{\delta x_i}{x_i} - \frac{\delta x_j}{x_j} \right) x_i x_j y_i^r y_j^s \frac{(1 - y_i^r)(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \\ &\quad + \frac{9}{4} \sum_{ijrs} \frac{\delta x_i}{x_i} x_i x_j y_i^r y_j^s \frac{(1 - y_i^r)(1 - y_j^s)(\chi_i^r + \chi_j^s)(\chi_i^r - \chi_j^s)/(kT)^2}{[1 + \delta_i^1(1 - y_i^r)][1 + \delta_j^1(1 - y_j^s)]} \\ &\quad + 3 \sum_{ijrs} \frac{\delta x_i}{x_i} x_i x_j y_i^r y_j^s \frac{(1 - y_i^r)(1 - y_j^s) \chi_i^r \chi_j^s (\chi_i^r - \chi_j^s)/(kT)^3}{[1 + \delta_i^1(1 - y_i^r)][1 + \delta_j^1(1 - y_j^s)]} \end{aligned} \quad (\text{A.14})$$

If we group symmetrical terms, the last two sums are bounded by

$$3^{\text{rd}} \text{ sum} < \frac{9}{4} \sum_{ir < js} \left(\frac{\delta x_i}{x_i} - \frac{\delta x_j}{x_j} \right) x_i x_j y_i^r y_j^s (1 - y_i^r) (1 - y_j^s) \times \left(\frac{\chi_i^r + \chi_j^s}{kT} \right) \left(\frac{\chi_i^r - \chi_j^s}{kT} \right) \quad (\text{A.15})$$

and

$$4^{\text{th}} \text{ sum} < 3 \sum_{ir < js} \left(\frac{\delta x_i}{x_i} - \frac{\delta x_j}{x_j} \right) x_i x_j y_i^r y_j^s (1 - y_i^r) (1 - y_j^s) \times \left(\frac{\chi_i^r}{kT} \right) \left(\frac{\chi_j^s}{kT} \right) \left(\frac{\chi_i^r - \chi_j^s}{kT} \right). \quad (\text{A.16})$$

For the same reasons as mentioned above, we consider that these two terms are negligible in the following. Because $\delta_{v,T}(\Gamma_1) = -2/3 \delta_{v,T}(\gamma_1)$, an analytical expression of the perturbation is given by

$$\delta_{v,T}(\Gamma_1) \simeq -\frac{1}{(\partial_{TT}^2 f)^2} \sum_{ir} x_i y_i^r \frac{(1 - y_i^r)(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \left[\frac{\delta x_i}{x_i} + \sum_{js} x_j y_j^s \left(\frac{\delta x_i}{x_i} - \frac{\delta x_j}{x_j} \right) \right]. \quad (\text{A.17})$$

A.3. Perturbing Γ_1 at given ρ, T values

The perturbation at given ρ and T values is more complex because $\delta_{\rho,T} y_i^r \neq 0$ because of the variable change $\rho/m_0 = N/V$. For any quantity α , the resulting perturbation change can be found by applying

$$\begin{aligned} \delta_{\rho,T}(\alpha) &= \delta_{v,T}(\alpha) - \delta_{v,T}(\rho) \left| \frac{\partial(\alpha, T)}{\partial(\rho, T)} \right| \\ &= \delta_{v,T}(\alpha) - \frac{\delta m_0}{m_0} \left(\frac{\partial \alpha}{\partial \ln \rho} \right)_T. \end{aligned} \quad (\text{A.18})$$

The perturbation of any quantity expressed in terms of y_i^r can be found using $\frac{\partial \ln y_i^r}{\partial \ln \rho} = \frac{d \ln y_i^r}{d \ln K_i^r} \frac{\partial \ln K_i^r}{\partial \ln \rho} = -\frac{1 - y_i^r}{1 + \delta_i^1(1 - y_i^r)}$.

To simplify the derivation, we present only the parts that do not result in terms containing $x_i x_j y_i^r y_j^s (1 - y_i^r) (1 - y_j^s)$, which are neglected,

$$\begin{aligned} \delta_{\rho,T}(\chi_2) &= \delta_{v,T}(\chi_2) + \frac{\delta m_0}{m_0} \sum_{ir} x_i y_i^r \left(\frac{(1 - y_i^r)(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \right) \\ &\quad \left(1 - \frac{y_i^r}{1 - y_i^r} + \frac{\delta_i^1 y_i^r}{1 + \delta_i^1(1 - y_i^r)} \right). \end{aligned} \quad (\text{A.19})$$

The resulting perturbation of the numerator is approximated by

$$\begin{aligned} (\partial_{TT}^2 f)^2 \delta_{\rho,T}(\gamma_1) &= (\partial_{TT}^2 f)^2 \delta_{v,T}(\gamma_1) + (\partial_{TT}^2 f)^2 [\delta_{\rho,T}(\chi_2) - \delta_{v,T}(\chi_2)] \\ &\quad - (\chi_2) [\delta_{\rho,T}(\partial_{TT}^2 f) - \delta_{v,T}(\partial_{TT}^2 f)] \\ &\simeq (\partial_{TT}^2 f)^2 \delta_{v,T}(\gamma_1) + \frac{3}{2} (\partial v f) [\delta_{\rho,T}(\chi_2) - \delta_{v,T}(\chi_2)] \\ &\simeq (\partial_{TT}^2 f)^2 \delta_{v,T}(\gamma_1) \\ &\quad + \frac{3}{2} \frac{\delta m_0}{m_0} \sum_{ir} x_i y_i^r \left(\frac{(1 - y_i^r)(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \right) \\ &\quad \left(\frac{2 - y_i^r}{1 - y_i^r} + \frac{\delta_i^1 y_i^r}{1 + \delta_i^1(1 - y_i^r)} \right) \left(1 + \sum_{js} x_j y_j^s \right). \end{aligned} \quad (\text{A.20})$$

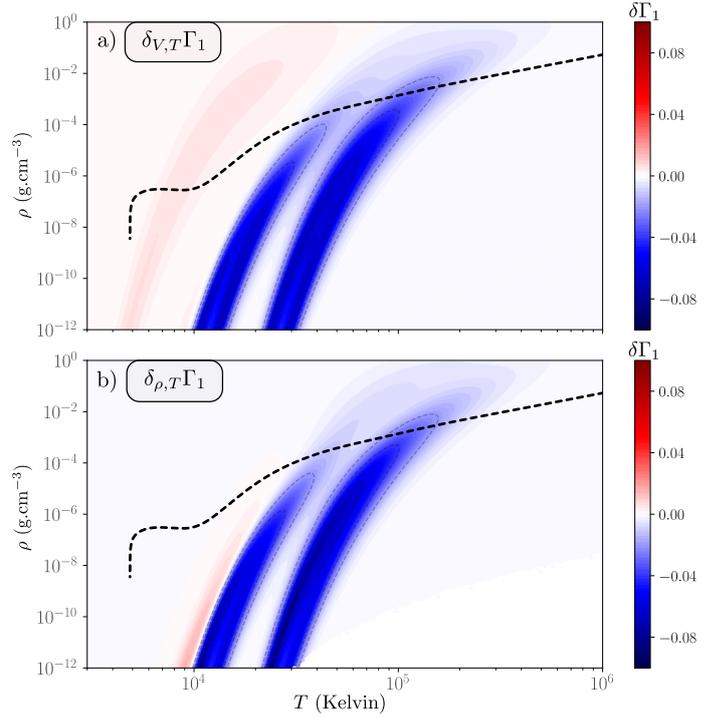


Fig. A.1. Representation of the approximations (A.17) (panel (a)) and (A.21) (panel (b)) given by a change $\delta Y = 0.1$ from the reference value $Y = 0.25$.

Finally, the overall perturbation on $\delta_{\rho,T}(\Gamma_1)$ is given by the two sums

$$\begin{aligned} \delta_{\rho,T}(\Gamma_1) &\simeq -\frac{1}{(\partial_{TT}^2 f)^2} \sum_{ir} x_i y_i^r \frac{(1 - y_i^r)(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \\ &\quad \left[\frac{\delta x_i}{x_i} + \sum_{js} x_j y_j^s \left(\frac{\delta x_i}{x_i} - \frac{\delta x_j}{x_j} \right) \right] \\ &\quad - \frac{\delta m_0/m_0}{(\partial_{TT}^2 f)^2} \sum_{ir} x_i y_i^r \left(\frac{(1 - y_i^r)(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \right) \\ &\quad \left(\frac{2 - y_i^r}{1 - y_i^r} + \frac{\delta_i^1 y_i^r}{1 + \delta_i^1(1 - y_i^r)} \right) \left(1 + \sum_{js} x_j y_j^s \right). \end{aligned} \quad (\text{A.21})$$

Although we have neglected many terms to obtain Eq. (A.21), this equation still gives a useful approximation. For instance, the difference between $\Gamma_1(\rho, T, 0.35) - \Gamma_1(\rho, T, 0.25)$ (shown in Fig. 5) and the relation (A.21) is bounded by 10^{-2} at any position of the (ρ, T) plane, and we verify that panel (b) of Fig. A.1 is very similar to what we found in panel (d) of Fig. 5. We also represent the perturbation given by (A.17), which appears to be closer to the shape that can be expected for the perturbation, at least based on panel (b) of Figure 5. The fact that the perturbation moves towards the borders of each ionisation region when it passes from $\delta_{v,T}(\Gamma_1)$ to $\delta_{\rho,T}(\Gamma_1)$ is then interpreted as the result of a shift in y_i^r caused by perturbing at constant ρ . In other words, at a given density, a perturbation on the abundances causes the ionisation regions to move, and this results in a less intuitive map. This effect can be seen as artificial, however, because it is obtained by fixing a variable that mixes the specific effects of various state variables (i.e. V and each N_i).

To give a qualitative understanding of the amplitude difference between the modification of the hydrogen and helium wells

shown in the top panel of Fig. A.1, we may consider the dominant term of the denominator in Eq. (A.17),

$$(\partial_{TT}^2 f)_{ir}^2 \sim 3x_i y_i^r \frac{(1 - y_i^r)(\chi_i^r/kT)^2}{1 + \delta_i^1(1 - y_i^r)} \left(1 + \sum_{js} x_j y_j^s \right), \quad (\text{A.22})$$

where the \sim notation stands for the dominant term in the ionisation region (i, r) . Comparing this with the numerator of

Eq. (A.17) gives an intuition of why the order of magnitude to keep in mind for the perturbation is $\delta x_i/x_i$ rather than δx_i . Thus, while we have $\delta x_1 = -\delta x_2$ for the hydrogen-helium case, we see that $\delta x_1/x_1 = -\delta x_2/x_1 = -(x_2/x_1)\delta x_2/x_2$. However, for $Y = 0.25$, $(x_2/x_1) = 1/12$, resulting in a perturbation about 12 times lower in the hydrogen region. This estimate is consistent with Fig. A.1, where the maximum in the hydrogen region is $\sim 6 \times 10^{-3}$ and the minimum in the helium region $\sim -7.2 \times 10^{-2}$.

Chapitre 5

Modélisation sismique du glitch de l'ionisation

Contenu du chapitre

- 5.1 Properties of the ionisation glitch. II. Seismic signature of the structural perturbation **136**
-

Properties of the ionisation glitch

II. Seismic signature of the structural perturbation

Pierre S. Houdayer¹, Daniel R. Reese, and Marie-Jo Goupil

LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 Place Jules Janssen, 92195 Meudon, France
e-mail: pierre.houdayer@obspm.fr

Received 9 February 2022 / Accepted 13 May 2022

ABSTRACT

Aims. In the present paper, we aim to constrain the properties of the ionisation region of a star from the oscillation frequency variation (a so-called glitch) caused by rapid structural variations in this very region. In particular, we seek to avoid the use of calibration based on stellar models, thus providing a truly independent estimate of these properties. These include both the helium abundance and other physical quantities that can have a significant impact on the oscillation frequencies, such as the electronic degeneracy parameter or the extent of the ionisation region.

Methods. Building on previous findings, we applied structural perturbations of the ionisation zone to the wave equation for radial oscillations in an isentropic region. The resulting glitch model is thus able to exploit the information contained in the fast frequency oscillation caused by the helium ionisation but also that in the slow trend accompanying the ionisation of hydrogen. This information can be directly expressed in terms of parameters related to the helium abundance, electronic degeneracy, and the extent of the ionisation region, respectively.

Results. Using Bayesian inference, we show that substantial recovery of the properties at the origin of the glitch is possible. We find a degeneracy between the helium abundance and the electronic degeneracy, which particularly affects the helium estimate. Extending the method to cases where the glitch is subject to contamination (e.g., surface effects), we note the importance of the slow glitch trend associated with hydrogen ionisation. We propose the use of a Gaussian process to disentangle the frequency glitch from surface effects.

Key words. asteroseismology – stars: oscillations – stars: interiors – stars: abundances

1. Introduction

When a star is subject to various forms of excitation, certain resonances may emerge that depend on its geometry and internal structure. When the star is spherically symmetric, each of these sustained oscillations can be described by quantum numbers such as its radial order n and degree ℓ , which reflect the geometric properties of the oscillations. The overall way in which the pulsation frequencies $\nu_{n,\ell}$ vary with radial order has been described with asymptotic developments (Vandakurov 1967; Tassoul 1980). These approximations have been very successful in describing solar-like oscillations, especially for large radial orders.

In addition to this asymptotic trend, the advent of high-precision photometry with CoRoT (Baglin et al. 2006), Kepler (Gilliland et al. 2010; Lund et al. 2017), and now TESS (Ricker et al. 2015; Stassun et al. 2019) has enabled the identification of more subtle contributions, including a particularly remarkable component, namely the ionisation glitch. Initially observed for the Sun, this deviation from the global trend manifests itself as an oscillatory component whose period, amplitude, and damping reflects local properties of the ionisation region (Gough 1990, 2002; Christensen-Dalsgaard & Perez Hernandez 1992). Numerous studies have then exploited this particular signature in solar-like stars in order to determine the position of their ionisation region (e.g., Verma et al. 2014, 2017; Broomhall et al. 2014) or

the amount of helium contained in this latter (e.g., Vorontsov et al. 1992; Perez Hernandez & Christensen-Dalsgaard 1994; Lopes et al. 1997; Verma et al. 2014, 2019). Most of the recent approaches are based on empirical models of the perturbation caused by the ionisation of helium (e.g., Monteiro & Thompson 2005; Houdek & Gough 2007) and relate their parameters to actual physical quantities via calibration on realistic models (Houdek & Gough 2011; Verma et al. 2014, 2019). In our first article (Houdayer et al. 2021, hereafter Paper I), we made various comments on this procedure and presented an alternative approach to modelling the ionisation region as well as the perturbation caused by a change in its properties. The derived structure is found to depend on three parameters that correspond respectively to the surface helium abundance, the electron degeneracy in the convective zone, and the extent of the ionisation region. Characterising the glitch based on such a model would thus allow us to extract physical quantities, such as the helium abundance, without the need for calibration, as well as other interesting constraints, such as the electronic degeneracy in the star's convection zone.

In order to reach this objective, we establish, in the present paper, the relation between the properties of the ionisation region and the resulting glitch. After providing some details on the glitch modelling in Sect. 2, we present the corresponding formalism (and the underlying assumptions) in Sect. 3. We subsequently present the tests that we carried out in order to

assess how well the properties of the ionisation region may be recovered from a glitch. The results are described in Sect. 4 for an ideal case and Sect. 5 when a contamination is included. The closing Sect. 6 is dedicated to our conclusions.

2. Creating an ionisation glitch model

2.1. Modelling the ionisation region

In Paper I, we presented a thermodynamically derived model of the structure inside the ionisation region. This is done by means of an analytical approximation of the first adiabatic exponent $\Gamma_1(\rho, T, X_{i>1})$ inside an adiabatic region; the $X_{i>1}$ refer here to the abundances of all elements except hydrogen. When applied to a hydrogen-helium mixture (of respective abundances $1 - Y$ and Y), solving the hydrostatic equilibrium as well as Poisson's equation defines the entirety of the associated structure (i.e., the density ρ , pressure P , sound speed c , and so on) at any location t (we often refer to the position in the star by its normalised acoustic depth $0 \leq t \leq 1$ in this paper). The resulting structural model, $\mathcal{S}(t)$, is therefore able to account for the impact of different amounts of helium in the ionisation region. More specifically, this model is fully determined by three parameters: (1) Y_s , the surface helium abundance, (2) ψ_{CZ} , the value of the electron degeneracy parameter inside the convective zone, and (3) ε , a small parameter defining the position and extent of the ionisation region. For a more compact notation, this parameterisation is hereafter designated by the vector $\theta = [Y_s, \psi_{CZ}, \varepsilon]^T$ (and thus the associated structural model $\mathcal{S}_\theta(t)$).

In the same paper, we proposed the use of a difference of these models $\delta\mathcal{S}_{\theta,\delta\theta}(t) = \mathcal{S}_{\theta+\delta\theta}(t) - \mathcal{S}_\theta(t)$ instead of an ad hoc profile in order to model the structural perturbation causing the ionisation glitch. Indeed, in exchange for an increased numerical cost (i.e., the time needed to compute the model $\mathcal{S}_{\theta+\delta\theta}(t)$ for each $\delta\theta$), this approach has many advantages, such as the possibility to account for more complex and realistic perturbations; the ability to model the perturbation of any structural profile and not only that of the first adiabatic exponent Γ_1 ; the way it relates these perturbations to actual physical properties of the region; and the fact that it is not limited to the impact of a pure helium change, and this for a limited number of parameters (3). In addition to the structural aspects raised here, we show that the method also leads to benefits regarding the exploitation of glitches.

2.2. Modelling the ionisation glitch using an ad hoc perturbation profile

Many approaches to obtaining an expression for the seismic glitch rely on formulae derived from the variational principle, which relate changes in oscillation frequencies to internal structure perturbations (Dziembowski et al. 1990; Gough & Thompson 1991; Antia & Basu 1994). Subsequently, the general aim would be to eliminate certain contributions to obtain a final expression that depends only on perturbations that can be described analytically. Therefore, a starting point to many derivations of theoretical forms of the frequency shift is the neglect of the contributions of slowly varying perturbations, such as differences in pressure $\delta_x P/P$ or density $\delta_x \rho/\rho$ compared to $\delta_x \Gamma_1/\Gamma_1$ (e.g., Monteiro & Thompson 2005; Houdek & Gough 2007, 2011) where δ_x designates a perturbation taken at fixed normalised radius $x = r/R$. As an example, the approximation used in Appendix A of Houdek & Gough (2007) gives the resulting (quadratic) frequency difference with respect to a given

reference:

$$\delta\omega^2 = \omega^2 - \omega_{\text{ref}}^2 \simeq \frac{\int_V \delta_x \Gamma_1 P (\text{div } \xi)^2 dV}{\int_V \rho \xi^T \xi dV}, \quad (1)$$

where $\omega = 2\pi\nu$ is the angular oscillation frequency, V designates the volume of the star, and ξ is the Lagrangian displacement caused by the oscillation. Such an expression can be highly useful; indeed, having an analytical expression for the single perturbation involved $\delta_x \Gamma_1$ is sufficient to provide a parametric form of the glitch. If it is also possible to approximate the eigenfunction (e.g. with the WKB formalism, cf. Bender & Orszag 1978), then a fully analytical expression for the frequency shift can be derived (Monteiro et al. 1994, 1998; Houdek & Gough 2007), although the many assumptions needed to achieve this may affect the reliability of the approximation. The main drawback of this analytical development lies in its primary interest: it eliminates the need to use a complete model of the structure. From then on, it is no longer possible to fit frequency differences from a reference model ($\delta\nu$). Instead, the frequencies (ν) must be fitted directly. Although the frequency shift can be approximated by Eq. (1), the reference frequencies, ν_{ref} , are not explicitly defined in this case. Rather than trying to model them roughly (e.g., via an asymptotic relationship), it is preferable to consider second differences in order to limit the error introduced (Gough 1990):

$$d^2\nu_{n,\ell} = \nu_{n+1,\ell} - 2\nu_{n,\ell} + \nu_{n-1,\ell}. \quad (2)$$

Applied to Eq. (1), the use of the second differences leads to

$$d^2\nu \simeq (\Delta\nu)^2 \frac{d^2\mathcal{F}}{d\nu^2} + d^2\nu_{\text{ref}}, \quad (3)$$

where $\Delta\nu$ corresponds to the asymptotic large separation and $\mathcal{F}(\nu)$ is an analytical approximation of

$$\mathcal{F}(\nu) \simeq \delta\nu \simeq \frac{\int_V \delta_x \Gamma_1 P (\text{div } \xi)^2 dV}{8\pi\nu \int_V \rho \xi^T \xi dV}. \quad (4)$$

This way, the $d^2\nu_{\text{ref}}$ component is generally assumed to be smooth, in the sense that it can be the result of a second derivative applied to the oscillation frequencies of a 'glitch-free' model. It is therefore usually eliminated from $d^2\nu$ by fitting a slowly varying function. We note that a significant assumption made here is the persistence of the linear perturbation regime, and thus that the perturbed model differs only slightly from the glitch-free model.

2.3. Modelling the ionisation glitch using a structural perturbation model

In contrast, the structural perturbation model $\delta\mathcal{S}_{\theta,\delta\theta}(t)$ on which we rely allows us to express the perturbation of each structural quantity, so that we are no longer limited to a formula that depends only on $\delta_x \Gamma_1$. It is therefore possible to rely on a more comprehensive form than Eq. (1), in which some of the contributions to the glitch are neglected. Also, it is possible to numerically compute the reference frequencies ω_{ref} as well as the eigenfunctions ξ based on $\mathcal{S}_\theta(t)$. With this in mind, we tried to construct a glitch model that makes the link explicit between a change in physical properties, $\delta\theta$, and frequency perturbations $\delta\omega_y^2(\delta\theta)$. The link between $\delta\theta$ and the structural perturbation $\delta\mathcal{S}_{\theta,\delta\theta}(t)$ having been established in our first paper, it is now a matter of deriving an expression analogous to Eq. (1) in order to express the relation between this structural perturbation and the glitch. A schematic representation of our approach is shown in Fig. 1.

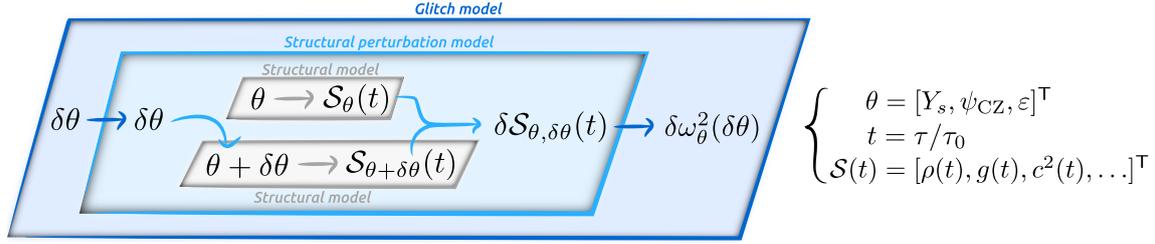


Fig. 1. Schematic diagram of the method adopted here to derive a glitch model. The aim of Sect. 3 is represented by the rightmost arrow, i.e., how to derive the glitch from a given structural difference.

3. Impact of structural perturbations on oscillation frequencies

Here, we aim to describe the behaviour of the oscillation frequencies when disturbed by a perturbation of the equilibrium quantities in the ionisation region. We rely on the following two assumptions: (1) the perturbation is localised in a convective zone, meaning that we consider the region to be strictly isentropic, and (2) the perturbation is close enough to the surface so that we can assume the displacement (at least for low degree pulsation modes) to be purely radial.

These approximations lead to the following simplifications in the calculations:

$$\ell = 0, \quad (5)$$

$$\gamma \equiv \frac{d \ln P}{d \ln \rho} = \left(\frac{\partial \ln P}{\partial \ln \rho} \right)_S \equiv \Gamma_1, \quad (6)$$

and therefore the exact cancellation of both Lamb S_ℓ and Brunt–Väisälä \mathcal{N} frequencies. Within this framework, the remaining constraints on the wave propagation (which are needed in order to define a cavity for a standing wave) are grouped under the term cut-off frequency and denoted ω_c .

3.1. The wave equation

Under the assumption of adiabatic and radial pulsations, the wave equation relating the oscillation frequency ω to the displacement amplitude ξ can be reduced to a Schrödinger equation expressed on the acoustic depth $\tau \equiv \int_r^R dr'/c$ (Roxburgh & Vorontsov 1994a,b):

$$\frac{\partial^2 \lambda \xi}{\partial \tau^2} + (\omega^2 - \omega_c^2) \lambda \xi = 0, \quad (7)$$

with,

$$\omega_c^2 = \frac{1}{4} \left(\frac{2c}{r} - \frac{g}{c} - \frac{dc}{dr} \right)^2 - \frac{c}{2} \frac{d}{dr} \left[\frac{2c}{r} - \frac{g}{c} - \frac{dc}{dr} \right] - 4\pi G \rho, \quad (8)$$

and the rescaling variable λ defined as

$$\lambda^2 \equiv -\frac{dm}{d\tau} = 4\pi r^2 \rho c. \quad (9)$$

In order to relate all the quantities that are homogeneous to a frequency in Eq. (8), it is useful to introduce

$$F_s \equiv \frac{d \ln |s|}{d\tau}, \quad (10)$$

the frequency scale of any structural quantity s . In this formalism, the cut-off frequency can be written as a quadratic form over a few frequency scales:

$$\omega_c^2 = 2F_r^2 + \frac{F_\rho^2}{4} + \frac{F_c^2}{4} + 3F_{Pr}F_\rho - 2F_rF_c + \frac{F_\rho F_g}{2} + \frac{F_c F_{dc/dr}}{2}, \quad (11)$$

by making use of the following properties:

$$F_{s_1 s_2} = F_{s_1} + F_{s_2},$$

$$F_r = -c/r \quad \text{and} \quad \frac{dF_r}{d\tau} = F_r(F_c - F_r),$$

$$F_\rho = g/c \quad \text{and} \quad \frac{dF_\rho}{d\tau} = F_\rho(F_g - F_c),$$

$$F_c = -\frac{dc}{dr} \quad \text{and} \quad \frac{dF_c}{d\tau} = F_c F_{dc/dr}.$$

From Eq. (11), we observe that the constraints to the wave propagation can be both of a geometrical (involving the terms in $F_r, F_g \sim c/r$ which dominate near the star's centre) and structural nature (composed of the terms in F_ρ, F_c , or $F_{dc/dr} \sim g/c$, dominating in the upper layers). A representation of the ratio between the resulting cut-off and oscillation frequencies ω_c^2/ω^2 with respect to the normalised acoustic depth $t = \tau/\tau_0$ ($\tau_0 = \int_0^R dr/c$ being the star's acoustic radius) is given in Fig. 2. From Eq. (7), we see that this ratio contains all the information about the propagation: as long as its value remains below 1 (blue area), the wave will adopt an oscillatory behaviour, whereas the wave will be evanescent if the ratio exceeds this limit (red area). The variations close to the surface, which break the symmetry between the centre and surface of the star, constitute a typical signature of the ionisation region.

Before going further, let us reiterate that we are interested here in describing the behaviour of the glitch and not of the frequencies that appear in Eq. (7). To this end, it is more convenient to normalise this relation by introducing quantities that can be compared from one structure to another. Accordingly, we rewrite the above Eq. (7) in terms of the following normalised variables: $\sigma = \omega\tau_0$, $\sigma_c = \omega_c\tau_0$, $t = \tau/\tau_0$, and $\eta = \lambda\xi/\|\lambda\xi\|$ with the notation:

$$\|\lambda\xi\| = \sqrt{\frac{1}{\tau_0} \int_0^{\tau_0} (\lambda\xi)^2 d\tau}. \quad (12)$$

The wave equation verified by the normalised eigenfunction $\eta(t, \sigma)$ can now be rewritten as follows:

$$\frac{\partial^2 \eta}{\partial t^2} + (\sigma^2 - \sigma_c^2) \eta = 0. \quad (13)$$

3.2. Obtaining a glitch model

Taking the product of Eq. (13) and η , and integrating the result from $t = 0$ to 1, we can naively write (after integrating by parts):

$$\sigma_{\text{mod}}^2 = \int_0^1 \left(\sigma_c^2 \eta^2 + \left(\frac{\partial \eta}{\partial t} \right)^2 \right) dt - \left[\eta \frac{\partial \eta}{\partial t} \right]_0^1, \quad (14)$$

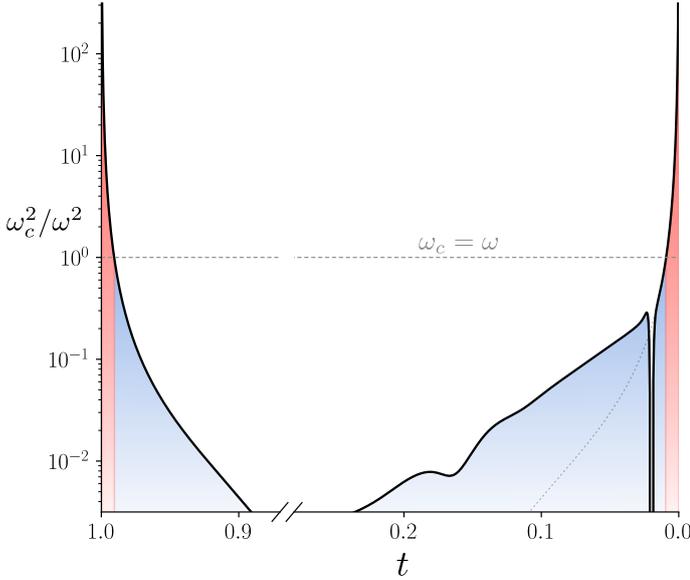


Fig. 2. Example of the ratio ω_c^2/ω^2 in a solar-like model with respect to the normalised variable $t = \tau/\tau_0$ (the surface is located on the right-hand side and the centre is made visible by a cut on the axis). In the above plot, we have chosen a typical solar pulsation frequency $\omega = 2\pi \times 3000 \mu\text{Hz}$. The wave propagates in the blue shaded region ($0.01 < t < 0.99$) below the horizontal dashed line ($\omega_c = \omega$) while being damped in the red shaded area. The dotted line near the surface has been added for the sake of comparison and represents the symmetric counterpart to the curve near the centre. It overlaps with the actual curve down to the ionisation region.

remembering that η is normalised and therefore $\int_0^1 \eta^2 dt = 1$. It is possible to show that the part in brackets can be considered to be null as long as it is possible to find a thin layer in both the surface and centre where the variations of Γ_1 are negligible. As this layer can be as thin as desired, we consider hereafter the following approximation:

$$\sigma_{\text{mod}}^2 = \int_0^1 \left(\sigma_c^2 \eta^2 + \left(\frac{\partial \eta}{\partial t} \right)^2 \right) dt. \quad (15)$$

It is clear that the actual oscillation frequencies (σ) can significantly differ from the modelled expression of oscillation frequencies (σ_{mod}) provided above. Indeed, if the assumptions of isentropy and radial oscillations appear to be coherent for pulsations located in a surface convective zone (and a fortiori in the ionisation region), these have been implicitly extended to the whole star during the integration. The resulting difference is accounted for by the function \mathcal{R} , such that:

$$\sigma^2 = \sigma_{\text{mod}}^2 + \mathcal{R} = \int_0^1 \left(\sigma_c^2 \eta^2 + \left(\frac{\partial \eta}{\partial t} \right)^2 \right) dt + \mathcal{R}. \quad (16)$$

In this form, Eq. (16) is particularly well-suited to studying the frequency difference $\delta\sigma = \sigma^B - \sigma^A$ between two models A and B . Using the variational principle (Chandrasekhar 1964) in order to perturb Eq. (16) merely provides, at first order:

$$\delta\sigma^2 = \delta\sigma_{\text{mod}}^2 + \delta\mathcal{R} \quad \text{with} \quad \delta\sigma_{\text{mod}}^2 \equiv \int_0^1 \delta_t(\sigma_c^2) \eta^2 dt, \quad (17)$$

where δ_t designates a difference in profile at identical t (where δ stands for a simple difference in quantity) between the two models (Paper I; Christensen-Dalsgaard & Thompson 1997).

The advantage of using normalised variables becomes apparent: while no problem arises with the difference in profile at a fixed value of t , it would have been different at fixed τ if ever $\tau_0^A \neq \tau_0^B$.

Equation (17) satisfies the objective set at the end of Sect. 2: it provides an explicit link between the perturbation of the oscillation frequencies $\delta\sigma^2$ and the perturbation of the structure through $\delta_t(\sigma_c^2)$. In this equation, $\delta\mathcal{R}$ characterises how closely the actual glitch $\delta\sigma^2$ is approximated by our modelled glitch expression $\delta\sigma_{\text{mod}}^2$. We note that it is not necessary for σ_{mod}^2 to be a good approximation of σ^2 for $\delta\mathcal{R}$ to be small. Indeed, it is only necessary for the assumptions to be verified in the regions where the structure differs between the models (as $\delta_t(\sigma_c^2)$ would be strictly null in the other regions in any case).

3.3. Comparison of glitch models with a numerically derived glitch

Although Eq. (17) does meet the criteria we set ourselves, it does not constitute the only possible approximation of the glitch. In particular, we propose to compare Eq. (17) (to which we refer with the index I) with the approximation of the glitch proposed by Eq. (1) (index II):

$$\delta\sigma_{\text{mod, I}}^2 = \int_0^1 \delta_t(\sigma_c^2) \eta^2 dt, \quad (18)$$

$$\delta\sigma_{\text{mod, II}}^2 = 2\sigma^2 \frac{\delta\tau_0}{\tau_0} + \frac{\int_V \delta_x \Gamma_1 P(\text{div } \xi)^2 dV}{\int_V \rho \xi^T \xi dV}. \quad (19)$$

In order to compare these two formulae, it is necessary to define the structural models A and B from which the glitch results. It may be noted that neither Eq. (19) nor Eq. (18) imposes a particular choice for model A or B as long as it is possible to derive the required structural perturbations. Here we choose two instances of the structural models defined in Paper I, $\mathcal{S}_{\theta^\circ}(t)$ and $\mathcal{S}_{\theta^\circ + \delta\theta}(t)$, respectively parameterised by $\theta^\circ = [Y_s^\circ = 0.255, \psi_{\text{CZ}}^\circ = -5, \varepsilon^\circ = 0.015]^T$ and $\theta^\circ + \delta\theta$. Such a situation can be seen as ideal, in the sense that the assumptions used to derive both glitch expressions I and II are verified by design (in particular, the two models differ only in the ionisation region). Expressions I and II should therefore be satisfactory approximations of $\delta\sigma^2$ (at first order) and this assertion is tested in Fig. 3. The reference for our comparison, the ‘true’ glitch, is merely the difference in oscillation frequency between the structures $\mathcal{S}_{\theta^\circ}(t)$ and $\mathcal{S}_{\theta^\circ + \delta\theta}(t)$. The frequencies of each structure have been calculated numerically using *InversionKit* (Reese et al. 2016) and appear in black symbols. This comparison is made for low degree modes $\ell \leq 2$ and a wide range of radial orders to show the asymptotic behaviour. Three glitches are considered (respectively in panels a, b, and c), corresponding to a difference in each parameter, the amplitude of which was chosen to cause a similar impact on the frequencies in terms of amplitude: $\delta\theta = \{[0.05, 0, 0]^T, [0, 0.2, 0]^T, [0, 0, 0.001]^T\}$. Finally, in order to better account for the departures, we have chosen to compare $\delta\sigma/\sigma$ with $\delta\sigma_{\text{mod}}^2/2\sigma^2$ rather than $\delta\sigma^2$ with $\delta\sigma_{\text{mod}}^2$ directly.

A first thing to note about the glitches calculated numerically is that they seem to share a unique relationship regardless of their degree, which is in line with our hypothesis, which states that the displacement resulting from these low degree modes is almost radial in the perturbed region. Also, in addition to the clearly visible oscillation caused by the ionisation of helium, which gradually fades away at high frequencies, a slowly varying component is apparent. This component, which is visible in every panel, persists for longer than the component of helium and also oscillates

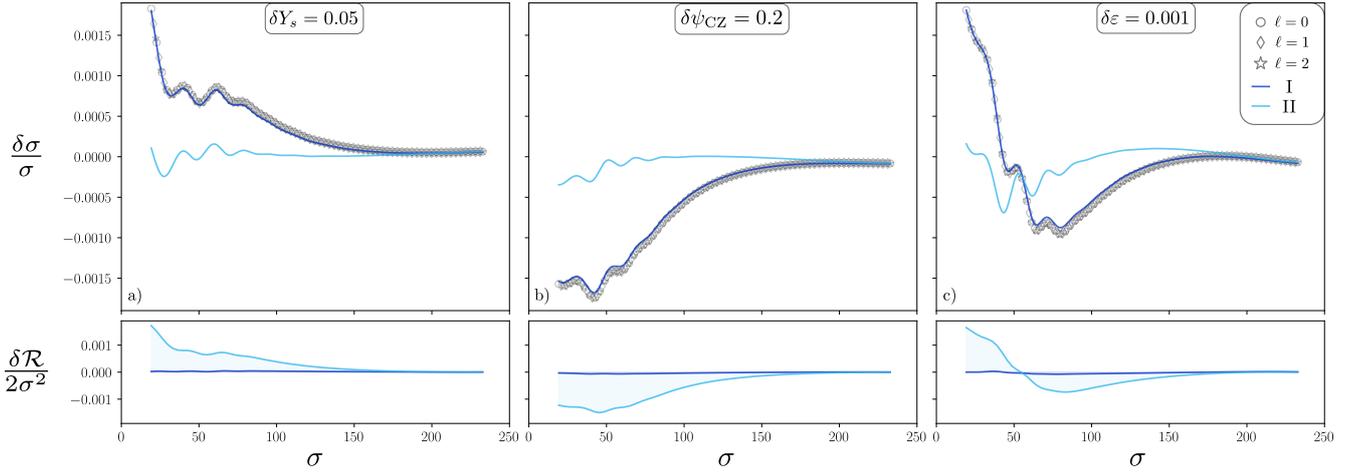


Fig. 3. Comparison of numerically derived glitches $\delta\sigma/\sigma$ (black symbols, $0 \leq \ell \leq 2$, $5 \leq n \leq 70$) with both modelled expressions $\delta\sigma_{\text{mod, I}}^2/2\sigma^2$ (dark blue curves) and $\delta\sigma_{\text{mod, II}}^2/2\sigma^2$ (light blue curves) given by Eqs. (18) and (19) for three differences in the parameters: (a) $\delta Y_s = 0.05$, (b) $\delta\psi_{\text{CZ}} = 0.2$, and (c) $\delta\varepsilon = 0.001$. The *bottom panels* show the residual functions $\delta\mathcal{R}/2\sigma^2$ for both approximations.

much more slowly. A standard interpretation in terms of glitch tends to suggest an origin that is both more localised and superficial than that of helium. We believe that this is the signature of hydrogen ionisation, which has an impact on σ_c^2 as can be seen in Fig. 2 at $t \sim 0.02$ (this is to be compared with the impact of helium ionisation at $0.1 < t < 0.2$).

Regarding the modelled expressions, the glitch evaluated using Eq. (18) (darker curves) provides a near perfect recovery of the ‘true’ glitch in each situation. Considering the residual function $\delta\mathcal{R}/2\sigma_{\text{num}}^2$, a subtle undershoot is visible which is to be expected from a first-order approximation. In comparison, Eq. (19) (lighter curves) provides poorer approximations, even though it can account for part of the oscillation (we note that $\delta\sigma_{\text{mod, II}}^2/2\sigma_{\text{num}}^2$ was corrected by an offset to better account for the asymptotic behaviour). However, by looking at the residuals, we see that by neglecting the pressure and inertia perturbations, Eq. (19) has not only discarded the slowly varying component but also some of the quickly oscillating part of the glitch.

4. Estimating the properties of the ionisation region from the glitch

As stated above, it is worth noting that we do not assume a specific shape for the structural perturbations up to Eqs. (18) and (19). For instance, we consider the structural perturbations $\delta_i\sigma_c^2$ and $\delta_x\Gamma_1/\Gamma_1$ as given to perform the comparisons in Fig. 3. As suggested in Sect. 2 (and in Fig. 1), our objective is now to combine the structural perturbation model $\delta\mathcal{S}_{\theta,\delta\theta}(t)$ with Eq. (18) to obtain a parameterised model of the ionisation glitch. With such a model, we investigate in this section how to infer the properties of the ionisation region $\delta\theta$ from the glitch $\delta\sigma^2$.

4.1. Constructing a glitch model from a structural perturbation model

Used in its present form, Eq. (17) could allow us to find, knowing the shape of the glitch, the profile of $\delta_i\sigma_c^2$ by inversion. However, it seems likely in practice that this recovery would be patchy, especially given the number of frequencies available and the complexity of the cut-off frequency profile. As pointed out in Sect. 2, we instead prefer to consider the shape of the perturbation as constrained by a parametric structural pertur-

bation model $\delta\mathcal{S}_{\theta,\delta\theta}(t)$. This implies the following assumption: $\delta_i\sigma_c^2 \approx \sigma_c^2(t, \theta + \delta\theta) - \sigma_c^2(t, \theta)$. The resulting perturbation $\delta_i\sigma_c^2$ can therefore be fully described using the pair $(\theta, \delta\theta)$. In practice, the quantity θ designates the value of the parameters in our reference model and is therefore known by design. In contrast, $\delta\theta$ accounts for the parameter differences that best explain the structural difference and must be inferred from the available data. For brevity, our glitch model is therefore written as a function of $\delta\theta$ only, although it depends on the chosen reference:

$$\delta\sigma^2 \approx \delta\sigma_{\text{mod}}^2(\delta\theta) = \int_0^1 \delta_i\sigma_c^2(\delta\theta) \eta^2 dt. \quad (20)$$

The quantity $\delta_i\sigma_c^2$ being perfectly defined as a function of $\delta\theta$, this expression can be used in its present form to estimate the variation in parameters responsible for the glitch. However, in the following, we mainly focus on the linearised form of this expression:

$$\delta\sigma_{\text{mod}}^2(\delta\theta) \approx (\nabla_{\theta}\sigma_{\text{mod}}^2)^{\text{T}} \delta\theta, \quad (21)$$

with

$$\nabla_{\theta}\sigma_{\text{mod}}^2 \equiv \int_0^1 (\nabla_{\theta}\sigma_c^2) \eta^2 dt. \quad (22)$$

As discussed in Sect. 2, the main trade-off between the use of this model and the fully analytical forms of the glitch is in numerical cost. Whatever the value of $\delta\theta$, it is necessary to compute the model $\mathcal{S}_{\theta+\delta\theta}(t)$ in order to find the profile $\sigma_c^2(t, \theta + \delta\theta)$ required to calculate Eq. (20). Equation (21) completely eliminates this constraint because it is only necessary to estimate $\nabla_{\theta}\sigma_{\text{mod}}^2$ once and for all, thus introducing an error in $\mathcal{O}(\delta\theta^{\text{T}}\delta\theta)$ in the process. We note that this linearisation assumes that the gradient $\nabla_{\theta}\sigma_{\text{mod}}^2$ only depends slightly on the parameters of the reference model, which means that our expression (20) only depends on $\delta\theta$ and is locally insensitive to θ (which justifies its omission in the glitch model).

4.2. The fitting method

We now quantify the overall error on the recovery of $\delta\theta$ from a glitch $\delta\sigma^2$ using Eq. (21). For this purpose, we consider the

example of a glitch caused by the change of parameters $\delta\theta_{\text{true}} = [0.05, 0.2, 0.001]^\top$. Regarding Fig. 3, this implies studying the glitch that results from combining the perturbation in each panel. This example offers two insights: firstly, it can reveal potential non-linearities as it is a non-negligible perturbation (each impact is of the order of the helium change $\delta Y_s = 0.05$, and therefore their combination constitutes a significant departure from the reference). Secondly, closer inspection of Fig. 3 leads us to expect a certain degree of degeneracy between the effects of δY_s and $\delta\psi_{\text{CZ}}$. Although the helium glitch is fairly similar in all panels (comparable amplitudes, periods, and damping), these two parameters, that is, δY_s and $\delta\psi_{\text{CZ}}$, also cause a comparable hydrogen glitch but with opposite sign. As we can expect a near cancellation by combining these two contributions, it is interesting to see how much information can be recovered about the parameter differences from Eq. (21).

Hereafter, $\delta\sigma^2$ denotes the vector of frequency shifts forming the glitch at all radial orders. The recovery is achieved by means of a probabilistic approach, that is assuming that our parameter vector $\delta\theta$ is a random variable, and we try to estimate its distribution given the observation of a certain glitch $\delta\sigma^2$. This random variable is referred to here as $\delta\theta|\delta\sigma^2$ and its probability density function is referred to as the posterior distribution. Such an approach may excessively complicated for a linear problem like this one, but it is nevertheless essential to account for the potential degeneracies of this problem as well as the uncertainties on our estimate.

The distribution $p(\delta\theta|\delta\sigma^2)$ is estimated with the help of the Python library *emcee*¹, which implements MCMC methods for sampling distributions. Further details about this procedure and the underlying assumptions can be found in Appendix A.1. The synthetic data, $\delta\sigma^2$, generated for this purpose result from the oscillation frequency differences between the structures $\mathcal{S}_{\theta^\circ}(t)$ and $\mathcal{S}_{\theta^\circ+\delta\theta}(t)$ as was done in the previous section. The parameter recovery can therefore be seen as an ‘inverse crime’, in the sense that the assumed ionisation region structure in Eq. (21) actually corresponds to the one used to calculate the synthetic glitch. It constitutes therefore an ideal framework that tests the derivation of Eq. (17) and the subsequent linearisation Eq. (21) rather than the relevance of the structural model itself. In contrast to Sect. 3, we consider here only purely radial modes with orders ranging from 11 to 27, in accordance with the modes we expect to observe. The standard deviation on the frequencies was set to $s_n = 5 \times 10^{-3}$, which would result in a $\sim 0.2 \mu\text{Hz}$ dispersion for frequencies typical of a solar-like star. Although optimistic, this value is not unrealistic either: it is typically of the order of the mean uncertainty on the radial and dipolar modes in the *Kepler* LEGACY sample (Lund et al. 2017).

4.3. The resulting posterior distribution

The resulting posterior distribution after applying the above procedure is shown in Fig. 4 and summaries of this same distribution can be found in Table 1. The corner plot shows the univariate (two times marginalised) distributions on the diagonal as well as each bivariate (one time marginalised) distributions below, which means that the figure, as a whole, is fairly representative of the underlying (3D) distribution.

Compared to the actual value of $\delta\theta$ (shown with dashed lines), the most plausible value according to the PDF, $\delta\theta_{\text{MAP}}$, is somewhat underestimated ($\sim 30\%$ on δY_s , $\sim 10\%$ on $\delta\psi_{\text{CZ}}$ and $\sim 5\%$ on $\delta\epsilon$) albeit fairly close. The extent of the PDF is directly

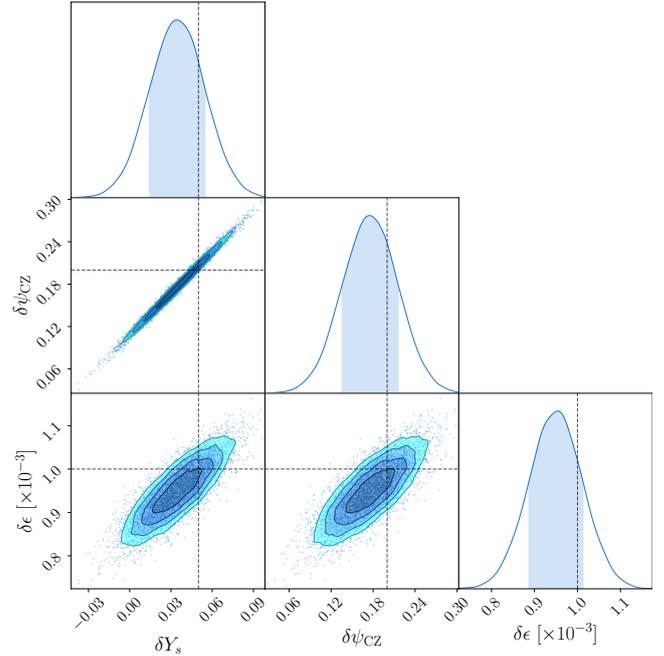


Fig. 4. MCMC estimate of the distribution $p(\delta\theta|\delta\sigma^2)$. The dashed lines indicate the value of the true $\delta\theta_{\text{true}} = [0.05, 0.2, 1 \times 10^{-3}]^\top$ used to generate $\delta\sigma^2$ (the PDF gives at this point: $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2) = 15.72$). The mode of this distribution is $\delta\theta_{\text{MAP}} = [0.036, 0.178, 0.951 \times 10^{-3}]^\top$. This plot was made with the Python package *ChainConsumer* (<https://samreay.github.io/ChainConsumer>).

Table 1. Comparison of the exact and fitted change in parameters.

Parameters	‘True’ values	Estimates (*)
δY_s	0.05	0.034 ± 0.020
$\delta\psi_{\text{CZ}}$	0.2	$0.175^{+0.041}_{-0.039}$
$\delta\epsilon$	1×10^{-3}	$(95.5^{+5.8}_{-6.8}) \times 10^{-5}$

Notes. (*) The values shown here are the distribution medians along with the 16th and 84th percentiles. We note that they differ slightly from the mode (cf. Fig. 4) because of the distribution asymmetries.

related to the noise level surrounding the oscillation frequencies s_n . Nevertheless, it is possible to draw some conclusions independently of a specific s_n value: the estimate of the helium abundance seems to be the most volatile, followed by that of the electronic degeneracy, while the estimate of $\delta\epsilon$ is the most stable. The reasons behind this statement are clearly apparent from Fig. 4. As expected, a strong degeneracy appears between δY_s and $\delta\psi_{\text{CZ}}$ (the two parameters share a correlation coefficient of 0.98) as they produce relatively similar effects on the glitch (cf. Fig. 3). It should be noted that this degeneracy was already expected from the structure itself because both parameters impact the depth of the helium wells in the Γ_1 profile (Paper I). This degeneracy especially increases the uncertainties on the helium abundance, making extreme values like $\delta Y_s = 0$ still relatively credible. It can also be noted that these two parameters also show a certain degree of correlation with $\delta\epsilon$. This, unlike the previous degeneracy, which is of physical origin, can be imputed to a lack of information in the data and in particular on the hydrogen component of the glitch. This particular point is investigated in Appendix B, where a recovery of the parameters with additional low-order frequencies is presented.

¹ <https://emcee.readthedocs.io>

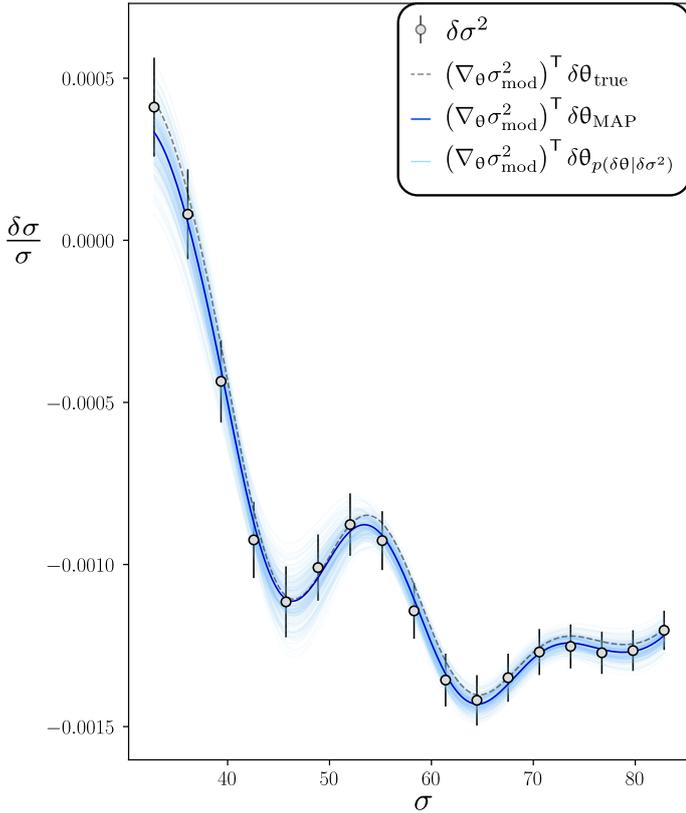


Fig. 5. Comparison of the numerical glitch $\delta\sigma^2/2\sigma^2$ (circular symbols, $\ell = 0$ and $11 \leq n \leq 27$) with the modelled glitch expression $(\nabla_{\theta}\sigma_{\text{mod}}^2)^{\top}\delta\theta/2\sigma^2$ in the case of (1: grey dashed curve) $\delta\theta = \delta\theta_{\text{true}}$, the actual value used to generate $\delta\sigma^2$, (2: dark blue curve) $\delta\theta = \delta\theta_{\text{MAP}}$, the MAP estimate, and (3: light blue curves) >300 different $\delta\theta_{p(\delta\theta|\delta\sigma^2)}$ sampled from the posterior distribution $p(\delta\theta|\delta\sigma^2)$ (see Fig. 4).

Finally, it is tempting to judge the quality of the fit based on a distance such as $\|\theta_{\text{true}} - \theta_{\text{MAP}}\|$ for instance, eventually defining a metric accounting for the distribution spread in each direction. However, it would be complicated for such a distance to account for strong correlations or degeneracies in the distribution, in which case the parameter space admits preferential directions. A very natural way to proceed is to consider the value of the PDF at θ_{true} , $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2)$, which at the same time penalises estimates that are too distant or too vague (because of the normalisation of the distribution) while taking into account the degeneracies². In the present case, this metric gives the value $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2) = 15.72$. This value on its own has only limited meaning (although it is clear that it should be as high as possible). It is nonetheless interesting to compare it with other parameter recoveries, as we show in the following section.

To visualise the posterior distribution in terms of frequency shifts, we show in Fig. 5 a comparison between the numerical glitch and our glitch model. Here, we compare the model for the true value of the parameters, $\delta\theta_{\text{true}}$, the model for the most probable value of the parameters given the data, $\delta\theta_{\text{MAP}}$, and more than 300 curves drawn according to the posterior distribution on the parameters. These curves give an idea of how the spread of the distribution effectively propagates through the glitch models.

² We may note that this quantity is nothing more than the (negative) cross-entropy between the ‘perfect’ Dirac distribution, $\delta(\delta\theta - \delta\theta_{\text{true}})$, and the posterior distribution $p(\delta\theta|\delta\sigma^2)$.

Regarding the appearance of the curves, it is not surprising to see that the synthetic glitch is fairly similar to the one shown in panel c of Fig. 3 (in the corresponding frequency range) because the helium and electronic degeneracy contributions to the glitch approximately cancel out. A first thing to note is that the model, even in this linear version, accurately reproduces the actual glitch for the true value of $\delta\theta$ and thus justifies the linearisation in Eq. (21). This curve is noticeably close to that of the maximum a posteriori (MAP) estimate, $\delta\theta_{\text{MAP}}$, which confirms the relevance of the retrieved solution. Also, the models drawn a posteriori account well for the uncertainties on the glitch based on the lighter curves. This illustrates the importance of the degeneracy on the helium abundance: glitch models that differ by more than 0.05 in δY_{s} may sometimes barely be distinguished.

5. Preliminary applications to seismic data

The results presented in the previous section suggest our method is able to recover most of the information on the ionisation region properties contained in the glitch, and this without needing any calibration on realistic models. Therefore, despite the degeneracy affecting the reliability of our helium estimate, the procedure introduced seems to provide reasonable predictions for the parameters (as an indication, the MAP estimate on the total parameters would give $\theta_{\text{MAP}} = [0.291, -4.82, 1.60 \times 10^{-2}]^{\top}$ to be compared with $\theta_{\text{true}} = [0.305, -4.80, 1.60 \times 10^{-2}]^{\top}$). However, at this point, it is important to bear in mind that the recovery of parameters in Sect. 4 is carried out in an ideal framework. Although we have only considered an observable range of radial modes with an acceptable noise level, the difference between the two models under consideration is a perturbation of the ionisation region. By design, this implies that the synthetic frequency differences only reflect the presence of the ionisation glitch.

It is fairly easy to draw the limits of this case when confronted with actual seismic data³: (1) by assuming $\gamma = \Gamma_1$, we neglected contributions of the Brunt–Väisälä frequency $\delta\mathcal{N}^2$ to the frequency shift $\delta\sigma^2$. Although they are indeed zero in the ionisation region, such differences could very well occur in a radiative region because of the μ -gradient for instance (Deheuvels & Michel 2010; Noll et al. 2021). Thus, the unexpected presence of a convective core in the observed star or a deficient modelling of the chemical composition in this region can impact the differences in the oscillation frequencies. However, it should be noted that the glitch model can easily be extended to the case $\delta\mathcal{N}^2 \neq 0$ by simply replacing the cut-off frequency perturbation $\delta\sigma_c^2$ by the total acoustic potential perturbation $\delta\mathcal{V}$ (Roxburgh & Vorontsov 1994b).

(2) By restricting ourselves to the ionisation region, we have neglected many of the perturbations to the cut-off frequency occurring elsewhere. Other well-known contributions are the signature of the transition between the convective and radiative regions, and surface effects. The first signature results from a much deeper layer than the ionisation region for the stars we are interested in and thus manifests itself as a higher frequency (and lower amplitude) oscillation in the glitch (e.g., Monteiro et al. 1994; Houdek & Gough 2007). By far the most important in terms of amplitude, surface effects are at the origin of much slower components. They can be considered as the main source of ‘contamination’ to the ionisation glitch fit and are one of the reasons for using second differences (cf. Eq. (2)) instead of frequencies directly.

³ We note that it is not necessary to focus on potential changes to the Lamb frequency δS_{ℓ}^2 as long as we only consider radial modes.

These effects, if not treated properly, are likely to have a severe impact on our estimate of the properties of the ionisation region. In the remainder of this section, we explore two approaches for applying our method despite the presence of interfering components in the data. The first is very common and seeks to reduce these components by combining frequencies and the second attempts to fit the frequencies directly using Gaussian processes.

5.1. Using frequency combinations as a diagnostic

Considering a diagnostic, \mathbf{y} , expressed as a linear combination of the squared frequencies, σ^2 , we write:

$$\mathbf{y} = \mathbf{C}\sigma^2, \quad (23)$$

where \mathbf{C} is a matrix for which we can assume a general form at this point. We note here that the diagnostic introduced above differs slightly from those that are used habitually, which are written as a combination of the frequencies directly. Using squared frequencies brings a more intuitive understanding in the formalism introduced so far, while giving identical results, as confirmed by supplementary numerical tests. The general idea, and a fortiori the elimination of the most superficial components when \mathbf{C} is the discrete derivative operator, remains very similar.

Subsequently, from Eqs. (21) and (22), it follows immediately from Eq. (23) that

$$\delta\mathbf{y} \simeq (\nabla_{\theta}\mathbf{y}_{\text{mod}})^{\text{T}} \delta\theta, \quad (24)$$

with

$$\nabla_{\theta}\mathbf{y}_{\text{mod}} \equiv \int_0^1 (\nabla_{\theta}\sigma_c^2) (\mathbf{C}\boldsymbol{\eta}^2)^{\text{T}} dt, \quad (25)$$

where $\boldsymbol{\eta}^2$ is the vector composed of all the normalised eigenfunctions η^2 . We note that in this vector form (accounting for all frequencies at once) $\nabla_{\theta}\mathbf{y}_{\text{mod}}$ designates a matrix, which is very similar to

$$\nabla_{\theta}\sigma_{\text{mod}}^2 \equiv \int_0^1 (\nabla_{\theta}\sigma_c^2) (\boldsymbol{\eta}^2)^{\text{T}} dt, \quad (26)$$

the vector form of Eq. (22). In fact, $\nabla_{\theta}\mathbf{y}_{\text{mod}}$ refers to the decomposition of the same function, $\nabla_{\theta}\sigma_c^2$, but on different kernels, namely $\mathbf{C}\boldsymbol{\eta}^2$ instead of $\boldsymbol{\eta}^2$. While $\boldsymbol{\eta}^2$ designates oscillating functions of constant amplitude in the major part of the star, the combinations $\mathbf{C}\boldsymbol{\eta}^2$ can take more general envelopes. In the specific case where $\mathbf{C} = \mathbf{D}^k$, with \mathbf{D}^k being the discrete derivative operator to a given power k , this envelope varies as $\sin^k(\pi t)$ independently of the frequency σ (Ballot et al. 2004). This combination can therefore be seen as enhancing the signal of $\nabla_{\theta}\sigma_c^2$ at $t = t_1$ compared to $t = t_2$ by a factor of $(\sin(\pi t_1)/\sin(\pi t_2))^k$. With regard to the concerns raised above, the main advantage of this method clearly appears when it comes to getting rid of the surface effects. Taking the example of a ‘contamination’ at $t_{\text{surf}} \sim 0.01$ and a helium glitch at $t_{\text{He}} \sim 0.15$, the signal of the latter is comparatively amplified by a factor of more than 200 for $k = 2$ (corresponding to a second derivative). However, the same reasoning holds for the hydrogen component – for which $t_{\text{H}} \sim 0.02$ – which is thereby completely removed. It is therefore important to assess the impact of this loss of information in our fit, especially as we expect this component to be essential when recovering the different parameters.

To this end, we performed a similar analysis to that conducted in Sect. 4, this time aiming to fit the perturbation of

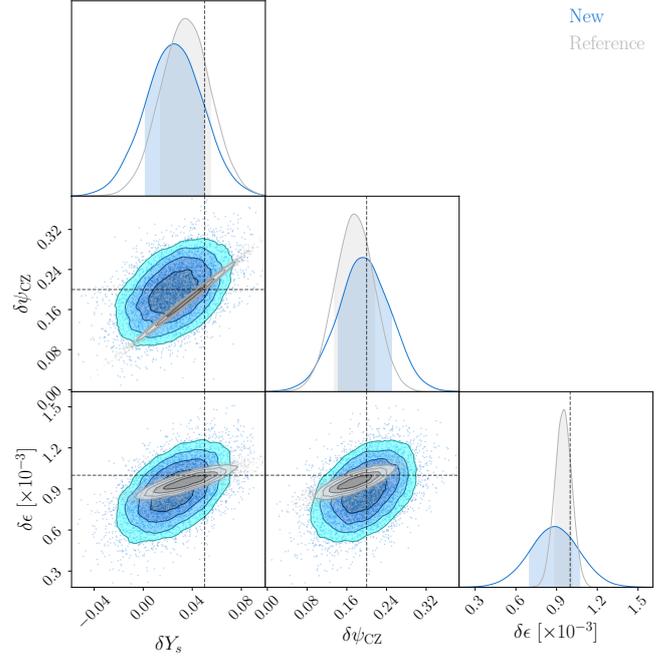


Fig. 6. MCMC estimate of the distribution $p(\delta\theta|\delta\mathbf{y})$ in the case $\mathbf{y} = \mathbf{D}^2\sigma^2$ (in blue) compared to the case $\mathbf{y} = \sigma^2$ shown in Fig. 4. The dashed lines indicate the true value of $\delta\theta_{\text{true}} = [0.05, 0.2, 1 \times 10^{-3}]^{\text{T}}$ used to generate $\delta\mathbf{y}$ (the PDF gives at this point: $\ln p(\delta\theta_{\text{true}}|\delta\mathbf{y}) = 12.17$). The mode of this distribution is $\delta\theta_{\text{MAP}} = [0.026, 0.195, 0.888 \times 10^{-3}]^{\text{T}}$.

the second differences $\delta\mathbf{y} = \delta(\mathbf{D}^2\sigma^2)$ instead of that of the frequencies $\delta\sigma^2$ directly. The method is completely analogous to that described above, and some details can be found in Appendix A.2. For the exact same value of $\delta\theta_{\text{true}} = [0.05, 0.2, 1 \times 10^{-3}]^{\text{T}}$, Fig. 6 gives a MCMC sampling of the posterior distribution $p(\delta\theta|\delta\mathbf{y})$ using the diagnostic $\delta\mathbf{y}$. Compared to the reference distribution $p(\delta\theta|\delta\sigma^2)$ found in Fig. 4 and shown in grey-scale here, this new PDF is notably more spread out. This is not surprising, because the elimination of the slowly varying component can only result in a loss of information; nevertheless it is worth quantifying this loss. We see that an important loss of information concerns the parameter $\delta\epsilon$, which can be seen in the spreading of its marginal distribution. As mentioned in Sect. 4, with the compensation of the contributions of δY_s and $\delta\psi_{\text{CZ}}$, it is expected that the hydrogen component mainly constrains $\delta\epsilon$. Thus, its absence here is most clearly manifested by this last parameter.

In addition, the shape of the distribution significantly changes on the δY_s vs. $\delta\psi_{\text{CZ}}$ panel, with the clear degeneracy between the two parameters now being more diffuse. Without the reference distribution for comparison, we might have concluded that the degeneracy was removed. However, the comparison makes it apparent that it is rather the presence of the helium oscillation alone (i.e. without the contribution of hydrogen) that allows more pairs $(\delta Y_s, \delta\psi_{\text{CZ}})$ to be credible. Thus, the physically based degeneracy between δY_s and $\delta\psi_{\text{CZ}}$ seems to be ‘resolved’ when sufficient information is available, while being hidden in the uncertainty when knowledge on the slowly varying component is lacking.

Finally, the mode of distribution, $\delta\theta_{\text{MAP}} = [0.026, 0.195, 0.888 \times 10^{-3}]^{\text{T}}$, has changed compared to the previous case, although not drastically regarding the extent of the distribution (however, we may note that the most plausible value of δY_s is underestimated by $\sim 50\%$ compared to the true value). However,

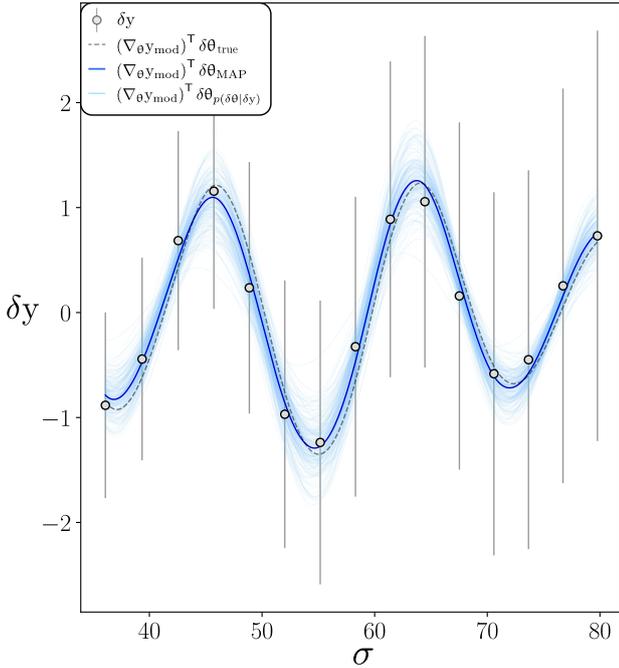


Fig. 7. Same as Fig. 5 but for the synthetic diagnostic $\delta\mathbf{y}$ and the modelled expression $(\nabla_{\theta}\mathbf{y}_{\text{mod}})^{\top}\delta\theta$.

as mentioned earlier, a more appropriate way to judge the quality of the fit is through the value of the distribution in $\delta\theta_{\text{true}}$, which gives $\ln p(\delta\theta_{\text{true}}|\delta\mathbf{y}) = 12.17$. This metric notably tells us that $p(\delta\theta_{\text{true}}|\delta\mathbf{y})$ is about 35 times lower than $p(\delta\theta_{\text{true}}|\delta\sigma^2)$ and thus quantifies the deterioration of the fit by using the second differences as a diagnostic.

We once more compared our model provided by Eq. (24) and the diagnostic obtained from numerical frequency deviations in Fig. 7. This time a single oscillatory component, that of helium, appears as expected. A first observation stands out, namely the considerable amplification of the uncertainty due to the linear combination. The errors shown in the figure are given by the square root of the diagonal coefficients of Σ_y , the covariance matrix, for lack of a better representation. Therefore, while being visually dominant, these error bars are not necessarily representative of the uncertainties on the diagnostic \mathbf{y} as they do not account for the correlated part of the noise (which helps to reduce some of the overall uncertainty). We may also note the proximity of the curves given by $(\nabla_{\theta}\mathbf{y}_{\text{mod}})^{\top}\delta\theta$ and $(\nabla_{\theta}\mathbf{y}_{\text{mod}})^{\top}\delta\theta_{\text{MAP}}$ despite their difference in terms of the parameters. More generally, many curves drawn from $p(\delta\theta|\delta\mathbf{y})$ are of comparable amplitude, whereas some of them differ by more than 0.1 in δY_s for instance.

Overall, we can see that the use of a diagnostic, while mitigating contamination of the ionisation glitch, such as that caused by surface effects, might hinder our ability to recover the solution. Even in the ideal case presented here (in the absence of contamination), finding a way to preserve the hydrogen component appears essential. Therefore, in the following section we present an approach to fit the ionisation glitch directly even in the presence of additional components.

5.2. Using frequencies directly

In this section, we take the explicit example of a contamination by surface effects and therefore consider a direct approximation

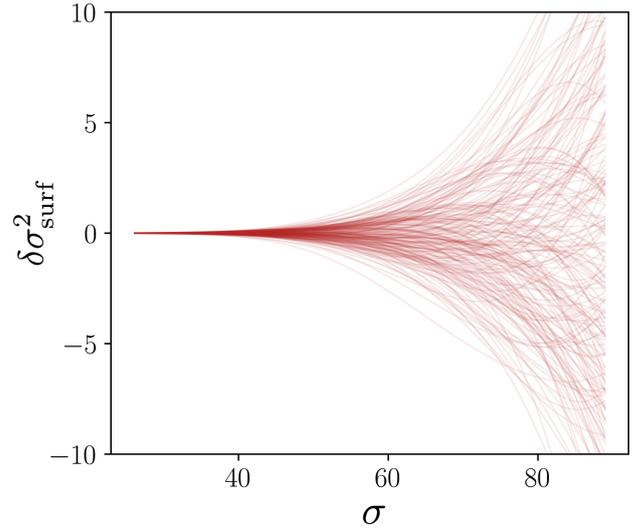


Fig. 8. Example of 200 realisations of the Gaussian process $\delta\sigma^2_{\text{surf},\mu}(\sigma)$ (cf. Eqs. (A.10) and (A.11)) for a fixed value of $\mu = [\log\beta^2, \log M]^{\top} = [50, 6]^{\top}$.

of $\delta\sigma^2$ by relying on the following model:

$$\delta\sigma^2_{\text{mod}}(\delta\theta, \mu) = \delta\sigma^2_{\text{ion}}(\delta\theta) + \delta\sigma^2_{\text{surf}}(\mu), \quad (27)$$

where $\delta\sigma^2_{\text{ion}}(\delta\theta)$ is given by Eq. (21) and $\delta\sigma^2_{\text{surf}}(\mu)$ designates a parameterised model of the surface effects. Expressing such a model explicitly is by no means straightforward, especially because what we refer to as ‘surface effects’ depend more on what is included (or not) in our reference structural model than on a well-defined physical phenomenon. It is therefore clear that the model describing the corresponding frequency shifts must be somewhat flexible. Meanwhile, this model must satisfy constraints, which are that the frequency shifts must (1) vary on sufficiently large scales – by definition, these effects include the most superficial deviations from the structure; and (2) be much weaker at low than at high frequencies. Indeed, the acoustic modes are becoming more and more sensitive to the surface effects with increasing radial orders.

To satisfy these conditions, here we have chosen to model these surface effects by a Gaussian process (GP) whose parameterisation by μ is specified in Appendix A.3. To illustrate the possibilities offered by this GP, Fig. 8 shows 200 realisations of the process given by Eqs. (A.10) and (A.11) for a fixed set of values for μ . As desired, the generated functions are indeed of low amplitude at low frequencies and exhibit a polynomial envelope. They also show a certain flexibility and generally end up deviating from a simple power law at higher frequencies, which allows them to mimic a more comprehensive behaviour. It is also clear that many of the functions in this sample do not correspond to a plausible observation of surface effects. Without observational constraints, this sample only provides us with an a priori view of the functions that the process generates.

On a practical level, the use of GPs is implemented via the Python library *george*⁴ while the distribution sampling is carried out as in Sect. 4. Naturally, we need to synthetically generate surface effects this time in order to test the method in a controlled framework. We followed the prescription of [Sonoji et al. \(2015\)](#) to generate ad hoc surface effects that we would expect to find from a solar-like star. Based on their scaling relations and

⁴ <https://george.readthedocs.io>

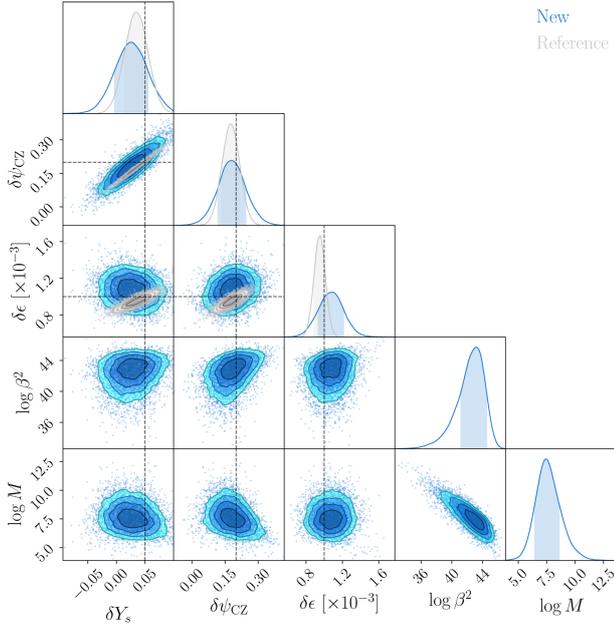


Fig. 9. MCMC estimate of the distribution $p(\delta\theta, \mu|\delta\sigma^2)$ (in blue) compared to the reference without any contamination found in Fig. 4 (in grey). The dashed lines indicate the value of $\delta\theta_{\text{true}} = [0.05, 0.2, 1 \times 10^{-3}]^T$ used to generate $\delta\sigma^2$ (the marginalised PDF over μ gives at this point: $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2) = 12.67$). The mode of this distribution is given by $\delta\theta_{\text{MAP}} = [0.027, 0.178, 1.051 \times 10^{-3}]^T$ and $\mu_{\text{MAP}} = [42.8, 7.63]^T$.

taking into account our non-dimensionalisation, we obtained the following expression:

$$\delta\sigma_{\text{surf, true}}^2(\sigma) = 2\sigma\delta\sigma_{\text{surf, true}}(\sigma) = -0.6 \times \sigma \left(1 - \frac{1}{1 + (\sigma/60)^5} \right). \quad (28)$$

This artificial frequency shift was added to the glitch studied in Sect. 4 in order to generate the ‘observed’ frequency deviation $\delta\sigma^2$. The probability distribution of the parameters given this data realisation, $p(\delta\theta, \mu|\delta\sigma^2)$, is shown in Fig. 9 (in blue scale) and compared with the one obtained in Fig. 4 in the absence of contamination (in grey-scale). As is the case in the previous section, the distribution has expanded, because of the uncertainty caused by surface effects on the one hand and the fixed amount of information that is now shared by five parameters on the other. The marginalised distribution over μ , visible in the six upper panels, shows many similarities with $p(\delta\theta|\delta y)$ in Fig. 6. In particular, we observe the more diffuse degeneracy between δY_s and $\delta\psi_{\text{CZ}}$ as well as the expansion of the distribution on the previously best-constrained parameter, $\delta\epsilon$. However, these effects occur to a lesser degree as indicated by the value of $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2) = 12.67$ being larger than $\ln p(\delta\theta_{\text{true}}|\delta y)$. In some ways, the fact that this method provides a better parameter recovery in the presence of synthetic surface effects than using second differences without any contamination emphasises the benefits of the procedure. Also, as was done with the contamination-free fit, we investigate the improvement of this recovery with the use of additional frequencies in Appendix B.

Once again, a more visual representation of this distribution is available in Fig. 10 through realisations of the models $\sigma_{\text{mod}}^2(\delta\theta, \mu) = \delta\sigma_{\text{ion}}^2(\delta\theta) + \delta\sigma_{\text{surf}}^2(\mu)$, $\delta\sigma_{\text{ion}}^2(\delta\theta)$ and $\delta\sigma_{\text{surf}}^2(\mu)$ according to the distribution $p(\delta\theta, \mu|\delta\sigma^2)$ (light curves). Figure 10 also allows us to assess the magnitude of the

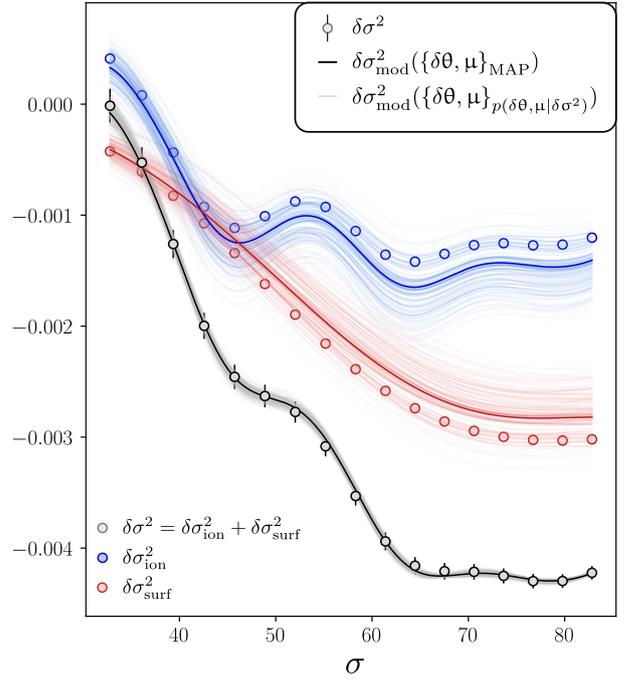


Fig. 10. Comparison of the synthetic frequency shift $\delta\sigma^2$ (black symbols) resulting from the ionisation glitch $\delta\sigma_{\text{ion}}^2$ (blue symbols), and a contamination $\delta\sigma_{\text{surf}}^2$ (red symbols) with the modelled glitch expression $\delta\sigma_{\text{mod}}^2(\delta\theta, \mu) = \delta\sigma_{\text{ion}}^2(\delta\theta) + \delta\sigma_{\text{surf}}^2(\mu)$ (respectively shown with black, blue, and red curves). For each component, the darker curve corresponds to the MAP estimate while the lighter ones result from the sampling of $p(\delta\theta, \mu|\delta\sigma^2)$.

surface effects (red symbols) compared to the glitch. In particular, their amplitude is three times higher at high frequencies, largely impacting the resulting frequency shift (black symbols) when compared to the ionisation glitch (blue symbols). Despite this, it is not surprising that the model in its entirety (black curves) manages to adequately reproduce this frequency shift considering the added flexibility it is given. Having said this, it is far more interesting to see what the model assumes about $\delta\sigma_{\text{surf}}^2(\mu)$ and especially $\delta\sigma_{\text{ion}}^2(\delta\theta)$ (red and blue curves respectively) based solely on the observation of $\delta\sigma^2$. Concerning the surface effects model, we can see in particular that the mean of the GP has clearly changed in comparison with the a priori process (cf. Fig. 8). All these curves do appear to be plausible realisations of surface effects (at least when compared to those we artificially imposed) and overall only slightly underestimate their amplitude. The shape of the glitch is also fairly well guessed, although the most likely model gradually deviates from that shape at high frequencies. We note that we obtain constraints not only on the fast oscillation but also on the slower trend of the glitch. Thus, some of the information on the hydrogen component is recovered and more can be retrieved if it is possible to obtain even lower radial orders (cf. Appendix B). The envelope formed by the sampling is naturally larger than in Fig. 5 and thus reflects the uncertainty introduced by the added contamination.

6. Conclusion

In the present paper, we develop a method for estimating the properties of the ionisation region from the glitch caused by rapid structural variations in this very region. In particular, the method presented here is designed to overcome the dependence

on realistic stellar models caused by calibration in order to provide a truly independent estimate of these properties. These include both the helium abundance and other physical quantities that can have a significant impact on the oscillation frequencies such as the electronic degeneracy parameter or the extent of the ionisation region.

This paper takes [Paper I](#) as a starting point and combines the structural perturbations of the ionisation zone presented there with the perturbation of a wave equation. To this end, we consider a framework suited to the ionisation region (purely radial oscillations in an isentropic region) and verify that the resulting formula is able to correctly reproduce glitches produced by various structural perturbations. Although it only involves the perturbation of the (normalised) cut-off frequency $\delta\sigma_c^2$, this formulation is more accurate than common glitch expressions that only consider the perturbation of the first adiabatic exponent $\delta\Gamma_1$. This gain in accuracy is especially useful for us: the resulting glitch model is thus able to exploit the information contained in the fast oscillation caused by the helium ionisation but also in the slow trend accompanying the ionisation of hydrogen. By using the work conducted in [Paper I](#), this information can be directly expressed in terms of parameters $\delta\theta = [\delta Y_s, \delta\psi_{CZ}, \delta\varepsilon]^T$, which represent a difference in helium abundance, the electronic degeneracy, and the extent of the ionisation region (from a fixed reference), respectively.

We quantify this parameter recovery through Bayesian inference in the context of limited (observable radial modes) and noisy synthetic data. By design, the ‘observed’ glitch is generated based only on structural differences localised in the ionisation region, and therefore constitutes a robust reference for evaluating the performance of the procedure in an ideal case (without any external contamination). We thus verified that the properties at the origin of the glitch can mostly be recovered. Specifically, the best constrained parameter is the extent of the region, followed by electronic degeneracy, while the most variable estimate concerns the helium abundance. At the origin of this uncertainty is a degeneracy between δY_s and $\delta\psi_{CZ}$ which was already expected in [Paper I](#) and which particularly extends over the helium abundance. Unlike the correlations between the other parameters which disappear when more data are available (see [Appendix B](#)), this degeneracy has a physical origin and does not seem to be specifically inherent to this approach. Indeed, both parameters have (locally) similar impacts within the structure of the star itself, which fundamentally makes it difficult to disentangle them. Therefore, an increase in the number of observed frequencies or even a decrease in the considered noise level does not make this degeneracy disappear but limits its extent.

We then considered how to extend the procedure to the non-ideal (but more realistic) case of a contamination of the observed glitch by an additional component, typically surface effects. A first possibility is to extend the formalism to other diagnostics such as frequency combinations. While this adaptation does not raise any fundamental problems, a strong degradation of the parameter recovery is already observed in the ideal case. We attribute this deterioration to the loss of information on the slow trend of the glitch. Indeed, while frequency combinations allow the amplitude of a slowly varying contamination to be mitigated, they also remove all trace of the glitch component resulting from the ionisation of hydrogen. We therefore considered a second possibility in order to preserve this component despite the contamination by exploiting Gaussian processes. The latter allows us to introduce an additional degree of flexibility in the model and takes into account additional components not explic-

itly defined while introducing a limited number of additional parameters (two in our case). A deterioration appears compared to the ideal case, but is mainly due to the impact of the surface effects (the amplitude of which is a few times higher than that of the glitch at high frequencies) and the fact that a fixed amount of information in the data is now used to constrain additional parameters. It was thus possible to graphically verify that the slowly varying component was partially recovered and the estimate of surface effects was plausible. We note that this fit provides better results (in the sense of cross-entropy with the target distribution) than a second differences fit in the absence of contamination.

This work now needs to be extended to the analysis of actual data and account for the accuracy and uncertainty on a retrieval of the ionisation region properties. To this end, we point out the benefit of combining a glitch model with a Gaussian process. The number of options offered by GPs is immense and the results presented here show only one particular attempt based on weak assumptions. It is clear that this procedure can be adapted to cases where the contamination is more significantly constrained based on physical grounds and can provide results accordingly. Therefore, although it remains a prospect for now, we would like to emphasise the information that this method can provide by focusing on the slow glitch trend and directly dealing with the frequencies rather than their combination.

Acknowledgements. The authors would like to thank the referee for his or her constructive report and willingness to improve and clarify the content of this article.

References

- Aigrain, S., Hodgkin, S. T., Irwin, M. J., Lewis, J. R., & Roberts, S. J. 2015, *MNRAS*, **447**, 2880
- Aigrain, S., Parviainen, H., & Pope, B. J. S. 2016, *MNRAS*, **459**, 2408
- Angus, R., Morton, T., Aigrain, S., Foreman-Mackey, D., & Rajpaul, V. 2018, *MNRAS*, **474**, 2094
- Antia, H. M., & Basu, S. 1994, *A&AS*, **107**, 421
- Baglin, A., Auvergne, M., Barge, P., et al. 2006, in *The CoRoT Mission Pre-Launch Status – Stellar Seismology and Planet Finding*, eds. M. Fridlund, A. Baglin, J. Lochard, & L. Conroy, *ESA Spec. Publ.*, **1306**, 33
- Ballot, J., Turck-Chièze, S., & García, R. A. 2004, *A&A*, **423**, 1051
- Bender, C. M., & Orszag, S. A. 1978, *Advanced Mathematical Methods for Scientists and Engineers* (New York: McGraw Hill)
- Broomhall, A. M., Miglio, A., Montalbán, J., et al. 2014, *MNRAS*, **440**, 1828
- Chandrasekhar, S. 1964, *A&A*, **139**, 664
- Christensen-Dalsgaard, J., & Perez Hernandez, F. 1992, *MNRAS*, **257**, 62
- Christensen-Dalsgaard, J., & Thompson, M. J. 1997, *MNRAS*, **284**, 527
- Deheuvels, S., & Michel, E. 2010, *Astron. Nachr.*, **331**, 929
- Dziembowski, W. A., Pamyatnykh, A. A., & Sienkiewicz, R. 1990, *MNRAS*, **244**, 542
- Gilliland, R. L., Brown, T. M., Christensen-Dalsgaard, J., et al. 2010, *PASP*, **122**, 131
- Gough, D. O. 1990, in *Comments on Helioseismic Inference*, eds. Y. Osaki, & H. Shibahashi, 367, 283
- Gough, D. O. 2002, in *Stellar Structure and Habitable Planet Finding*, eds. B. Battrick, F. Favata, I. W. Roxburgh, & D. Galadi, *ESA Spec. Publ.*, **485**, 65
- Gough, D. O., & Thompson, M. J. 1991, *The Inversion Problem*, 519
- Houdayer, P. S., Reese, D. R., Goupil, M.-J., & Lebreton, Y. 2021, *A&A*, **655**, A85
- Houdek, G., & Gough, D. O. 2007, *MNRAS*, **375**, 861
- Houdek, G., & Gough, D. O. 2011, *MNRAS*, **418**, 1217
- Kjeldsen, H., Bedding, T. R., & Christensen-Dalsgaard, J. 2008, *ApJ*, **683**, L175
- Lopes, I., Turck-Chièze, S., Michel, E., & Goupil, M.-J. 1997, *ApJ*, **480**, 794
- Lund, M. N., Silva Aguirre, V., Davies, G. R., et al. 2017, *ApJ*, **835**, 172
- Monteiro, M. J. P. F. G., & Thompson, M. J. 1998, in *New Eyes to See Inside the Sun and Stars*, eds. F. L. Deubner, J. Christensen-Dalsgaard, & D. Kurtz, *IAU Symp.*, **185**, 317
- Monteiro, M. J. P. F. G., & Thompson, M. J. 2005, *MNRAS*, **361**, 1187

- Monteiro, M. J. P. F. G., Christensen-Dalsgaard, J., & Thompson, M. J. 1994, [A&A](#), 283, 247
- Noll, A., Deheuvels, S., & Ballot, J. 2021, [A&A](#), 647, A187
- Perez Hernandez, F., & Christensen-Dalsgaard, J. 1994, [MNRAS](#), 269, 475
- Rasmussen, C. E., & Williams, C. K. I. 2006, [Gaussian Processes for Machine Learning](#) (MIT Press)
- Reese, D., Buldgen, G., & Zharkov, S. 2016, [InversionKit: Linear Inversions from Frequency Data](#)
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, [J. Astron. Telesc. Instrum. Sys.](#), 1, 014003
- Roxburgh, I. W., & Vorontsov, S. V. 1994a, [MNRAS](#), 268, 143
- Roxburgh, I. W., & Vorontsov, S. V. 1994b, [MNRAS](#), 268, 880
- Sonoi, T., Samadi, R., Belkacem, K., et al. 2015, [A&A](#), 583, A112
- Stassun, K. G., Oelkers, R. J., Paegert, M., et al. 2019, [AJ](#), 158, 138
- Tassoul, M. 1980, [ApJS](#), 43, 469
- Vandakurov, Y. V. 1967, [AZh](#), 44, 786
- Verma, K., Faria, J. P., Antia, H. M., et al. 2014, [ApJ](#), 790, 138
- Verma, K., Raodeo, K., Antia, H. M., et al. 2017, [ApJ](#), 837, 47
- Verma, K., Raodeo, K., Basu, S., et al. 2019, [MNRAS](#), 483, 4678
- Vorontsov, S. V., Baturin, V. A., & Pamiatnykh, A. A. 1992, [MNRAS](#), 257, 32

Appendix A: Details about the recovery of the posterior distributions

Here, we provide details about the Bayesian inference on which Sections 4 and 5 rely.

A.1. Fitting the frequencies using $\delta\sigma^2_{\text{mod}}(\delta\theta)$

We first provide a prior distribution on θ by decomposing it on each parameter (assumed to be a priori independent):

$$p(\theta) = \prod_i p(\theta_i) = p(Y_s)p(\psi_{\text{CZ}})p(\varepsilon), \quad (\text{A.1})$$

with $p(X)$ designating the probability density function (PDF) of the random variable X . Specifically, we use the following distribution laws

$$\begin{aligned} Y_s &\sim \text{Beta}(\mu = 0.25, s = 0.1), \\ -\psi_{\text{CZ}} &\sim \text{BetaPrime}(\mu = 5, s = 2), \\ \varepsilon &\sim \text{BetaPrime}(\mu = 0.02, s = 0.01). \end{aligned}$$

The natural parameters (α, β) of these two laws can easily be determined from our parameterisation via the mean, μ , and standard deviation, s :

$$\begin{aligned} \text{Beta} : \quad \alpha &= \mu(v_- - 1), \quad \beta = (1 - \mu)(v_- - 1) \\ \text{BetaPrime} : \quad \alpha &= \mu(v_+ + 1), \quad \beta = v_+ + 2, \end{aligned}$$

with $v_{\pm} = \mu(1 \pm \mu)/s^2$. We chose the above distributions because they are defined on appropriate domains and have advantageous properties: $0 \leq Y_s \leq 1$, $\psi_{\text{CZ}} < 0$ (the convective region in solar-like pulsators cannot be degenerate), $\varepsilon > 0$ (by design). The means given correspond more or less to solar values but with sufficient variability so as to remain rather uninformative. The prior distribution on $\delta\theta$ trivially follows from:

$$\delta\theta = \theta - \theta^\circ, \quad (\text{A.2})$$

with $\theta^\circ = [0.255, -5, 0.015]^\top$ being our (fixed) reference. We assume that the observed glitch distribution differs from the model only by an unbiased multivariate Gaussian, that is:

$$\delta\sigma^2|\delta\theta = \delta\sigma^2_{\text{mod}}(\delta\theta) + \mathbf{n} \text{ with } \mathbf{n} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\delta\sigma^2}), \quad (\text{A.3})$$

with $\delta\sigma^2_{\text{mod}}(\delta\theta)$ denoting the vector composed of all modelled (quadratic) frequency deviations from the reference given by Eq. (21). Our modelling of the noise \mathbf{n} remains very rudimentary: we assume a frequency-independent and uncorrelated error distribution on $\delta\sigma$, that is, $\Sigma_{\delta\sigma} = s_{\mathbf{n}}^2\mathbf{I}$. From there, the noise distribution is easily propagated from $\delta\sigma$ to $\delta\sigma^2$:

$$\Sigma_{\delta\sigma^2} = (2\text{diag}(\sigma))(s_{\mathbf{n}}^2\mathbf{I})(2\text{diag}(\sigma))^\top = (2s_{\mathbf{n}}\text{diag}(\sigma))^2, \quad (\text{A.4})$$

with $\text{diag}(\sigma)$ being the diagonal matrix composed of σ . From Bayes' theorem, the most plausible values of $\delta\theta|\delta\sigma^2$ minimise the following function (explicitly written in terms of the distributions introduced above):

$$-\ln p(\delta\theta|\delta\sigma^2) = -\ln p(\delta\theta) - \ln p(\delta\sigma^2|\delta\theta) + \ln p(\delta\sigma^2), \quad (\text{A.5})$$

with

$$p(\delta\sigma^2) = \int p(\delta\theta)p(\delta\sigma^2|\delta\theta) d\delta\theta, \quad (\text{A.6})$$

the marginal likelihood that can be treated as a normalising constant in this case.

A.2. Fitting the frequencies using $\delta\mathbf{y}_{\text{mod}}(\delta\theta)$

In the presence of a contamination by surface effects, the observed frequency differences will be broken down into two components (in addition to the noise \mathbf{n}):

$$\delta\sigma^2 = \delta\sigma^2_{\text{ion}} + \delta\sigma^2_{\text{surf}} + \mathbf{n}, \quad (\text{A.7})$$

where $\delta\sigma^2_{\text{ion}}$ designates the ionisation glitch and $\delta\sigma^2_{\text{surf}}$ the deviation caused by surface effects. The method presented in Section 5.1 consists in introducing the combination:

$$\delta\mathbf{y} = C\delta\sigma^2 = C\delta\sigma^2_{\text{ion}} + C\delta\sigma^2_{\text{surf}} + C\mathbf{n}, \quad (\text{A.8})$$

in the hope that $C\delta\sigma^2_{\text{surf}} \ll C\delta\sigma^2_{\text{ion}}$. In such a case, it is possible to approximate $\delta\mathbf{y}$ using the ionisation diagnostic model $C\delta\sigma^2_{\text{ion}}(\delta\theta)$ (given by Eq. (24)) together with $C\mathbf{n}$. In the case considered in the main text where $C = D^2$, the noise covariance matrix is given by:

$$\Sigma_{\delta\mathbf{y}} = D^2(2s_{\mathbf{n}}\text{diag}(\sigma))^2(D^2)^\top. \quad (\text{A.9})$$

A.3. Fitting the frequencies using $\delta\sigma^2_{\text{mod}}(\delta\theta, \mu)$

A comprehensive introduction to the theory of GPs is given in [Rasmussen & Williams \(2006\)](#) and most of the notation appearing in this Appendix is based on that paper. For instance, we write:

$$\delta\sigma^2_{\text{surf}}(\sigma) \sim \mathcal{GP}(0, k_{\text{surf}}(\sigma, \sigma')) \quad (\text{A.10})$$

to designate a model of the surface effects following the GP with mean function $m_{\text{surf}}(\sigma) = 0$ and covariance function $k_{\text{surf}}(\sigma, \sigma')$. Although it may seem surprising, the choice of a zero mean a priori is rather common and enables us to avoid expressing a parametric form of $m_{\text{surf}}(\sigma)$ explicitly. Constrained by the observations $\mathcal{D} = \{\sigma, \delta\sigma^2\}$, it is clear that this mean will no longer be zero but will instead be expressed in terms of $k_{\text{surf}}(\cdot, \sigma)$ and $\delta\sigma^2$. The explicit definition of the process in Eq. (A.10) then boils down to expressing the covariance function, $k_{\text{surf}}(\sigma, \sigma')$, which gives the covariance between any choice of two frequencies (σ, σ') . Many astrophysical applications of GPs such as corrections of systematics and detrending of K2 light curves or estimates of stellar rotation periods (e.g. [Aigrain et al. 2015, 2016](#); [Angus et al. 2018](#)) already make use of the ‘squared exponential’ (SE) and ‘quasi-periodic’ (QP) covariance functions. In this paper, in order to satisfy the constraints described above, we choose k_{surf} to be the product of a sixth order ‘homogeneous polynomial’ (HP) and a SE covariance function:

$$\begin{aligned} k_{\text{surf}}(\sigma, \sigma') &= k_{\text{HP}}(\sigma, \sigma') \times k_{\text{SE}}(\sigma, \sigma') \\ &= \frac{(\sigma\sigma')^6}{\beta^2} \times \exp\left(-\frac{M(\sigma - \sigma')^2}{2}\right). \end{aligned} \quad (\text{A.11})$$

In this function, each part acts as follows:

- The HP covariance function reduces the variance of the frequency, σ , when its value diminishes. Thus the process will take on values close to its mean (in this case 0) at low frequencies and allow a larger dispersion at higher frequencies. The ‘envelope’ of the process (in the sense of equally likely values as a function of frequency) is polynomial and varies as σ^6/β . The choice of the exponent (6) is due to the observation that the surface effects vary approximately as $\delta\sigma_{\text{surf}} \propto \sigma^5$ in the solar case ([Kjeldsen et al. 2008](#)), thus making $\delta\sigma^2_{\text{surf}} \propto \sigma^6$. This choice is somewhat arbitrary although we see that the process can widely deviate from this envelope.

- The SE covariance function reduces the covariance between two frequencies, σ and σ' , that are distant in value. This introduces a characteristic scale of variation that behaves as $\sqrt{1/M}$ and allows for a progressive departure from a strict polynomial relationship. This covariance function therefore adds flexibility to the process.

The process $\delta\sigma_{\text{surf}}^2(\sigma)$ can then be seen as parameterised by the two values β^2 and M (strictly speaking, they are hyperparameters), which influence the probability distribution of finding the value $\delta\sigma_{\text{surf}}^2(\sigma)$ at frequency σ . Both of these parameters being positive (and potentially very large), we refer to this parameterisation as $\boldsymbol{\mu} = [\log \beta^2, \log M]^T$ (and to the corresponding GP as $\delta\sigma_{\text{surf},\boldsymbol{\mu}}^2(\sigma)$). Based on this GP, our parameterised model of surface effects in Eq. (27) directly follows from the estimate of the mean process on σ for a fixed value of $\boldsymbol{\mu}$:

$$\delta\sigma_{\text{surf}}^2(\boldsymbol{\mu}) = \overline{\delta\sigma_{\text{surf},\boldsymbol{\mu}}^2(\sigma)}. \quad (\text{A.12})$$

We also adopt the notation

$$\mathbf{K}_{\text{surf}}(\boldsymbol{\mu}) = k_{\text{surf},\boldsymbol{\mu}}(\sigma, \sigma') \quad (\text{A.13})$$

to designate the covariance matrix resulting from applying $k_{\text{surf},\boldsymbol{\mu}}$ to any pair of frequencies in σ .

Given an observation of surface effects, it would then be fairly easy to deduce the process a posteriori from the GP prior (given in Eq. (A.10)). However, such a case never appears in practice. As highlighted by Eq. (A.7), the observation of a surface effect will never occur independently of that of the glitch and of a stochastic component. Even if the ionisation glitch component is purely deterministic (unlike the other two), it is useful to write the entire sum as a GP itself to exploit its properties. Thus, the process representing the entirety of the observed frequency shift can be written as:

$$\delta\sigma^2(\sigma) \sim \mathcal{GP}\left(\delta\sigma_{\text{ion},\delta\theta}^2(\sigma), k_{\text{surf},\boldsymbol{\mu}}(\sigma, \sigma') + k_n(\sigma, \sigma')\right), \quad (\text{A.14})$$

with $\delta\sigma_{\text{ion},\delta\theta}^2(\sigma)$ the prediction of the model given by Eq. (21) at σ , $k_n(\sigma, \sigma')$ the (non-stationary) white noise covariance function defined as

$$k_n(\sigma, \sigma') = (2s_n\sigma)^2\delta(\sigma' - \sigma), \quad (\text{A.15})$$

and associated with the covariance matrix $\Sigma_{\delta\sigma^2}$, $\delta(x)$ being the Dirac function. Indeed, it is clear from Eqs. (A.4) and (A.15) that we have $k_n(\sigma, \sigma) = \Sigma_{\delta\sigma^2}$. We note that the deterministic aspect of the ionisation glitch component results in the addition of a mean value to the process only without any additional covariance compared to Eq. (A.10). Also, as for $\boldsymbol{\mu}$, $\delta\theta$ can be considered here as a hyperparameter in the sense that it designates a parameter of the underlying process (its mean function) and not of a particular realisation $\delta\sigma^2$.

By noticing that $\delta\sigma^2|\delta\theta, \boldsymbol{\mu}$ simply follows the normal law $\mathcal{N}(\delta\sigma_{\text{ion}}^2(\delta\theta), \mathbf{K}_{\text{surf}}(\boldsymbol{\mu}) + \Sigma_{\delta\sigma^2})$ from Eq. (A.14), the resulting marginalised likelihood is given by:

$$\begin{aligned} \ln p(\delta\sigma^2|\delta\theta, \boldsymbol{\mu}) = & \\ & - \frac{1}{2} \left(\delta\sigma^2 - \delta\sigma_{\text{ion}}^2(\delta\theta) \right)^T \left[\mathbf{K}_{\text{surf}}(\boldsymbol{\mu}) + \Sigma_{\delta\sigma^2} \right]^{-1} \left(\delta\sigma^2 - \delta\sigma_{\text{ion}}^2(\delta\theta) \right) \\ & - \frac{1}{2} \ln |\mathbf{K}_{\text{surf}}(\boldsymbol{\mu}) + \Sigma_{\delta\sigma^2}| - \frac{N}{2} \ln 2\pi, \end{aligned} \quad (\text{A.16})$$

with N denoting the number of radial orders considered. Here, the designation ‘marginal’ may seem obscure when compared to

Eq. (A.6), where it referred to an integration over any possible $\delta\theta$. In Eq. (A.16), we refer instead to a marginalisation over the realisations of the GP for fixed values of $\delta\theta$ and $\boldsymbol{\mu}$ (hence their designation as hyperparameters). We note that this expression is analytically tractable only because we assume our frequency shift model to be a GP. Subsequently, accounting for the prior distributions, we sample the following hyperparameter posterior:

$$\begin{aligned} \ln p(\delta\theta, \boldsymbol{\mu}|\delta\sigma^2) = & \ln p(\delta\theta, \boldsymbol{\mu}) + \ln p(\delta\sigma^2|\delta\theta, \boldsymbol{\mu}) - \ln p(\delta\sigma^2) \\ = & \ln p(\delta\theta) + \ln p(\boldsymbol{\mu}) + \ln p(\delta\sigma^2|\delta\theta, \boldsymbol{\mu}) - \ln p(\delta\sigma^2), \end{aligned} \quad (\text{A.17})$$

with $p(\delta\theta)$ resulting from Eq. (A.1), $p(\boldsymbol{\mu})$ being taken as an improper (in the sense of Bayesian inference) uniform prior, $p(\delta\sigma^2|\delta\theta, \boldsymbol{\mu})$ being given by Eq. (A.16) and

$$p(\delta\sigma^2) = \iint p(\delta\theta)p(\boldsymbol{\mu})p(\delta\sigma^2|\delta\theta, \boldsymbol{\mu}) d\delta\theta d\boldsymbol{\mu} \quad (\text{A.18})$$

being the normalising constant.

Appendix B: Parameter recovery with additional frequencies

Considering all the distribution sampling performed so far, it may be interesting to try to quantify the impact of the various nuisances on the parameter recovery. In our case, besides potential imperfections of the model we presented, a clearly identifiable source of error lies in the absence of very low degree frequencies. Indeed, these frequencies contain the most information on the slowly varying component of the glitch, which we expect to be essential for a complete retrieval of the parameters. To this end, we present in Figs. B.1 and B.2 the sampling of the distributions shown in Figs. 4 and 9 with additional frequencies corresponding to $7 \leq n \leq 10$ (the noise level on the oscillation frequencies, s_n , remains the same).

Figure B.1 shows the significant improvement in the parameter recovery when fitting the frequencies $\delta\sigma^2$ in the ideal case (to be compared with the distribution found in Fig. 4, shown in grey). The mode, $\delta\theta_{\text{MAP}} = [0.047, 0.200, 0.989 \times 10^{-3}]^T$, is now much closer to the true value $\delta\theta_{\text{true}}$ and the extent of the distribution in general has considerably narrowed, now giving $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2) = 16.91$. We also note the clearly distinct nature of the $\delta Y_s - \delta\psi_{\text{CZ}}$ degeneracy with its physical origin and the correlation of these two parameters with $\delta\varepsilon$. The first one, although shortened, conserves the same shape while the second one completely disappears to give fully uncorrelated parameters. Therefore, if the latter is indeed the result of a lack of information on the slow component associated with hydrogen (cf. Sections 4 and 5), we witness here the fundamentally different nature of the $\delta Y_s - \delta\psi_{\text{CZ}}$ degeneracy. If its extent can be reduced as is done here (a decrease in the noise level s_n also diminishes its size), any glitch estimate of the helium abundance will be confronted with this elongated shape, which results from an ambivalence within the structure of the star itself.

Figure B.2 shows a similar test for the case of a contamination by surface effects. We therefore choose the distribution $p(\delta\theta, \boldsymbol{\mu}|\delta\sigma^2)$ as a reference, shown in Fig. 9 and appearing here in grey. Interestingly, the additional frequencies do not have any impact on the distribution marginalised over δY_s and $\delta\psi_{\text{CZ}}$ (6 rightmost panels) where the new distribution perfectly overlaps with the reference. This new information only effectively constrains the distribution on δY_s and $\delta\psi_{\text{CZ}}$ (visible in the uppermost bivariate distribution) and the latter is largely improved.

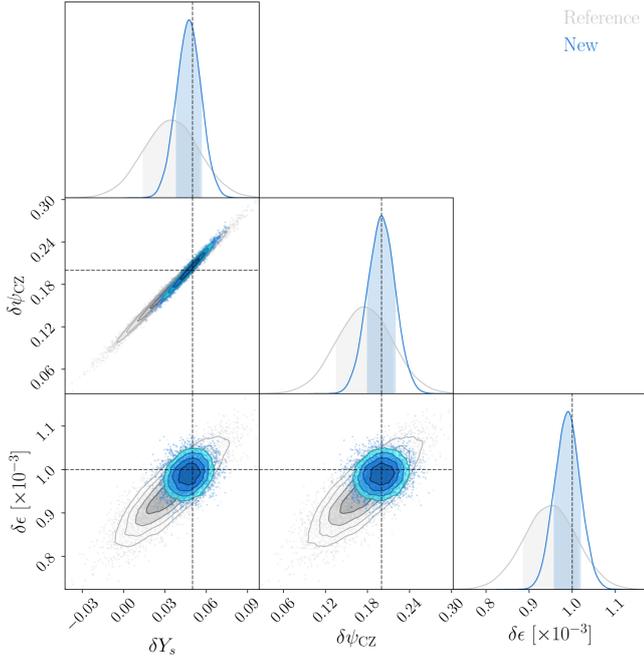


Fig. B.1. MCMC estimate of the distribution $p(\delta\theta|\delta\sigma^2)$ with the additional low order frequencies $7 \leq n \leq 10$ (in blue) in comparison with the reference shown in Fig. 4 (in grey). The dashed lines indicate the value of the true $\delta\theta_{\text{true}} = [0.05, 0.2, 1 \times 10^{-3}]^T$ used to generate $\delta\sigma^2$ (the PDF gives at this point: $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2) = 16.91$). The mode of this distribution is $\delta\theta_{\text{MAP}} = [0.047, 0.200, 0.989 \times 10^{-3}]^T$.

The MAP estimate, $\delta\theta_{\text{MAP}} = [0.046, 0.214, 1.083 \times 10^{-3}]^T$, has moved closer to $\delta\theta_{\text{true}}$ and the distribution extent has narrowed, resulting in a value of $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2) = 14.37$.

We therefore emphasize the importance of these very low radial orders, in particular when constraining the physical quantities, especially the helium abundance and the electronic

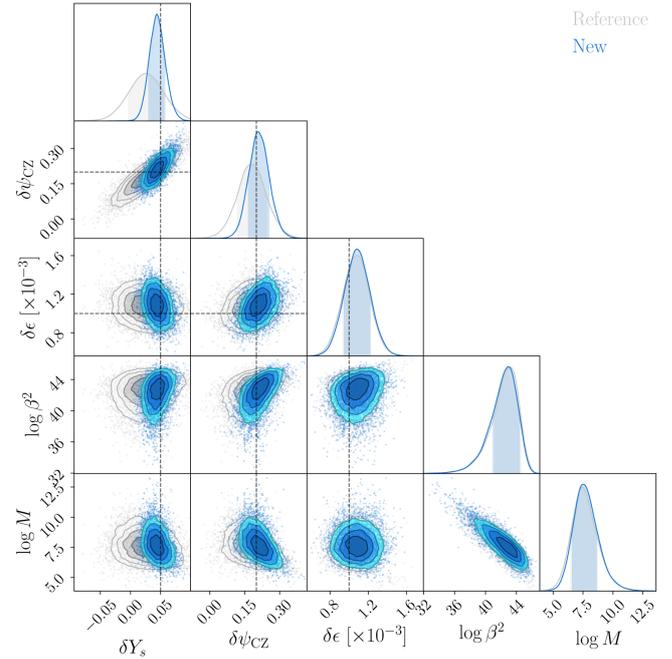


Fig. B.2. MCMC estimate of the distribution $p(\delta\theta, \mu|\delta\sigma^2)$ with the additional low order frequencies $7 \leq n \leq 10$ (in blue) in comparison with the reference shown in Fig. 9 (in grey). The dashed lines indicate the value of $\delta\theta_{\text{true}} = [0.05, 0.2, 1 \times 10^{-3}]^T$ used to generate $\delta\sigma^2$ (the marginalised PDF over μ gives at this point: $\ln p(\delta\theta_{\text{true}}|\delta\sigma^2) = 14.37$). The mode of this distribution is given by $\delta\theta_{\text{MAP}} = [0.046, 0.214, 1.083 \times 10^{-3}]^T$ and $\mu_{\text{MAP}} = [43.2, 7.57]^T$.

degeneracy parameter. Let us note that adding high radial orders (> 27), on the other hand, hardly changes the retrieved distributions.

Chapitre 6

Conclusion et perspectives

Contenu du chapitre

6.1	Modélisation de la région d'ionisation	151
6.2	Modélisation du glitch de l'ionisation	152
6.3	Perspectives	154

Ce travail de thèse s'inscrit dans la recherche de contraintes pour la modélisation stellaire et ce par l'exploitation des données sismiques maintenant à disposition pour ces astres. Plus particulièrement, c'est ici la question de la détermination de l'abondance en hélium des étoiles de type solaire qui est traitée, enjeu transverse à de nombreux pans de la physique stellaire et aux implications multiples : caractérisation des populations stellaires, contraintes sur la physique interne des étoiles ou encore détermination de leur masse par la levée de la dégénérescence masse-hélium (cf. chapitre 1). Dans cette optique, l'exploitation des glitches sismiques, et en particulier celle du glitch de l'ionisation, semble vouée à jouer un rôle essentiel dans un contexte où les données photométriques se sont faites toujours plus précises (par CoRoT (BAGLIN et al., 2002; BAGLIN et al., 2009), *Kepler* (GILLILAND et al., 2010; LUND et al., 2017) puis TESS (RICKER et al., 2015; STASSUN et al., 2019)) et qui devrait bientôt voir advenir l'ère PLATO (RAUER et al., 2014). Néanmoins, une analyse de la région d'ionisation et des signatures sismiques induites fait apparaître un problème complexe et dual par essence, requérant une étude approfondie de la structure stellaire comme des modes d'oscillation qui s'y développent (cf. chapitres 2 et 3). L'exploitation du glitch de l'ionisation repose ainsi sur un certain nombre d'hypothèses qui définissent les limites de cette méthode telle qu'actuellement formulée (cf. section 3.6), elles-mêmes à l'origine des objectifs auxquels ce travail cherche à répondre.

6.1 Modélisation de la région d'ionisation

Une des priorités établies était la mise au point d'une modélisation physique de la région d'ionisation dans le but de dispenser les estimations de l'abondance en hélium de calibration sur des modèles stellaires réalistes. Au moyen de relations thermodynamiques et d'une approximation aux équations de Saha, nous avons ainsi proposé une approximation analytique du premier exposant adiabatique, Γ_1 , dans la région d'ionisation. Celle-ci s'exprime comme une fonction des conditions de densité ρ et de température T mais aussi des abondances de chaque élément X_j . Alliée à une stratification cohérente satisfaisant aux équilibre hydrostatique et équation de Poisson et dans un mélange d'hydrogène et d'hélium, le modèle de la région d'ionisation fait alors intervenir trois paramètres. Ces trois grandeurs peuvent ainsi être vues comme impactant le plus la structure de cette région localement. Il s'agit de l'abondance en hélium de la région (notée Y_s), du paramètre de dégénérescence électronique dans la zone convective (ψ_{CZ}) ainsi que de l'étendue de la région d'ionisation (ε). Chacune de ces quantités impacte sensiblement la structure de la région et il est dès lors possible d'interpréter les variations d'une région d'ionisation à l'autre

comme la conséquence de changements parmi ces trois propriétés physiques. Cette conception tranche ainsi avec une hypothèse commune selon laquelle les variations d'abondances δY constitueraient la principale origine au glitch de l'ionisation.

Par l'étude du premier exposant adiabatique dans la région d'ionisation, nous avons pu évaluer l'incidence de chaque paramètre sur la région. Ainsi, l'abondance en hélium Y_s influe majoritairement sur l'énergie localement allouée à l'ionisation de l'hélium, ce qui se traduit par des variations d'aire dans les puits d'ionisation du profil de Γ_1 . La dégénérescence électronique ψ_{CZ} joue sur la "vitesse" à laquelle s'effectue chaque ionisation (hydrogène (H), première (HeI) et seconde (HeII) ionisation de l'hélium) et impacte donc la dispersion des puits. Enfin, l'étendue de la région ε fait naturellement varier les positions où les ionisations prennent place dans la structure. Ainsi, la modification de ces paramètres, et donc des amplitudes, étendues et positions des puits d'ionisation, permet de produire une grande diversité de régions d'ionisation de façon réaliste tout en les dotant d'une interprétation physique.

Notant θ ces trois paramètres, l'étude des perturbations de structures se déduit directement des variations induites dans la structure par un changement de propriétés $\delta\theta$. Nous nous sommes proposé d'étudier plus particulièrement les perturbations induites au niveau du premier exposant adiabatique $\delta_t\Gamma_1$. Ainsi, les variations causées par un changement d'abondance en hélium δY_s ont mis en évidence des profils complexes en comparaison des fonctions gaussiennes généralement utilisées pour modéliser son impact, caractère s'accroissant lorsque ces variations prennent place dans des modèles à faible dégénérescence (et donc plus massifs). Par rapport aux gammes de valeurs caractéristiques des pulsateurs de type solaire, les variations en dégénérescence électronique $\delta\psi_{CZ}$ ainsi qu'en étendue de la région $\delta\varepsilon$ ont ici un impact notable. Les profils de $\delta_t\Gamma_1$ engendrés sont complexes, soulignant dans un premier temps l'intérêt d'une approche physique plutôt que visuelle. Mais surtout, les variations $\delta\psi_{CZ}$ et $\delta\varepsilon$ entraînent des perturbations de plus fortes amplitudes que celles causées par δY_s . Il est dès lors à considérer que ces deux paramètres puissent avoir un impact plus prononcé que l'abondance d'hélium elle-même sur la structure, d'où la nécessité de les prendre en compte. Un point intéressant et commun aux impacts des trois paramètres est la présence d'une variation très abrupte et locale de la structure à l'entrée de la région d'ionisation de l'hydrogène. Une contrepartie fréquentielle à cette signature, un glitch de l'hydrogène est donc à attendre dans les fréquences propres.

Notons que cette modélisation de la région permet d'avancer des explications aux résultats obtenus par différentes études antérieures. Elle donne d'abord une raison physique à la forte augmentation de l'amplitude du glitch avec la masse des modèles stellaires (VERMA et al., 2014a), puisque cette dernière varie de façon monotone avec la dégénérescence électronique, ψ_{CZ} , de la région. Enfin, les perturbations de structure engendrées par ce modèle fournissent également des explications aux localisations de l'origine du glitch en dehors des puits d'ionisation (BROOMHALL et al., 2014; VERMA et al., 2014a).

6.2 Modélisation du glitch de l'ionisation

Lier la signature sismique du glitch de l'ionisation aux perturbations de structure, et donc aux travaux déjà menés, nécessite le passage par l'étude des modes d'oscillation dans la région d'ionisation. À cet égard, nous avons cherché à obtenir une expression permettant d'exploiter le modèle de perturbation mis au point, provenant donc de profils de perturbation arbitraires et non pas causés par une seule variation d'abondance δY_s mais de l'ensemble des paramètres $\delta\theta$. À partir d'une équation d'onde obtenue dans la limite d'oscillations quasi-radiales (adaptée pour l'étude des modes de bas degrés, les seuls visibles pour les étoiles autres que le soleil), une forme intégrale du glitch a pu être dérivée. Celle-ci ne fait intervenir que la perturbation du potentiel acoustique et diffère ainsi des formules traditionnelles s'exprimant sur les perturbations de deux variables. Les travaux préalablement menés permettant de relier une variation des propriétés de la région d'ionisation à une variation du potentiel acoustique, nous disposons dorénavant d'un

lien direct entre les propriétés structurelles et fréquentielles. L'expression du glitch obtenue ne requérant pas de se situer à hautes fréquences, nous avons pu vérifier que celle-ci était bien à même de reproduire l'écart en fréquence obtenu numériquement dans l'ensemble du spectre. L'observation du glitch de l'ionisation a également permis de mettre en évidence deux composantes distinctes aux comportements opposés. La première a pour origine les régions de l'ionisation de l'hélium, possède une période courte (dans le domaine fréquentiel) et s'atténue rapidement. La seconde exhibe une variation beaucoup plus lente, qui persiste à hautes fréquences. Nous lui donnons ainsi pour origine la région d'ionisation de l'hydrogène et en particulier la variation très localisée à l'entrée de la région observée dans les perturbations de structure.

Afin de quantifier les capacités prédictives de la méthode alors définie, nous avons cherché à la tester dans un cadre idéal. On qualifie ici d'idéale la seule présence du glitch de l'ionisation dans les fréquences synthétiques considérées, malgré une gamme de fréquence et une incertitude représentatives de véritables données *Kepler*. Le cadre ainsi défini nous permet donc de quantifier précisément la perte d'information lors de l'ajustement du glitch par notre forme intégrale et comment cette perte se traduit sur l'estimation des propriétés de la région, θ . La quantification a été menée dans un formalisme Bayésien et permet d'établir plusieurs constats. Il est d'abord à noter que la plus grande part de l'information sur les propriétés de la région d'ionisation peuvent être retrouvées du glitch de l'ionisation. Celle-ci se répartit néanmoins de façon inégale entre les différents paramètres. Une variation en étendue de la région d'ionisation $\delta\epsilon$ semble ainsi mieux contrainte qu'un changement en dégénérescence électronique $\delta\psi_{CZ}$, lui-même mieux contraint qu'une différence en abondance d'hélium δY_s . À l'origine de ce phénomène, une forte dégénérescence entre δY_s et $\delta\psi_{CZ}$, s'étendant particulièrement sur δY_s . Si les autres paramètres sont également corrélés deux à deux, cette dégénérescence semble de nature différente. Là où les corrélations entre paramètres disparaissent avec l'ajout de quelques fréquences propres dans les données, la forme de la dégénérescence persiste même si son étendue décroît. On attribue à cette dégénérescence une origine plus profonde, au sein de la structure elle-même, où les perturbations causées par δY_s et $\delta\psi_{CZ}$ prennent des apparences comparables sur les puits d'ionisation de l'hélium.

Notre analyse s'est ensuite portée sur un cas plus réaliste en considérant une composante autre que le glitch de l'ionisation dans les écarts de fréquence. En l'occurrence, cette composante additionnelle a été choisie pour refléter l'impact d'effets de surface, la contamination du signal que nous supposons être la plus problématique dans l'optique d'une application à des données réelles. De par leur origine superficielle dans l'étoile, les effets de surface sont en effet les plus à même de produire une signature sismique difficilement dissociable du glitch de l'ionisation et en particulier de la composante de l'hydrogène.

Deux approches ont alors été considérées pour permettre une estimation de $\delta\theta$ malgré la présence d'une contamination par les effets de surface. La première est l'utilisation d'un diagnostic sismique afin d'éliminer l'impact des effets de surface. S'il est facile d'adapter la méthode présentée à l'exploitation d'un diagnostic sismique comme les secondes différences plutôt que des fréquences elles-mêmes, le test de cette procédure sur des données synthétiques se heurte à des difficultés. En réalité, en présence d'effets de surface synthétiques, la procédure ne converge pas vers une estimation de $\delta\theta$. En fait, un test dans le cadre idéal précédemment considéré fait déjà apparaître une perte d'information importante lors de l'utilisation des secondes différences par rapport aux fréquences elles-mêmes. La raison principale est la disparition de la composante de l'hydrogène des secondes différences puisque cette composante constitue elle-même une variation lente dans les fréquences propres à l'instar des effets de surface. Nous pensons donc qu'une grande part de l'information sur la structure est à attribuer à cette composante. Cet aspect est mis en évidence par l'amélioration drastique des estimations lors de l'ajout de basses fréquences dans les données (là où la composante de l'hydrogène est la plus prononcée).

La seconde approche vise donc à préserver cette composante de l'hydrogène en s'intéressant directement aux fréquences d'oscillation. Il est alors nécessaire de proposer une modélisation des effets de surface et l'on a utilisé pour cela un Processus Gaussien (GP). L'avantage du GP, par

comparaison à un modèle polynomial par exemple, réside dans sa grande flexibilité pour une introduction de paramètres moindre. S'il les résultats sont quelque peu dégradés par rapport à ceux obtenus dans le cadre idéal, ces impacts sont inhérents à la contamination du signal introduite ainsi qu'à la répartition d'une même quantité d'information sur 5 (et non plus 3) paramètres. En particulier, on notera que les estimations obtenues en présence d'effets de surface traités par un GP sont meilleures que celles déduites du cadre idéal par l'exploitation des secondes différences. Ainsi, bien que nécessitant un travail de modélisation préalable (ici, l'introduction d'un GP), nous considérons que cette seconde approche est à privilégier dans l'optique d'une exploitation de données réelles.

6.3 Perspectives

La perspective directe de la modélisation du glitch de l'ionisation mise au point durant ce travail de thèse est son application aux données réelles, et en particulier à l'échantillon d'étoiles LEGACY de *Kepler*. De par le formalisme introduit, cette méthode est à même d'apporter des contraintes sur l'abondance en hélium de ces étoiles et donc d'un vif intérêt pour l'estimation de leur masse et la compréhension des mécanismes régissant leurs intérieurs (cf. chapitre 1). La procédure exposée ne se limite cependant pas à la seule estimation de l'abondance en hélium, la modélisation physique a également permis d'établir un lien entre le glitch de l'ionisation et l'étendue de la région, ε , ainsi qu'avec la dégénérescence électronique de la zone convective ψ_{CZ} . Des analyses préliminaires de la structure de modèles réalistes indiquent notamment une variation de la dégénérescence électronique avec le choix du modèle de diffusion en plus de sa forte dépendance à la masse de l'étoile, là où l'étendue de la région d'ionisation semble sensible à l'âge des étoiles de masse $> 1M_{\odot}$. Ainsi, l'estimation des trois propriétés de la région d'ionisation dans leur ensemble pourrait fournir de fortes contraintes structurelles sur la physique interne.

TABLE 6.1 – Comparaison des hypothèses et limites entre l'approche présentée dans le chapitre 3 (cf. TABLE 3.1) et celle introduite dans cette thèse pour l'exploitation du glitch de l'ionisation

	Hypothèses / Limites (approches précédentes)	Hypothèses / Limites (dans cette thèse)
1	Développement asymptotique à l'ordre 0	Modes de bas degrés ℓ
2	Variations suffisamment lentes dans la structure interne	- (*)
3	Relation $A_{II} = f(Y)$ établie par calibration	-
4	Impact des autres régions d'ionisation (H, HeI) négligées	-
5	Nombre de paramètres non physiques	Seuls 3 paramètres physiques
6	Choix de variable à fixer différent entre perturbation de structure et kernel	-
7	Perturbation de la pression $\delta_x p / p$ négligée devant $\delta_x \Gamma_1 / \Gamma_1$	-
8	Modèle de référence non établi	-
9	Persistance du régime de perturbation linéaire	-

(*) : Le "-" indique que la méthode ne requiert pas l'hypothèse / n'est pas concernée par la limite en question.

La question est alors de savoir à quel point ces propriétés peuvent être contraintes lors de l'application à des données réelles. Au regard des résultats obtenus dans le cadre idéal étudié, la modélisation du glitch proposée me semble être une approximation du glitch de l'ionisation satisfaisante au premier ordre. La majeure partie de l'information sur la structure a pu être retrouvée et les incertitudes plus prononcées sur l'abondance en hélium sont la conséquence directe d'un manque d'information à basse fréquence, sur lequel le choix du modèle ne peut avoir d'influence. D'un point de vue théorique ensuite, la modélisation du glitch introduite permet de

répondre aux objectifs soulevés dans la section 3.6 et, en particulier, améliore les approches précédentes sur l'ensemble des points soulevés en TABLE 6.1 (cf. TABLE 3.1). Il serait bien sûr possible de proposer une modélisation plus exacte de la région d'ionisation (en prenant en considération la pression coulombienne lors de l'ionisation, par exemple), mais les gains en exactitude risquent d'être marginaux là où l'imprécision due à des paramètres supplémentaires (en l'occurrence, une échelle pour la longueur de Debye, BRAUS, 2019) augmente elle nécessairement.

Un aspect plus problématique me semble résider dans l'interaction du glitch de l'ionisation avec d'autres composantes présentes dans les différences de fréquence. En prenant encore une fois l'exemple des effets de surface, le test que l'on a mené reste, bien que plus réaliste qu'en leur absence, idéal en soi. En effet, on a ajouté dans ce cadre des effets de surface synthétiques aux glitch de l'ionisation mais supposé aucun couplage entre ces deux composantes. Or, l'étude de modèles réalistes et "patchés" par l'ajout d'une enveloppe issue de simulation 3D fait apparaître une forte interaction entre les phénomènes à l'origine des effets de surface et la région d'ionisation de l'hydrogène (SONOI et al., 2015), ce qui est à attendre de par leur proximité dans la structure stellaire. Il est à donc à prévoir que la composante de l'hydrogène ne puisse être distinguée des autres effets de surface, voir que notre modèle l'estime de façon incorrecte si l'effet du couplage est dominant.

Cette analyse souligne l'intérêt d'une prise en compte simultanée de ces effets lors de la modélisation. Tout comme cela a été fait pour cette étude, où les trois zones d'ionisation (H), (HeI) et (HeII) apparaissent dans un même modèle ne contenant que trois paramètres (là où il en aurait fallu 9 en les approximant par trois gaussiennes), il me semble qu'une prise en compte de ce couplage serait d'un grand intérêt. Notons que la perspective d'une prise en compte simultanée de plusieurs composantes dans les différences de fréquence pourrait aussi permettre une meilleure détermination des propriétés de la région d'ionisation. En prenant cette fois-ci l'exemple du glitch de la base de la zone convective (BCZ), il paraît vraisemblable que la profondeur de cette transition, τ_{BCZ} , ne varie pas indépendamment de l'étendue de la région d'ionisation ε par exemple. L'élaboration d'un modèle regroupant les deux glitches de l'ionisation et de la BCZ pourrait alors permettre une détermination d'autant plus précise du paramètre ε , maintenant contraint par deux processus physiques. De manière générale, la prise en compte simultanée de plusieurs processus physiques dans les modèles donne comme perspective l'élaboration de modèles complexes mais faiblement paramétrés (ceux-ci se recoupant) et donc extrêmement prédictifs. Une telle exploitation se rapproche alors de l'inversion globale (cf. paragraphe 2.5.5), mais sur une base de fonctions établie physiquement et moins nombreuses car correspondant au nombre de paramètres du modèle¹.

La mise au point de ces modèles passe cependant par une modélisation physique des mécanismes en jeu et donc des travaux théoriques poussés sur des sujets complexes. La prise en compte des effets de surface pourrait passer par une modélisation de la pression turbulente pour étudier son interaction avec la région d'ionisation de l'hélium, ces deux phénomènes influençant la stratification effective de la région et donc l'exposant γ . Pour ce qui est de l'inclusion du glitch de la BCZ, une modélisation s'inspirant des travaux de REMPEL, 2004 me paraît prometteuse, ceux-ci présentant déjà une modélisation paramétrée de la structure et incluant γ .

1. Notons qu'il s'agit là de la façon dont on a introduit le glitch sismique dans ce manuscrit, une inversion paramétrée (cf. paragraphe 2.5.6).

Chapitre 7

Conclusion and perspectives [en]

Content

7.1	Modelling of the ionisation region	157
7.2	Modelling the ionisation glitch	158
7.3	Perspectives	160

The work carried out during this thesis work is part of ongoing efforts to search for constraints for stellar modelling by exploiting the seismic data now available for solar-type stars. More precisely, the question of helium abundance determination in these stars is addressed here, an issue transverse to many areas of stellar physics and with multiple implications : characterising stellar populations, constraining the internal physics of stars, or even determining their mass by removing of the mass-helium degeneracy (cf. chapter 1). In this regard, it seems that the exploitation of seismic glitches, and in particular that of the ionisation glitch, is bound to play an essential role as photometric data become ever more precise (e.g. with CoRoT (BAGLIN et al., 2002 ; BAGLIN et al., 2009), *Kepler* (GILLILAND et al., 2010 ; LUND et al., 2017), TESS (RICKER et al., 2015 ; STASSUN et al., 2019) and the forthcoming PLATO era (RAUER et al., 2014)). However, an analysis of the ionisation region and its induced seismic signatures brings to light a complex and inherently dual problem that requires an in-depth study of both the stellar structure and the oscillation modes that develop within (cf. chapters 2 and 3). The exploitation of ionisation glitches thus relies on a number of assumptions that, together, set the limits of the ionisation-glitch exploitation as it is currently formalised (cf. section 3.7). These limits are the starting point from which the objectives of this thesis are drawn.

7.1 Modelling of the ionisation region

One of the established priorities is the development of a physical modelling of the ionisation region, in order to avoid estimating helium abundance through calibration on realistic stellar models. Through the use of thermodynamic relations together with an approximation of the Saha equations, we propose an analytical approximation of the first adiabatic exponent, Γ_1 , in the ionisation region. The latter is expressed as a function of the density ρ , the temperature T , and the abundances of each element X_i . Coupled with a consistent stratification that satisfies hydrostatic equilibrium and the Poisson equation, and considering a mixture of hydrogen and helium, our model of the ionisation region then involves three parameters, namely the helium abundance (denoted Y_s), the electronic-degeneracy parameter in the convection zone (ψ_{CZ}) and the extent of the ionisation region (ϵ). Each of these quantities has a significant impact on the local structure, and it is therefore possible to interpret variations from one ionisation region to another as the consequence of changes in these three physical properties. This conception thus contrasts with the common assumption, that the variations of abundances δY are the main cause of the ionisation glitch.

By studying the first adiabatic exponent, we were able to assess the impact of each parameter on the ionisation region. The helium abundance mainly affects the energy that is locally allocated to helium ionisation, which in turn leads to variations in the surface area of the ionisation wells of the Γ_1 profile. The electronic degeneracy ψ_{CZ} influences the "rate" at which each ionisation takes place (hydrogen (H), first (HeI) and second (HeII) helium ionisation), and thus impacts the spread of the ionisation wells. Finally, the extent of the ionisation region ε naturally shifts the locations where the ionisations occur within the structure. The modification of these parameters (and thus of the amplitudes, spreads and positions of the ionisation wells) therefore enables us to model in a realistic way a large diversity of ionisation regions, while providing them with a physical interpretation.

If we denote these three parameters as θ , the study of the structural perturbations directly stems from the variations induced by a change of properties $\delta\theta$ in the structure. In particular, we studied the perturbations of the first adiabatic exponent $\delta_t\Gamma_1$ resulting from this change. In this regard, the variations caused by a change of helium abundance δY_s lead to some complex profiles compared to the Gaussian functions usually used to model its impact. This complexity is further accentuated when the variations take place in models with a low degeneracy (i.e. for more massive stars). Compared to the characteristic values of solar-like pulsators, the variations in electronic degeneracy $\delta\psi_{CZ}$ as well as the region's extent $\delta\varepsilon$ both have a significant impact. The resulting $\delta_t\Gamma_1$ profiles are complex, underlining the benefits of a physical approach compared to a visual one. More particularly, variations in $\delta\psi_{CZ}$ and $\delta\varepsilon$, typical of solar-like pulsators, produce $\delta_t\Gamma_1$ profiles of greater amplitude than those caused by δY_s variations. Hence, these two parameters may have a more pronounced impact on the structure than the helium abundance itself, and it is thus necessary to take them into account. An interesting point, common to the impacts of the three parameters, is the presence of a very steep local variation of the structure at the beginning of the hydrogen ionisation region. A frequency response of this signature, a hydrogen glitch, is therefore to be expected in the star's oscillation frequencies.

We note that this modelling of the ionisation region enables us to provide explanations for the results obtained by previous studies. First, it gives a physical reason for the strong increase of the glitch amplitude with the mass of stellar models (VERMA et al., 2014a), as indeed the latter varies monotonically with the electronic degeneracy ψ_{CZ} . In addition, the structural perturbations produced by this model also explain why the glitch can sometimes be localised outside of the ionisation wells (BROOMHALL et al., 2014; VERMA et al., 2014a).

7.2 Modelling the ionisation glitch

To relate the seismic signature of the ionisation glitch to the structure perturbations (and hence to previous works) requires studying of the oscillation modes in the ionisation region. In this regard, we searched for an expression that enables us to exploit the previously developed perturbation model, an expression applicable to arbitrary perturbation profiles, not caused by a variation in abundance δY_s alone but rather involving the complete set of parameters $\delta\theta$. From a wave equation obtained in the limit of quasi-radial oscillations (which is suitable for studying low degree modes, the only ones visible in stars other than the Sun), we derived an integral form of the glitch. The latter only involves the perturbation of the acoustic potential, and thus differs from traditional formulae that usually involve the perturbation of two variables. As our previous work allows us to relate a variation of properties in the ionisation region to a variation of the acoustic potential, we now have at our disposal a direct link between the structural and frequency properties. As the obtained glitch expression does not require remaining in the high-frequency limit, we could verify that it indeed reproduces the frequency shift obtained numerically in the whole spectrum. The observation of the ionisation glitch also highlighted two distinct components presenting opposite behaviours. The first originates from the helium ionisation regions, is characterised by a short period (in the frequency domain) and is quickly damped. The second is

characterised by a much slower variation that is still seen at high frequencies. We trace it back to the hydrogen ionisation region, and particularly the very localised variation at the entrance of the region that is observed in the structural perturbations.

In order to quantify the predictive capacity of our method, we then tested it in an ideal case. Here, "ideal" refers to the fact that only an ionisation glitch is present in the synthetic frequencies, despite a frequency range and an uncertainty that are representative of actual *Kepler* data. The framework thus defined allows us to quantify the loss of information occurring during the adjustment of the glitch by our integral form, as well as how this loss affects estimates of the region's properties, θ . We carry out this evaluation using a Bayesian formalism, thus allowing us to make several observations. First, we note that the major part of the information regarding the ionisation region's properties can be found in the ionisation glitch. The latter is however split unevenly between the different parameters. A variation in the extent of the ionisation region $\delta\epsilon$ thus seems better constrained than a change in electronic degeneracy $\delta\psi_{CZ}$, itself better constrained than a difference in helium abundance δY_s . This stems from a strong degeneracy between δY_s and $\delta\psi_{CZ}$, particularly spread out on δY_s . While the other parameters are also correlated two by two, this degeneracy seems to be of a different nature. Indeed, where parameter correlations tend to disappear with the addition of few oscillation frequencies in the data, the shape of this degeneracy persists even though its extent decreases. Its origin is probably deeper, from within the structure itself, where the perturbations caused by δY_s and $\delta\psi_{CZ}$ take on comparable appearances on the helium ionisation wells.

Our analysis then focused on a more realistic case by considering an additional component – unrelated to the ionisation glitch – in the frequency deviations. This component was chosen to reflect the impact of surface effects, which is the contamination we take to be the most problematic while analysing real data. Indeed, because of their superficial origin, these surface effects are the most susceptible to produce a seismic signature that will not be easily disentangled from the ionisation glitch, particularly its hydrogen component.

Two approaches have been considered to allow us to estimate $\delta\theta$ despite this surface-effect contamination. The first one is the use of a seismic diagnostic to remove the impact of surface effects. While the method we presented is easily adaptable to the exploitation of a seismic diagnostic like the second differences, its actual application to synthetic data encounters some difficulties. Indeed, with the presence of synthetic surface effects, the procedure does not converge toward an estimate of $\delta\theta$. In fact, even in the ideal case previously considered, the use of the second differences leads to an important loss of information compared to the direct use of the frequencies. The principal reason is the disappearance of the hydrogen component in the second differences, since this component is itself a slow variation within the oscillation frequencies just like the surface effects. We therefore believe that a large part of the structural information can be attributed to this component. This aspect is highlighted by the drastic improvement of the estimates when adding low frequencies to the data (where the hydrogen component is the most pronounced).

The second approach thus aims to preserve this hydrogen component by addressing the oscillating frequencies directly. It is then necessary to propose a modelling of the surface effects, and for that purpose we used a Gaussian Process (GP). The advantage of a GP, compared to a polynomial model for instance, lies in its high flexibility while introducing few (free) parameters. If the results are somewhat degraded compared to those obtained in the ideal framework, these impacts are inherent to the signal contamination introduced, as well as to the distribution of the same amount of information on 5 parameters rather than 3. In particular, we note that the estimates we obtain in the presence of surface effects addressed by a GP are better than those deduced from the ideal case when exploiting the second differences. Therefore, and even though a modelling work was required beforehand (here the introduction of a GP), we consider that this second approach should be favoured with a view to exploiting real data.

7.3 Perspectives

The direct perspective of the ionisation-glitch modelling that we devised during this thesis is its application to real data, and in particular to *Kepler's LEGACY* stellar sample. Thanks to the formalism introduced, this method is able to constrain the helium abundance of these stars, and therefore is highly useful for estimating their mass together with understanding the mechanisms governing their interiors (cf. chapter 1). The procedure presented here is not limited to only estimating the helium abundance, the physical modelling also allows us to establish a link between the ionisation glitch, the extent of the region, ε , and the electronic degeneracy in the convective zone ψ_{CZ} . Preliminary analyses of the structure of realistic models indicate in particular a variation of the electronic degeneracy with the choice of diffusion model in addition to its strong dependence on the mass of the star, while the extent of the ionisation region seems to be sensitive instead to the age of stars with a mass $> 1M_{\odot}$. Therefore, simultaneously estimating the three properties of the ionisation region could provide strong structural constraints on the internal physics of stars.

TABLE 7.1 – Comparison of the assumptions and limits of previous approaches presented in chapter 3 (cf. TABLE 3.1) with the one introduced in this thesis for the exploitation of the ionisation glitch.

	Assumptions / Limits (previous approaches)	Assumptions / Limits (in this thesis)
1	o^{th} order asymptotic expansion	Low degree modes ℓ
2	Sufficiently slow variations in the internal structure	- (*)
3	$A_{\text{II}} = f(Y)$ relation established through calibration	-
4	Impact of other ionisation regions (H, HeI) neglected	-
5	Number of non-physical parameters	3 physical parameters only
6	Inconsistent choice of the variable to fix between the structure perturbation and kernel	-
7	Pressure perturbation $\delta_x p / p$ neglected compared to $\delta_x \Gamma_1 / \Gamma_1$	-
8	Reference model non-established	-
9	Persistence of the linear perturbation regime	-

(*) : The “-” indicates that the method does not require the assumption / is not concerned by the limit under consideration.

The question is then to know to what extent can the properties be constrained when exploiting real data. With regards to the results obtained in the ideal framework, the proposed glitch modelling seems to be a satisfying approximation to the first order. The major part of the structure information could be retrieved, and the more pronounced uncertainties on helium abundance are the direct consequence of a lack of low-frequency information, upon which the choice of model has no influence. From a theoretical point of view, the presented glitch modelling addresses the objectives raised in section 3.7, improving previous approaches regarding the points underlined in TABLE 7.1 (see also TABLE 3.2). It would of course be possible to propose a more exact modelling of the ionisation region (by considering the Coulombian pressure during ionisation for instance), but the gains in accuracy are likely to be marginal while the imprecision due to additional parameters (in this case, a scale for the Debye length, BRAUS, 2019) will necessarily increase.

A more problematic aspect resides, I believe, in the interaction of the ionisation glitch with other components that are present in the frequency differences. Taking again the example of the surface effects, the test we carried out remains ideal in itself, even if it is more realistic than in the absence of surface effects. Indeed, we added in this framework some synthetic surface effects to the ionisation glitch, but we did not assume any coupling between the two components. However, the study of realistic models that are “patched” by the addition of an envelop coming from 3D simulations highlights a strong interaction between the phenomena at the origin of

surface effects and the ionisation region of hydrogen (SONOI et al., 2015), which is to be expected due to their proximity in the stellar structure. Hence, it should be expected that the hydrogen component cannot be distinguished from other surface effects, or even that our model estimates it incorrectly if the coupling effect is dominant.

This analysis underlines the benefit of taking these effects into account simultaneously when modelling. As was done in this study, where the three ionisation zones (H), (HeI) and (HeII) appear together in the same model which contains only three parameters (nine would have been required for an approximation using three Gaussians), it seems to me that including this additional coupling would be highly beneficial. Let us note that the simultaneous consideration of several components in the frequency differences could also lead to a better determination of the ionisation region properties. By taking this time the example of the glitch associated with the base of the convection zone (BCZ), it seems likely that the depth of this transition τ_{BCZ} is not independent from the extent of the ionisation region ε for instance. The development of a model gathering both the ionisation and BCZ glitches could then enable an even more precise determination of the ε parameter, which would then be constrained by two physical processes. More generally, the simultaneous consideration of several physical processes in models gives the prospect of developing complex but weakly parameterised models (as they overlap) which are therefore extremely predictive. Such an exploitation draws closer to global inversion (cf. paragraph 2.5.5), but uses a physically established basis involving fewer functions, as they correspond to the number of parameters of the model¹.

The development of these models however requires a physical modelling of the mechanisms in action, and therefore advanced theoretical work on complex subjects. The consideration of surface effects could be achieved by modelling the turbulent pressure in order to study its interaction with the helium ionisation region, as these two phenomena both influence the effective stratification of the region, and as such the γ exponent. Regarding the BCZ glitch, a modelling inspired from the work of REMPEL, 2004 seems promising to me, as they already present a parameterised modelling of the structure that includes γ .

1. Let us note that this is how we introduced the seismic glitch in this manuscript, as a parameterised inversion (cf. paragraph 2.5.6).

Troisième partie

Annexes

Annexe A

Compléments thermodynamiques

Contenu de l'annexe

A.1	Définition de l'entropie	165
A.2	Expression des potentiels chimiques	167
A.3	Relations thermodynamiques locales	168
A.4	Variables d'états	169
A.5	Quantités thermodynamiques pour le fluide partiellement ionisé	170
A.6	Perturbation sous contrainte	175

A.1 Définition de l'entropie

Dans ce manuscrit, l'entropie a été exprimée à partir de la fonction de distribution f . Le lien entre ces deux grandeurs n'est pas aussi évident qu'il n'y paraît, il est facile de créer des incohérences dans le formalisme qui peuvent être exacerbées par le fait que l'entropie est souvent définie "à une constante près". On va donc commencer par une dérivation schématique de cette grandeur avant d'aboutir à une expression formelle. L'entropie, au sens informationnel, est définie comme le logarithme du nombre de configurations possibles pour un système donné (exprimée en JK^{-1}) :

$$S \equiv k \ln \Omega \quad (\text{A.1.1})$$

où Ω dénombre un nombre de configurations total. Ce nombre peut être partitionné sur les différents états *distinguishables* indexés par j : $\Omega = \prod_j \Omega_j$ avec Ω_j le nombre de configurations de l'état j . Généralement, tous les états du système ne peuvent être distingués, on notera g_j le nombre d'états *indistinguishables* au sein de l'état macroscopique j . Si N_j particules identiques peuplent l'état j , le nombre de configurations s'élève donc à :

$$\Omega_j = \frac{g_j^{N_j}}{N_j!} \quad (\text{A.1.2})$$

ce qui correspond aux différentes possibilités d'agencer N_j particules sur g_j états. L'entropie totale du système s'écrit alors :

$$S = k \sum_j (N_j \ln g_j - \ln N_j!). \quad (\text{A.1.3})$$

Puisque N_j sera toujours très grand devant 1, l'approximation de Stirling peut être utilisée en ne conservant que les deux premiers termes : $\ln N_j! \sim N_j \ln N_j - N_j$. On pourra donc rassembler les termes comme suit :

$$S = k \sum_j N_j \ln \frac{e g_j}{N_j}. \quad (\text{A.1.4})$$

Cherchons maintenant à adapter cette expression au cas étudié dans la partie principale : celui d'une distribution sur les états semi-classique. On dispose d'une fonction de distribution $f(\mathbf{v}, \mathbf{r}, q, t)$ qui s'exprime à l'équilibre de la façon suivante :

$$f(\mathbf{v}, \mathbf{r}, q, t) = n \left(\frac{m}{2\pi kT} \right)^{3/2} \exp \left(-\frac{m(\mathbf{v} - \bar{\mathbf{v}})^2}{2kT} \right) \left(\frac{g_q e^{-\varepsilon_q/kT}}{\mathcal{Z}_{\text{int}}} \right) \quad (\text{A.1.5})$$

Notons l'énergie cinétique d'une particule par rapport à l'ensemble $K(\mathbf{v}, \mathbf{r}, t) \equiv (1/2)m(\mathbf{v} - \bar{\mathbf{v}})^2$. Puisque

$$\int e^{-K(\mathbf{v})/kT} d\mathbf{v} = \left(\frac{2\pi kT}{m} \right)^{3/2}, \quad (\text{A.1.6})$$

la fonction de distribution se réécrit :

$$f(\mathbf{v}, \mathbf{r}, q, t) = N \left(\frac{n}{\mathcal{Z}_{\text{pos}}} \right) \left(\frac{e^{-K(\mathbf{v})/kT}}{\mathcal{Z}_{\text{cin}}} \right) \left(\frac{g_q e^{-\varepsilon_q/kT}}{\mathcal{Z}_{\text{int}}} \right) \quad (\text{A.1.7})$$

en définissant les fonctions de partition suivantes :

$$\mathcal{Z}_{\text{pos}} \equiv \int n(\mathbf{r}) d\mathbf{r} = N \quad (\text{A.1.8})$$

$$\mathcal{Z}_{\text{cin}} \equiv \int e^{-K(\mathbf{v})/kT} d\mathbf{v} = \left(\frac{2\pi kT}{m} \right)^{3/2} \quad (\text{A.1.9})$$

$$\mathcal{Z}_{\text{int}} \equiv \sum_q g_q e^{-\varepsilon_q/kT}. \quad (\text{A.1.10})$$

Dans l'expression (A.1.7), chaque parenthèse définit une densité de probabilité. Elles portent respectivement sur la position, la vitesse et l'état interne de la particule. Cette écriture permet de voir clairement que f/N définit une densité de probabilité sur les variables $(\mathbf{v}, \mathbf{r}, q)$ qui caractérisent l'état d'une particule. Ainsi, $f d\mathbf{v} d\mathbf{r}$ donne le nombre moyen de particules dans l'état $\{\mathbf{r}' \in \mathcal{B}(\mathbf{r}, |d\mathbf{r}|), \mathbf{v}' \in \mathcal{B}(\mathbf{v}, |d\mathbf{v}|), q' = q\}$ (avec $\mathcal{B}(\mathbf{x}, r)$ la boule centrée en \mathbf{x} de rayon r). Au sein de ces états distinguables, il nous faut trouver le nombre d'états indistinguables les composant avant d'exprimer l'entropie. Il s'agit des états quantiques qui occupent en moyenne un volume h^3 de l'espace des phases ; ils sont donc au nombre de $d\mathbf{v} d\mathbf{r} / h^3$ par état. On généralisera donc l'expression (A.1.4) à l'aide des substitutions

$$N_j \rightarrow f d\mathbf{v} d\mathbf{r} \quad \text{et} \quad \frac{g_j}{N_j} \rightarrow \frac{h^{-3} d\mathbf{v} d\mathbf{r}}{f d\mathbf{v} d\mathbf{r}} = \frac{1}{h^3 f}. \quad (\text{A.1.11})$$

L'entropie du système composé des N particules s'écrit alors :

$$S(t) = -k \int \sum_q f(\mathbf{v}, \mathbf{r}, q, t) \ln \left(\frac{h^3 f(\mathbf{v}, \mathbf{r}, q, t)}{e} \right) d\mathbf{v} d\mathbf{r}. \quad (\text{A.1.12})$$

ce qui donne le lien avec la fonction de distribution du système.

Comme affirmé dans la partie principale, cette fonction s'écrit bien comme $S = -kH + \text{cste}$ avec H la fonction H de Boltzmann. En réalité, tout l'intérêt du raisonnement présent est de permettre l'évaluation de cette constante additive et l'on va maintenant montrer qu'elle permet bien la cohérence thermodynamique du formalisme. À l'équilibre thermodynamique global, les grandeurs $n, \bar{\mathbf{v}}$ et T sont des constantes sur l'espace et l'on peut donner une expression explicite

de l'entropie :

$$\begin{aligned}
S &= -k \left[N \ln \left(\frac{Nh^3}{eV \mathcal{Z}_{\text{cin}} \mathcal{Z}_{\text{int}}} \right) - \frac{1}{kT} \int \sum_q f(K + \varepsilon_q) d\mathbf{v} d\mathbf{r} + \int \sum_q f \ln g_q d\mathbf{v} d\mathbf{r} \right] \\
&= -k \left[N \ln \left(\frac{Nh^3}{eV \mathcal{Z}_{\text{cin}} \mathcal{Z}_{\text{int}}} \right) - \frac{1}{kT} \left(\frac{3}{2} NkT + N \langle \varepsilon_q \rangle \right) + N \langle \ln g_q \rangle \right] \\
&= -kN \left[\ln \left(\frac{Nh^3}{V \mathcal{Z}_{\text{cin}} \mathcal{Z}_{\text{int}}} \right) - \frac{5}{2} - \frac{\langle \varepsilon_q \rangle}{kT} + \langle \ln g_q \rangle \right]
\end{aligned} \tag{A.1.13}$$

On remarquera que $h^3/\mathcal{Z}_{\text{cin}}$ est homogène à une distance au cube et définit ainsi la longueur d'onde thermique de De Broglie :

$$\lambda^3 \equiv \left(\frac{mh^2}{2\pi kT} \right)^{3/2}. \tag{A.1.14}$$

Dans un système où la structure fine est non dégénérée ($\langle \ln g_q \rangle = 0$) et ne peut être distinguée du niveau d'énergie fondamental ($\langle \varepsilon_q \rangle \ll kT$), l'expression (A.1.13) permet de retrouver la formule de Sackur-Tetrode (SACKUR, 1911 ; TETRODE, 1912) :

$$S = -kN \left[\ln \left(\frac{N\lambda^3}{V \mathcal{Z}_{\text{int}}} \right) - \frac{5}{2} \right]. \tag{A.1.15}$$

Notons que l'on peut adapter sans difficulté l'expression de l'entropie (A.1.12) à une mixture de particules dont l'on distinguera les espèces par l'index α . Par définition, chaque espèce peut être distinguée des autres et le nombre de configurations total s'obtient donc comme le produit des configurations individuelles. L'entropie totale de la mixture s'obtient donc par une somme :

$$S(t) = -k \sum_{\alpha} \int \sum_q f_{\alpha} \ln \left(\frac{h^3 f_{\alpha}}{e} \right) d\mathbf{v}_{\alpha} d\mathbf{r}. \tag{A.1.16}$$

De façon intéressante, la formulation (A.1.16) permet la généralisation de la notion d'entropie aux systèmes hors équilibre en définissant une entropie (volumique) locale s :

$$\rho s(\mathbf{r}, t) \equiv k \sum_{\alpha} \int \sum_q f_{\alpha} \ln \left(\frac{h^3 f_{\alpha}}{e} \right) d\mathbf{v}_{\alpha} \tag{A.1.17}$$

où l'on a simplement retiré l'intégration sur l'espace physique. Un calcul analogue à (A.1.13) permet d'exprimer ρs dans le cas où f est localement maxwellienne :

$$\rho s = -k \sum_{\alpha} n_{\alpha} \left[\ln \left(\frac{n_{\alpha} \lambda_{\alpha}^3}{\mathcal{Z}_{\alpha}} \right) - \frac{5}{2} \right] \tag{A.1.18}$$

avec \mathcal{Z}_{α} la fonction de partition interne à chaque espèce et où l'on a encore une fois négligé les niveaux d'énergie de la structure fine. Plutôt qu'au nombre de configurations total du système, cette entropie volumique est à associer à la densité d'état menant au même état macroscopique en \mathbf{r} .

A.2 Expression des potentiels chimiques

Puisque l'énergie interne de la mixture est donnée par $\rho \varepsilon = (3/2)kT \sum_{\alpha} n_{\alpha}$, l'équation (A.1.18) permet d'exprimer l'énergie libre f localement :

$$\rho f \equiv \rho(\varepsilon - Ts) = \sum_{\alpha} n_{\alpha} kT \left[\ln \left(\frac{n_{\alpha} \lambda_{\alpha}^3}{\mathcal{Z}_{\alpha}} \right) - 1 \right]. \tag{A.2.1}$$

Notons que, tout comme $\rho\varepsilon$ et ρs , l'énergie libre par unité de volume ρf s'exprime uniquement sur la température T et les densités n_α . En particulier, ces fonctions ne dépendent pas de la masse volumique ρ . Les potentiels chimiques s'obtiennent à partir de la dérivée partielle :

$$\mu_\alpha \equiv \left(\frac{\partial \rho f}{\partial n_\alpha} \right)_{T, n_\beta} = kT \ln \left(\frac{n_\alpha \lambda_\alpha^3}{\mathcal{Z}_\alpha} \right). \quad (\text{A.2.2})$$

Cette grandeur permet d'écrire de façon compacte les entropie et énergie libre locales¹ :

$$\rho s = \frac{1}{T} \left(\frac{5}{2} p - \sum_\alpha n_\alpha \mu_\alpha \right) \quad (\text{A.2.3})$$

$$\rho f = \sum_\alpha n_\alpha \mu_\alpha - p \quad (\text{A.2.4})$$

avec la pression du fluide $p = kT \sum_\alpha n_\alpha$.

A.3 Relations thermodynamiques locales

L'objectif est maintenant d'obtenir une forme locale des identités thermodynamiques comme :

$$dE = TdS - pdV + \sum_\alpha \mu_\alpha dN_\alpha \quad (\text{A.3.1})$$

pour vérifier la cohérence du formalisme local. Cette identité s'écrit localement :

$$d(\rho\varepsilon) = Td(\rho s) - \underbrace{pd(1)}_0 + \sum_\alpha \mu_\alpha dn_\alpha, \quad (\text{A.3.2})$$

ce qui explique pourquoi les grandeurs locales semblent dépendre d'une variable de moins. La différentielle sur ρf s'obtient par transformation de Legendre en soustrayant $d(\rho Ts)$:

$$d(\rho f) = \rho s dT + \sum_\alpha \mu_\alpha dn_\alpha, \quad (\text{A.3.3})$$

et justifie la définition du potentiel chimique (A.2.2) donnée plus haut. À noter qu'il est directement possible d'exprimer ces différentielles sur les $\rho_\alpha \equiv m_\alpha n_\alpha$ plutôt que sur les n_α puisque les m_α sont constants :

$$d(\rho\varepsilon) = Td(\rho s) + \sum_\alpha \tilde{\mu}_\alpha d\rho_\alpha, \quad (\text{A.3.4})$$

$$d(\rho f) = -\rho s dT + \sum_\alpha \tilde{\mu}_\alpha d\rho_\alpha \quad (\text{A.3.5})$$

avec $\tilde{\mu}_\alpha \equiv \mu_\alpha / m_\alpha$ un potentiel chimique massique. La densité ρ réapparaît comme variable d'état sitôt que l'on cherche à exprimer les différentielles sur les grandeurs massiques :

$$\rho d\varepsilon = -\underbrace{(\varepsilon - Ts)}_f d\rho + \rho T ds + \sum_\alpha \mu_\alpha dn_\alpha. \quad (\text{A.3.6})$$

1. La quantité $\sum_\alpha N_\alpha \mu_\alpha$ est parfois définie comme le *potentiel thermodynamique*, Φ (par exemple, LANDAU et LIFSHITZ, 1969). Il est donc possible de définir, par analogie avec les autres grandeurs locales, un potentiel local $\varphi \equiv (1/m) \sum_\alpha x_\alpha \mu_\alpha$ tel que $\rho\varphi = \sum_\alpha n_\alpha \mu_\alpha$.

En utilisant la relation (A.2.4), la différentielle sur l'énergie s'écrit :

$$\begin{aligned} \rho d\varepsilon &= p \frac{d\rho}{\rho} + \rho T ds + \sum_{\alpha} \mu_{\alpha} n_{\alpha} \left(\frac{dn_{\alpha}}{n_{\alpha}} - \frac{d\rho}{\rho} \right) \\ &= \rho T ds - \rho p d \left(\frac{1}{\rho} \right) + \rho \sum_{\alpha} \mu_{\alpha} d \left(\frac{n_{\alpha}}{\rho} \right). \end{aligned} \quad (\text{A.3.7})$$

Enfin, puisque $n_{\alpha}/\rho = X_{\alpha}/m_{\alpha}$ (avec $X_{\alpha} \equiv \rho_{\alpha}/\rho$), on obtient une expression des différentielles sur les fractions massiques :

$$d\varepsilon = T ds - p d \left(\frac{1}{\rho} \right) + \sum_{\alpha} \tilde{\mu}_{\alpha} dX_{\alpha}, \quad (\text{A.3.8})$$

$$df = -s dT - p d \left(\frac{1}{\rho} \right) + \sum_{\alpha} \tilde{\mu}_{\alpha} dX_{\alpha}, \quad (\text{A.3.9})$$

la transformée de Legendre sur s suivant immédiatement. Le jeu de différentielles le plus complexe à trouver s'exprime sur les fractions volumiques $x_{\alpha} \equiv n_{\alpha}/n$. En repartant de l'équation (A.3.2) :

$$d(\rho\varepsilon) = T d(\rho s) + \left(\sum_{\alpha} \mu_{\alpha} n_{\alpha} \right) \frac{dn}{n} + \sum_{\alpha} \mu_{\alpha} n d x_{\alpha} \quad (\text{A.3.10})$$

en utilisant la propriété $dn_{\alpha} = n dx_{\alpha} + (n_{\alpha}/n) dn$. L'expression de l'énergie libre (A.2.4) nous permet d'écrire $\sum_{\alpha} \mu_{\alpha} n_{\alpha} = \rho f + p = \rho\varepsilon - \rho Ts + p$, ce qui donne :

$$\underbrace{d(\rho\varepsilon) - \rho\varepsilon \frac{dn}{n}}_{nd(\rho\varepsilon/n)=nd(m\varepsilon)} = \underbrace{T d(\rho s) - \rho Ts \frac{dn}{n}}_{nTd(\rho s/n)=nTd(ms)} + p \frac{dn}{n} + \sum_{\alpha} \mu_{\alpha} n d x_{\alpha} \quad (\text{A.3.11})$$

avec la masse moyenne $m \equiv \rho/n$. Ce dernier couple de différentielle s'écrit donc

$$d(m\varepsilon) = T d(ms) - p d \left(\frac{1}{n} \right) + \sum_{\alpha} \mu_{\alpha} d x_{\alpha}, \quad (\text{A.3.12})$$

$$d(mf) = -m s dT - p d \left(\frac{1}{n} \right) + \sum_{\alpha} \mu_{\alpha} d x_{\alpha}. \quad (\text{A.3.13})$$

Si la forme des différentielles volumiques (A.3.2)-(A.3.3) & (A.3.4)-(A.3.5), massiques (A.3.8)-(A.3.9) et atomiques (A.3.12)-(A.3.13) peut être intuitée à partir de la forme globale (A.3.1), leur définition exacte n'a en réalité rien de trivial. Celle-ci dépend notamment de la définition de l'entropie faite dans la section A.1 et de la constante apparaissant dans la relation $S = -kH + \text{cste}$. En particulier, la définition $S = -kH$ que l'on trouve parfois mène soit à une énergie libre indépendante de la pression soit à une expression incorrecte des potentiels chimiques (et donc des équations de Saha invalides). On remarque que l'équation (A.3.8) a été utilisée sous l'équation (2.2.60) de la partie principale pour définir l'entropie d'un fluide composé d'une seule espèce.

A.4 Variables d'états

La différentielle classique sur l'énergie (A.3.1) fait apparaître $N_{\text{sp}} + 2$ variables d'états nécessaires à la description du système. Ce nombre passe à $N_{\text{sp}} + 1$ pour les expressions locales (A.3.2)-(A.3.3) & (A.3.4)-(A.3.5) puisque les grandeurs s'expriment par unité de volume (elles ne dépendent donc plus de V). Il peut cependant sembler curieux qu'une variable d'état supplémentaire ρ apparaisse en plus dans les différentielles massiques (A.3.8)-(A.3.9) et atomiques (A.3.12)-(A.3.13), alors que le système physique lui-même ne dépend pas de la représentation

choisie. En réalité le nombre de variable d'état n'a pas changé puisque l'on dispose par définition de la contrainte implicite $\sum_{\alpha} X_{\alpha} = 1$ (ou $\sum_{\alpha} x_{\alpha} = 1$). Formellement, les potentiels ε et ρ peuvent donc être vues comme des fonctions de ρ, T et des $X_{\alpha \neq \beta}$ avec β un choix d'espèce arbitraire dont la fraction peut-être retrouvée par $X_{\beta} = 1 - \sum_{\alpha \neq \beta} X_{\alpha}$. Notons que le passage des fractions volumiques x_{α} aux fractions massiques X_{α} (très utilisées en physique stellaire) ne pose aucun problème puisque :

$$\forall \alpha \neq \beta, X_{\alpha} = \frac{m_{\alpha} n_{\alpha}}{\rho} = \frac{m_{\alpha} x_{\alpha}}{m} = \frac{m_{\alpha} x_{\alpha}}{m_{\beta} + \sum_{\gamma \neq \beta} (m_{\gamma} - m_{\beta}) x_{\gamma}} \quad (\text{A.4.1})$$

ne s'exprime qu'en fonction des $x_{\alpha \neq \beta}$. De manière générale, on remarque la correspondance entre les représentations $(n, x_{\alpha \neq \beta}) \leftrightarrow (n_{\alpha}) \leftrightarrow (\rho_{\alpha}) \leftrightarrow (\rho, X_{\alpha \neq \beta})$, schématisée par la relation :

$$m_{\alpha} n x_{\alpha} = m_{\alpha} n_{\alpha} = \rho_{\alpha} = \rho X_{\alpha}. \quad (\text{A.4.2})$$

Celle-ci met en évidence que le nombre de dépendances du système s'élève à $N_{\text{sp}} + 1$ variables d'état quelque soit la représentation choisie.

A.5 Quantités thermodynamiques pour le fluide partiellement ionisé

A.5.1 Différentielles atomiques à l'équilibre chimique

Pour le cas spécifique de l'ionisation, les espèces α sont les différents éléments i dans l'état d'excitation r en plus des électrons, dénotés par l'indice e . Il est important de noter que les fractions d'ionisation sont définies comme $x_i^r \equiv n_i^r / n_i$ et non pas par n_i^r / n . Cela induit le changement suivant dans les différentielles massiques :

$$\sum_{\alpha} \mu_{\alpha} d \left(\frac{n_{\alpha}}{n} \right) = \sum_{i=1}^{N_{\text{sp}}} \sum_{r=0}^i \mu_i^r d \left(\frac{n_i^r n_i}{n_i n} \right) + \mu_e dx_e = \sum_{i=1}^{N_{\text{sp}}} \sum_{r=0}^i x_i \mu_i^r dx_i^r + \sum_{i=1}^{N_{\text{sp}}} \sum_{r=0}^i x_i^r \mu_i^r dx_i + \mu_e dx_e \quad (\text{A.5.1})$$

avec $\langle \mu_i^r \rangle \equiv \sum_r x_i^r \mu_i^r$. Cette expression nous permet de définir les potentiels chimiques en présence d'ionisation avec les dérivées partielles (peu intuitives) de l'équation (A.3.12) :

$$\left(\frac{\partial m \varepsilon}{\partial x_i^r} \right)_{s, \rho, x_i, x_i^s, x_e} = \mu_i^r \quad \text{et} \quad \left(\frac{\partial m \varepsilon}{\partial x_e} \right)_{s, \rho, x_i, x_i^r} = \mu_e. \quad (\text{A.5.2})$$

Il s'agit des définitions utilisées dans le paragraphe 2.3.7. On peut donner une expression de la somme (A.5.1) en exploitant les relations dérivées dans le paragraphe 2.3.7 et qui traduisent respectivement la réaction d'ionisation elle-même, l'homogénéité du mélange (région convective) et l'électroneutralité :

$$\forall 1 \leq i \leq N_{\text{sp}}, 1 \leq r \leq i, \quad \mu_i^r + \mu_e - \mu_i^{r-1} = 0, \quad (\text{A.5.3})$$

$$\forall 1 \leq i \leq N_{\text{sp}}, \quad x_i = \bar{x}_i, \quad (\text{A.5.4})$$

$$x_e = \sum_{ir} r x_i x_i^r \quad (\text{A.5.5})$$

avec $\bar{x}_i = \text{cste}$, la fraction (en nombre) moyenne en élément i dans la région convective et la convention d'écriture² :

$$\sum_{ir} \equiv \sum_{i=1}^{N_{\text{sp}}} \sum_{r=1}^i. \quad (\text{A.5.6})$$

2. De manière générale, on omettra d'expliciter les bornes d'une somme lorsque qu'elle commence par le terme d'indice 1 et finit par l'indice le plus grand possible (N_{sp} pour les sommes sur i , i pour les sommes sur r et r pour les sommes sur s).

Cela donne immédiatement $dx_i = 0$ et $dx_e = \sum_{ir} rx_i dx_i^r$, ce qui permet de rassembler les termes de la façon suivante :

$$\sum_{\alpha} \mu_{\alpha} dx_{\alpha} = \sum_{ir} x_i (\mu_i^r + r\mu_e) dx_i^r + \sum_i x_i \mu_i dx_i^0 \quad (\text{A.5.7})$$

avec $\mu_i \equiv \mu_i^0$, le potentiel chimique de l'élément neutre. Remarquons maintenant la chose suivante : le terme $\mu_i^r + r\mu_e$ fait intervenir l'équation (A.5.3) de la façon suivante :

$$\mu_i^r + r\mu_e = \mu_i^0 + \sum_{s \leq r} (\mu_i^s + \mu_e - \mu_i^{s-1}). \quad (\text{A.5.8})$$

En utilisant le théorème d'interversion des sommes

$$\sum_r a_r \sum_{s \leq r} b_s = \sum_r b_r \sum_{s \geq r} a_s, \quad (\text{A.5.9})$$

on pourra écrire l'équation (A.5.7) comme :

$$\sum_{\alpha} \mu_{\alpha} dx_{\alpha} = \sum_{ir} x_i (\mu_i^r + \mu_e - \mu_i^{r-1}) dy_i^r + \sum_i x_i \mu_i \sum_{r=0}^i dx_i^r \quad (\text{A.5.10})$$

en faisant intervenir la *fraction d'ionisation cumulée* :

$$y_i^r \equiv \sum_{s \geq r} x_i^s. \quad (\text{A.5.11})$$

L'équation (A.5.3) et la propriété $\forall i, \sum_{r=0}^i x_i^r = 1$ nous donne alors :

$$\sum_{\alpha} \mu_{\alpha} dx_{\alpha} = 0. \quad (\text{A.5.12})$$

Ainsi, on peut écrire les différentielles atomiques à l'intérieur de la région d'ionisation de la façon suivante :

$$d(m\varepsilon) = Td(ms) - pd \left(\frac{1}{n} \right), \quad (\text{A.5.13})$$

$$d(mf) = -msdT - pd \left(\frac{1}{n} \right). \quad (\text{A.5.14})$$

A.5.2 Potentiel thermodynamique, énergie et énergie libre

En fait, il est possible de démontrer une propriété encore plus remarquable en nous intéressant directement au potentiel thermodynamique $m\varphi \equiv \sum_{\alpha} x_{\alpha} \mu_{\alpha}$. En retirant, les "d" des différentielles de l'équation (A.5.10), il vient immédiatement :

$$m\varphi = \sum_i x_i \mu_i. \quad (\text{A.5.15})$$

Cette propriété n'a rien d'évident. Elle nous affirme que le potentiel thermodynamique reste toujours égal au potentiel du gaz neutre, $\sum_i x_i \mu_i$, ou, autrement dit, le potentiel ne varie pas dans la région d'ionisation. Un corollaire direct en découle : l'énergie libre, f , ne varie qu'indirectement dans la zone d'ionisation et ce au travers des variations de pression³. En effet, la pression $p =$

3. Le même phénomène se produit si on l'on cherche à exprimer la part de l'énergie libre causée par la radiation plutôt que l'ionisation (GRABOSKE, HARWOOD et ROGERS, 1969). Celle-ci n'intervient que dans la pression par le terme $p^R = 4\sigma T^4/3c$, dérivé dans le paragraphe 2.2.9. Cette contribution n'apparaît pas explicitement dans le calcul, celle-ci étant généralement négligeable dans les régions d'ionisation en plus de ne pas influencer sur les fractions d'ionisation (MIHALAS, DAPPEN et HUMMER, 1988).

$kT \sum_{\alpha} n_{\alpha}$ se réécrit dans le cas d'une ionisation partielle comme :

$$p = nkT(1 + x_e) = nkT\left(1 + \sum_{ir} rx_i x_i^r\right) = nkT\left(1 + \sum_{ir} x_i y_i^r\right) \quad (\text{A.5.16})$$

et l'équation (A.2.4) donne :

$$\rho f = \rho \varphi - p = n \left[\sum_i x_i \mu_i - kT \left(1 + \sum_{ir} x_i y_i^r \right) \right]. \quad (\text{A.5.17})$$

Avant de donner une expression de l'énergie dans cette région, il nous faut exprimer l'entropie ρs . En reprenant le développement (A.1.13), il n'est cette fois plus possible de négliger les différences entre les énergies des niveaux fondamentaux pour les différentes espèces. En négligeant une nouvelle fois la structure fine, on pourra écrire

$$\langle \varepsilon_{iq}^r \rangle \simeq \sum_{r \leq s} \chi_i^r \quad (\text{A.5.18})$$

et donc par interversion des sommes :

$$\sum_{ir} n_i^r \langle \varepsilon_{iq}^r \rangle = n \sum_{ir} x_i x_i^r \sum_{r \leq s} \chi_i^r = n \sum_{ir} x_i y_i^r \chi_i^r. \quad (\text{A.5.19})$$

L'entropie volumique prend alors la forme :

$$\rho s = \frac{n}{T} \left[\frac{5}{2} kT \left(1 + \sum_{ir} x_i y_i^r \right) + \sum_{ir} x_i y_i^r \chi_i^r - \sum_i x_i \mu_i \right]. \quad (\text{A.5.20})$$

Enfin, puisque l'énergie s'exprime comme $\varepsilon = f + Ts$:

$$\rho \varepsilon = n \left[\frac{3}{2} kT \left(1 + \sum_{ir} x_i y_i^r \right) + \sum_{ir} x_i y_i^r \chi_i^r \right]. \quad (\text{A.5.21})$$

On peut définir une énergie interne d'ionisation de la façon suivante :

$$\rho \varepsilon_{\text{ion}} \equiv n \sum_{ir} x_i y_i^r \chi_i^r \quad (\text{A.5.22})$$

pour retrouver l'écriture de l'énergie interne donnée dans le paragraphe 3.2 :

$$\rho \varepsilon = \frac{3}{2} p + \rho \varepsilon_{\text{ion}}. \quad (\text{A.5.23})$$

On effectuera cependant la remarque suivante. On peut définir la quantité adimensionnée

$$\phi_i^r \equiv \frac{3}{2} + \frac{\chi_i^r}{kT}, \quad (\text{A.5.24})$$

pour arriver au second découpage suivant :

$$\rho \varepsilon = nkT \left[\frac{3}{2} + \sum_{ir} x_i y_i^r \phi_i^r \right]. \quad (\text{A.5.25})$$

Cette écriture permet de rendre explicite l'énergie interne réellement apportée par l'ionisation. Celle-ci possède de fait deux origines distinctes : premièrement, l'élévation du nombre de particule (i.e. les électrons) qui contribue à augmenter la pression totale (impact indirect, responsable du 3/2 dans l'équation (A.5.24)) et deuxièmement, l'excitation des particules déjà

comptabilisées par le phénomène d'ionisation lui-même (impact direct, responsable du χ_i^r/kT).

A.5.3 Premier exposant adiabatique

En introduisant par praticité la fonction :

$$\mathbf{EQ}(\rho, T, x_i, x_i^r, x_e) \equiv \begin{pmatrix} \mathbf{EQ}_e \\ \vdots \\ \mathbf{EQ}_i \\ \vdots \\ \mathbf{EQ}_i^r \\ \vdots \end{pmatrix} \begin{matrix} (1 \leq i \leq N_{\text{sp}}) \\ \\ (1 \leq r \leq i \leq N_{\text{sp}}) \end{matrix} (\rho, T, x_i, x_i^r, x_e) = \begin{pmatrix} x_e - \sum_{ir} r x_i x_i^r \\ \vdots \\ x_i - \bar{x}_i \\ \vdots \\ \mu_i^r + \mu_e - \mu_i^{r-1} \\ \vdots \end{pmatrix}, \quad (\text{A.5.26})$$

on voit que les conditions (A.5.3)-(A.5.5) régissant l'équilibre chimique de l'ionisation dans la région convective se résument par la contrainte :

$$\mathbf{EQ} = \mathbf{0}. \quad (\text{A.5.27})$$

Afin d'assurer la correspondance avec l'exposant effectif γ dans une telle région (tant que l'hypothèse d'adiabaticité $s = \text{cste}$ est vérifiée), on peut alors définir l'exposant adiabatique dans le fluide partiellement ionisé par :

$$\Gamma_1 \equiv \left(\frac{\partial \ln p}{\partial \ln \rho} \right)_{s, \mathbf{EQ}}. \quad (\text{A.5.28})$$

Les dérivées prises à entropie fixée n'étant que difficilement exploitables, on pourra chercher à réexprimer cet exposant adiabatique sur les variables de température T et de volume moyen par unité de masse $v \equiv 1/\rho$, très pratique pour ce calcul. On utilisera pour cela le déterminant des matrices jacobiniennes :

$$\begin{aligned} \Gamma_1 &= -\frac{v}{p} \left(\frac{\partial p}{\partial v} \right)_{s, \mathbf{EQ}} = -\frac{v}{p} \left| \frac{\partial(p, s, \mathbf{EQ})}{\partial(v, s, \mathbf{EQ})} \right| \\ &= -\frac{v}{p} \left| \frac{\partial(p, s, \mathbf{EQ})}{\partial(v, T, \mathbf{EQ})} \cdot \frac{\partial(v, T, \mathbf{EQ})}{\partial(v, s, \mathbf{EQ})} \right| \\ &= -\frac{v}{p} \left(\frac{\partial s}{\partial T} \right)_{v, \mathbf{EQ}}^{-1} \left[\left(\frac{\partial p}{\partial v} \right)_{T, \mathbf{EQ}} \left(\frac{\partial s}{\partial T} \right)_{v, \mathbf{EQ}} - \left(\frac{\partial p}{\partial T} \right)_{v, \mathbf{EQ}} \left(\frac{\partial s}{\partial v} \right)_{T, \mathbf{EQ}} \right]. \end{aligned} \quad (\text{A.5.29})$$

En utilisant maintenant la différentielle de l'énergie libre (A.3.9), il est possible de réexprimer l'entropie et la pression en fonction de l'énergie libre :

$$s \equiv - \left(\frac{\partial f}{\partial T} \right)_{v, X_\alpha} \quad \& \quad p \equiv - \left(\frac{\partial f}{\partial v} \right)_{T, X_\alpha}. \quad (\text{A.5.30})$$

L'intérêt de cette manoeuvre vient du fait que l'énergie libre est le potentiel s'exprimant naturellement sur les variables T et v . On définira à présent les opérateurs :

$$\partial_\alpha \equiv (-1)^{\delta_\alpha^v} \frac{m\alpha}{kT} \left(\frac{\partial}{\partial \alpha} \right)_{\beta, \mathbf{EQ}} \quad (\text{A.5.31})$$

$$\partial_{\alpha\beta}^2 \equiv (-1)^{\delta_\alpha^v \delta_\beta^v + 1} \frac{m\alpha\beta}{kT} \left[\frac{\partial}{\partial \alpha} \left(\frac{\partial}{\partial \beta} \right)_{\alpha, X_\alpha} \right]_{\beta, \mathbf{EQ}} \quad (\text{A.5.32})$$

avec $\delta_v^\alpha = 1$ si $\alpha = v$ et 0 sinon. Ces définitions peuvent sembler complexes mais ont pour but de ne manipuler que des quantités adimensionnées et positives. À l'aide de ces introductions et des définitions (A.5.30), l'équation peut se réécrire uniquement en termes des dérivées première et seconde de l'énergie libre f :

$$\Gamma_1 = \frac{(\partial_{TT}^2 f)(\partial_{vv}^2 f) + (\partial_{Tv}^2 f)(\partial_{vT}^2 f)}{(\partial_v f)(\partial_{TT}^2 f)}. \quad (\text{A.5.33})$$

On a par définition de la pression p :

$$\partial_v f = x_e = 1 + \sum_{ir} x_i y_i^r, \quad (\text{A.5.34})$$

ce qui nous permet de déduire les dérivées ∂_{vv}^2 et ∂_{Tv}^2 :

$$\partial_{vv}^2 f = 1 + \sum_{ir} x_i y_i^r \left[1 - \left(\frac{\partial \ln y_i^r}{\partial \ln v} \right)_{T, \mathbf{EQ}} \right], \quad (\text{A.5.35})$$

$$\partial_{Tv}^2 f = 1 + \sum_{ir} x_i y_i^r \left[1 + \left(\frac{\partial \ln y_i^r}{\partial \ln T} \right)_{v, \mathbf{EQ}} \right]. \quad (\text{A.5.36})$$

Il est à présent plus simple de passer par l'énergie pour exprimer les dernières dérivées. Puisque $m = \text{cste}$ ($x_i = \bar{x}_i$), la différentielle à $\mathbf{EQ} = \mathbf{0}$ (A.5.13) s'écrit simplement :

$$d\varepsilon = Tds - p dv, \quad (\text{A.5.37})$$

ce qui nous donne avec la définition de l'entropie (A.5.30) :

$$\left(\frac{\partial \varepsilon}{\partial v} \right)_{T, \mathbf{EQ}} = T \left[\frac{\partial}{\partial v} \left(\frac{\partial f}{\partial T} \right)_{v, X_\alpha} \right]_{T, \mathbf{EQ}} - p, \quad (\text{A.5.38})$$

$$\left(\frac{\partial \varepsilon}{\partial T} \right)_{v, \mathbf{EQ}} = T \left[\frac{\partial}{\partial T} \left(\frac{\partial f}{\partial T} \right)_{v, X_\alpha} \right]_{v, \mathbf{EQ}}. \quad (\text{A.5.39})$$

La dérivation de la relation (A.5.25) fournit alors :

$$\partial_{vT}^2 f = 1 + \sum_{ir} x_i y_i^r \left[1 + \phi_i^r \left(\frac{\partial \ln y_i^r}{\partial \ln v} \right)_{T, \mathbf{EQ}} \right], \quad (\text{A.5.40})$$

$$\partial_{TT}^2 f = \frac{3}{2} + \sum_{ir} x_i y_i^r \left[\frac{3}{2} + \phi_i^r \left(\frac{\partial \ln y_i^r}{\partial \ln T} \right)_{v, \mathbf{EQ}} \right]. \quad (\text{A.5.41})$$

On voit alors qu'il suffit de déterminer une relation $y_i^r(v, T, x_i)$ pour que les dérivées (A.5.34), (A.5.35), (A.5.36), (A.5.40) et (A.5.41) nous permettent de fournir explicitement le premier exposant adiabatique Γ_1 comme une fonction de v , T et x_i . Notons que ce type de développement ne se limite pas au premier exposant adiabatique mais permet d'exprimer tous les exposants. Par exemple, pour le second exposant adiabatique Γ_2 :

$$\begin{aligned} \frac{\Gamma_2}{\Gamma_2 - 1} &\equiv \left(\frac{\partial \ln p}{\partial \ln T} \right)_{s, \mathbf{EQ}} = \frac{T}{p} \left(\frac{\partial p}{\partial T} \right)_{s, \mathbf{EQ}} = \frac{T}{p} \left| \frac{\partial(p, s, \mathbf{EQ})}{\partial(T, s, \mathbf{EQ})} \right| = \frac{T}{p} \left| \frac{\partial(p, s, \mathbf{EQ})}{\partial(T, v, \mathbf{EQ})} \cdot \frac{\partial(T, v, \mathbf{EQ})}{\partial(T, s, \mathbf{EQ})} \right| \\ &= -\frac{T}{p} \left(\frac{\partial s}{\partial v} \right)_{v, \mathbf{EQ}}^{-1} \left[\left(\frac{\partial p}{\partial v} \right)_{T, \mathbf{EQ}} \left(\frac{\partial s}{\partial T} \right)_{v, \mathbf{EQ}} - \left(\frac{\partial p}{\partial T} \right)_{v, \mathbf{EQ}} \left(\frac{\partial s}{\partial v} \right)_{T, \mathbf{EQ}} \right]. \end{aligned} \quad (\text{A.5.42})$$

Cette écriture est complètement analogue à l'équation (A.5.29) et l'on obtient :

$$\frac{\Gamma_2}{\Gamma_2 - 1} = \frac{(\partial_{TT}^2 f)(\partial_{vv}^2 f) + (\partial_{Tv}^2 f)(\partial_{vT}^2 f)}{(\partial_v f)(\partial_{vT}^2 f)} = \frac{\partial_{TT}^2 f}{\partial_{vT}^2 f} \Gamma_1. \quad (\text{A.5.43})$$

A.6 Perturbation sous contrainte

Considérons une grandeur y fonction de N variables d'états x_1, \dots, x_N ainsi qu'une contrainte sur ces variables $\mathcal{F}(x_1, \dots, x_N) = \text{cste}$. Par un théorème d'inversion local, il est possible d'écrire (au moins localement) x_1 comme une fonction des variables restantes : $x_1 = \mathcal{F}^{-1}(x_2, \dots, x_N)$. Ainsi, sur la surface définie par la contrainte, il est possible de faire coïncider y avec une grandeur z fonction de $N - 1$ variables :

$$y(x_1, \dots, x_N) = y(\mathcal{F}^{-1}(x_2, \dots, x_N), x_2, \dots, x_N) \equiv z(x_2, \dots, x_N). \quad (\text{A.6.1})$$

Ainsi, si l'on souhaite étudier la perturbation de cette grandeur δy , il semble à priori qu'il y ait deux manières de procéder.

1. Perturber y en fonction des δx_i , puis exprimer δx_1 en fonction des $\delta x_{i>1}$ grâce à la contrainte.

On a ainsi :

$$\delta y = \sum_i \frac{\partial y}{\partial x_i} \delta x_i. \quad (\text{A.6.2})$$

Or, puisque x_1 est fonction des $x_{i>1}$:

$$\delta x_1 = \sum_{i>1} \frac{\partial \mathcal{F}^{-1}}{\partial x_i} \delta x_i. \quad (\text{A.6.3})$$

On peut maintenant combiner ces deux expressions :

$$\delta y = \sum_{i>1} \left(\frac{\partial y}{\partial x_i} + \frac{\partial y}{\partial x_1} \frac{\partial \mathcal{F}^{-1}}{\partial x_i} \right) \delta x_i. \quad (\text{A.6.4})$$

2. Trouver la fonction z directement en fonction des $x_{i>1}$ à l'aide de la contrainte, puis la perturber. L'équation (A.6.1) permet d'obtenir le lien suivant entre les dérivées des deux fonctions :

$$\forall i > 1, \quad \frac{\partial z}{\partial x_i} = \frac{\partial y}{\partial x_i} + \frac{\partial y}{\partial x_1} \frac{\partial \mathcal{F}^{-1}}{\partial x_i} \quad (\text{A.6.5})$$

et donc

$$\delta z = \sum_{i>1} \frac{\partial z}{\partial x_i} \delta x_i = \sum_{i>1} \left(\frac{\partial y}{\partial x_i} + \frac{\partial y}{\partial x_1} \frac{\partial \mathcal{F}^{-1}}{\partial x_i} \right) \delta x_i = \delta y. \quad (\text{A.6.6})$$

Ce dernier résultat est rassurant puisque par définition $z \equiv y$ sur la surface définie par la contrainte. Regardons maintenant ce qu'il se passe pour une perturbation non linéaire, par exemple à l'ordre suivant :

$$\delta^2 y = \frac{1}{2} \sum_{ij} \frac{\partial^2 y}{\partial x_i \partial x_j} \delta x_i \delta x_j. \quad (\text{A.6.7})$$

En utilisant l'équation (A.6.3), cette somme se divise en 4 termes :

$$\begin{aligned} \delta^2 y &= \frac{1}{2} \sum_{ij>1} \frac{\partial^2 y}{\partial x_i \partial x_j} \delta x_i \delta x_j + \frac{1}{2} \sum_{i>1} \frac{\partial^2 y}{\partial x_i \partial x_1} \delta x_i \delta x_1 + \frac{1}{2} \sum_{j>1} \frac{\partial^2 y}{\partial x_1 \partial x_j} \delta x_1 \delta x_j + \frac{1}{2} \frac{\partial^2 y}{\partial x_1^2} \delta x_1 \delta x_1 \\ &= \frac{1}{2} \sum_{ij>1} \left[\frac{\partial^2 y}{\partial x_i \partial x_j} + \frac{\partial^2 y}{\partial x_i \partial x_1} \frac{\partial \mathcal{F}^{-1}}{\partial x_j} + \frac{\partial^2 y}{\partial x_1 \partial x_j} \frac{\partial \mathcal{F}^{-1}}{\partial x_i} + \frac{\partial^2 y}{\partial x_1^2} \frac{\partial \mathcal{F}^{-1}}{\partial x_i} \frac{\partial \mathcal{F}^{-1}}{\partial x_j} \right] \delta x_i \delta x_j. \end{aligned} \quad (\text{A.6.8})$$

Si l'on veut retrouver ce résultat en passant par z , il nous faut d'abord l'expression de sa dérivée seconde :

$$\begin{aligned} \frac{\partial^2 z}{\partial x_i \partial x_j} &= \frac{\partial}{\partial x_j} \left(\frac{\partial y}{\partial x_i} + \frac{\partial y}{\partial x_1} \frac{\partial \mathcal{T}^{-1}}{\partial x_i} \right) \\ &= \frac{\partial^2 y}{\partial x_i \partial x_j} + \frac{\partial^2 y}{\partial x_i \partial x_1} \frac{\partial \mathcal{T}^{-1}}{\partial x_j} + \frac{\partial^2 y}{\partial x_1 \partial x_j} \frac{\partial \mathcal{T}^{-1}}{\partial x_i} + \frac{\partial^2 y}{\partial x_1^2} \frac{\partial \mathcal{T}^{-1}}{\partial x_i} \frac{\partial \mathcal{T}^{-1}}{\partial x_j} + \frac{\partial y}{\partial x_1} \frac{\partial^2 \mathcal{T}^{-1}}{\partial x_i \partial x_j}. \end{aligned} \quad (\text{A.6.9})$$

On remarque donc que

$$\delta^2 z - \delta^2 y = \frac{1}{2} \sum_{ij>1} \frac{\partial y}{\partial x_1} \frac{\partial^2 \mathcal{T}^{-1}}{\partial x_i \partial x_j} \delta x_i \delta x_j \neq 0, \quad (\text{A.6.10})$$

ce qui est assez embêtant sachant que l'on est parti de $z \equiv y$. En réalité, la première façon de procéder, perturber y comme si elle dépendait de N variables alors qu'elle n'en dépend que de $N - 1$ avec la contrainte, est incorrecte. C'est pourtant ce que l'on fait souvent en thermodynamique ! En effet, il est rare en pratique de trouver les grandeurs physiques paramétrées de façon à faire explicitement apparaître les contraintes. Par exemple, l'énergie massique d'un fluide monoatomique est rarement exprimée comme une fonction de (s, ρ) (elle l'est pourtant en vertu de (A.3.8)) et on la trouvera souvent sous la forme :

$$\varepsilon = \frac{3kT}{2m} \quad (\text{A.6.11})$$

bien plus intuitive. Si l'on étudie les perturbations isentropiques de ε , on verra souvent écrit (cf. KIPPENHAHN, WEIGERT et WEISS, 2012, paragraphe 4.1 par exemple) :

$$\delta \varepsilon = \left(\frac{\partial \varepsilon}{\partial T} \right)_\rho \delta T + \left(\frac{\partial \varepsilon}{\partial (1/\rho)} \right)_T \delta \left(\frac{1}{\rho} \right) \quad (\text{A.6.12})$$

avant de chercher une relation $\delta T(\delta \rho)$ sous la contrainte $s = \text{cste}$. Puisqu'il s'agit d'une perturbation linéaire, la correspondance $\delta y = \delta z$ nous assure que cette procédure aboutira bien au bon résultat. L'équation (A.6.10) nous met cependant en garde contre une généralisation de cette approche. Pour écrire la perturbation (A.6.12), on a en effet considéré que ε était fonction de 2 variables alors qu'elle ne dépend que de la densité ρ sur la surface $s = \text{cste}$. Une telle approche peut donc mener à des incohérences dès que l'on s'intéresse à des effets non linéaires.

En dehors du cadre linéaire, on trouve deux cas où ces effets n'interviennent pas :

1. Le cas où $\mathcal{T}(x_1, \dots, x_N) = \text{cste}$ conduit à une projection affine. En effet, dans ce cas \mathcal{T}^{-1} est linéaire et

$$\forall i > 1, \forall j > 1, \frac{\partial^2 \mathcal{T}^{-1}}{\partial x_i \partial x_j} = 0 \quad \Rightarrow \quad \delta^2 y = \delta^2 z. \quad (\text{A.6.13})$$

C'est ce qu'il se passe notamment pour les contraintes du type $\sum_\alpha X_\alpha = 1$, ce qui justifie que l'on ait été assez peu précautionneux vis-à-vis de ce point dans la partie principale.

2. Le cas où les perturbations sont suffisamment faibles pour pouvoir approximer la surface par son plan tangent en (x_1, \dots, x_N) , auquel cas \mathcal{T}^{-1} est de nouveau linéaire localement. Notons que ce cas correspond souvent à $\delta^2 y \ll \delta y$.

Annexe B

Compléments sismiques

Contenu de l'annexe

B.1	Dérivation de la perturbation de l'énergie gravitationnelle	177
B.2	Dérivation des kernels	180
B.3	Obtention de l'équation d'onde	184

B.1 Dérivation de la perturbation de l'énergie gravitationnelle

On réfère par perturbation de l'énergie gravitationnelle (ou parfois terme de Cowling) au terme

$$\mathcal{G}' \equiv \int_V \rho \tilde{\xi}^* \cdot \nabla \Phi' dV \quad (\text{B.1.1})$$

apparaissant dans l'équation (2.5.4). Comme évoqué dans le paragraphe 2.5.1, cette perturbation devient négligeable pour les modes d'ordre radial n et de degré ℓ suffisamment élevés. Elle est ainsi fréquemment négligé lors de l'obtention d'une forme intégrale des fréquences propres, où l'on privilégie souvent une description asymptotique dans la limite des hautes fréquences ($n \gg 1$). On en propose cependant une expression analytique exacte ici, ne serait-ce que par soucis de complétude (et pour l'intérêt de l'exercice en soi). Commençons par remarquer que :

$$\int \xi^* \cdot \nabla \Phi' d\Omega = \xi_r \frac{d\Phi'}{dr} + \ell(\ell+1) \xi_h \frac{\Phi'}{r} \quad (\text{B.1.2})$$

en vertu des décompositions (2.4.72) et (2.4.85). Suit donc le terme de Cowling intégré sur le volume de l'étoile :

$$\mathcal{G}' = \int_0^R \rho r^2 \xi_r \frac{d\Phi'}{dr} + \rho r \ell(\ell+1) \xi_h \Phi' dr \quad (\text{B.1.3})$$

et, en intégrant le premier terme par parties :

$$\mathcal{G}' = \rho(R) R^2 \xi_r \Phi' \Big|_R + \int_0^R \Phi' \left[-\frac{d(\rho r^2 \xi_r)}{dr} + \rho r \ell(\ell+1) \xi_h \right] dr \quad (\text{B.1.4})$$

L'astuce consiste maintenant à identifier les termes avec la décomposition de la perturbation eulérienne de la densité $\rho' = -\nabla \cdot (\rho \xi)$:

$$\rho' = -\frac{1}{r^2} \frac{d(\rho r^2 \xi_r)}{dr} + \frac{\ell(\ell+1)}{r} \rho \xi_h. \quad (\text{B.1.5})$$

On obtient ainsi la forme compacte :

$$\mathcal{G}' = \int_0^R r^2 \rho' \Phi' dr + \rho(R) R^2 \xi_r \Phi' \Big|_R. \quad (\text{B.1.6})$$

Remarquons que, aussi élégante que soit cette expression, elle n'est pas beaucoup plus explicite que la forme originale puisqu'elle fait intervenir à la fois les perturbations ρ' et Φ' . Ces deux quantités ne sont cependant pas indépendantes, elles sont liées par la perturbation de l'équation de Poisson, dont la solution est donnée par l'intégrale (2.4.34). Il est important maintenant de garder à l'esprit que toutes les grandeurs que l'on manipule dans ce paragraphe sont les coefficients du développement de chaque grandeur sur les harmoniques sphériques¹ ($\rho'_{\ell m}$, $\Phi'_{\ell m}$, ...) et non pas les grandeurs elles-mêmes (ρ' , Φ' , ...). Ainsi, l'équation (2.4.34) n'est pas utilisable en l'état, il nous faut d'abord expliciter le lien entre les coefficients des développements $\rho'_{\ell m}$ et $\Phi'_{\ell m}$. Commençons par remarquer que :

$$|\mathbf{r} - \mathbf{s}|^{-1} = (r^2 - 2rs \cos \vartheta + s^2)^{-1/2}, \quad (\text{B.1.7})$$

avec ϑ l'angle entre les vecteurs \mathbf{r} et \mathbf{s} . On pourra commencer notre décomposition en utilisant la propriété fondamentale des harmoniques sphériques :

$$\forall x < 1, \quad (1 - 2x \cos \vartheta + x^2)^{-1/2} = \sum_{\ell=0}^{\infty} \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{+\ell} x^{\ell} Y_{\ell}^{m*}(\mathbf{s}) Y_{\ell}^m(\mathbf{r}). \quad (\text{B.1.8})$$

En faisant attention à la disjonction de cas (la norme de \mathbf{s} n'a aucune raison d'être inférieure à celle de \mathbf{r} si l'on ne se situe pas à la surface), on écrira donc :

$$|\mathbf{r} - \mathbf{s}|^{-1} = \sum_{\ell=0}^{\infty} \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{+\ell} \frac{r_{<}^{\ell}}{r_{>}^{\ell+1}} Y_{\ell}^{m*}(\tilde{\mathbf{r}}) Y_{\ell}^m(\mathbf{r}) \quad (\text{B.1.9})$$

où $r_{<}$ ($r_{>}$) désigne le $\min\{r, s\}$ (respectivement \max). On pourra scinder l'intégrale de l'équation (2.4.34) pour exprimer Φ' :

$$\begin{aligned} \Phi' = -4\pi G \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{+\ell} \frac{1}{2\ell+1} & \left[r^{\ell} \int_r^R s^{-\ell+1} \int \rho'(\mathbf{s}) Y_{\ell}^{m*}(\mathbf{s}) d\Omega ds \right. \\ & \left. + r^{-(\ell+1)} \int_0^r s^{\ell+2} \int \rho'(\mathbf{s}) Y_{\ell}^{m*}(\mathbf{s}) d\Omega ds \right] Y_{\ell}^m. \end{aligned} \quad (\text{B.1.10})$$

Or, par définition :

$$\rho'_{\ell m}(s) \equiv \int \rho'(\mathbf{s}) Y_{\ell}^{m*}(\mathbf{s}) d\Omega, \quad (\text{B.1.11})$$

ce qui nous conduit à l'expression :

$$\Phi' = -4\pi G \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{+\ell} \frac{1}{2\ell+1} \left[r^{\ell} \int_r^R s^{-\ell+1} \rho'_{\ell m} ds + r^{-(\ell+1)} \int_0^r s^{\ell+2} \rho'_{\ell m} ds \right] Y_{\ell}^m. \quad (\text{B.1.12})$$

Enfin, puisque la décomposition en harmoniques sphériques est unique, il nous suffit d'identifier les coefficients de la série pour obtenir le lien souhaité entre $\rho'_{\ell m}$ et $\Phi'_{\ell m}$:

$$\Phi'_{\ell m} = -\frac{4\pi G}{2\ell+1} \left[r^{\ell} \int_r^R s^{-\ell+1} \rho'_{\ell m} ds + r^{-(\ell+1)} \int_0^r s^{\ell+2} \rho'_{\ell m} ds \right]. \quad (\text{B.1.13})$$

Le premier terme de Cowling prend donc l'expression (on se dispensera à nouveau de spécifier les indices ℓ et m à partir de maintenant) :

$$\int_0^R r^2 \rho' \Phi' dr = -\frac{4\pi G}{2\ell+1} \left[\int_0^R r^{\ell+2} \rho' \int_r^R s^{-\ell+1} \rho' ds dr + \int_0^R r^{-\ell+1} \rho' \int_0^r s^{\ell+2} \rho' ds dr \right]. \quad (\text{B.1.14})$$

1. On indiquera temporairement les indices ℓ et m pour éviter les confusions dans le développement qui suivra.

Or, par interversion des intégrales :

$$\int_0^R r^{-\ell+1} \rho' \int_0^r s^{\ell+2} \rho' ds dr = \int_0^R r^{\ell+2} \rho' \int_r^R s^{-\ell+1} \rho' ds dr \quad (\text{B.1.15})$$

et les deux termes sont donc identiques :

$$\int_0^R r^2 \rho' \Phi' dr = -\frac{8\pi G}{2\ell+1} \int_0^R r^{\ell+2} \rho' \int_r^R s^{-\ell+1} \rho' ds dr. \quad (\text{B.1.16})$$

Une très belle astuce pour simplifier encore cette expression nous vient de **CHANDRASEKHAR, 1964**. Puisque :

$$\frac{d}{dr} \left[\left(\int_r^R s^{-\ell+1} \rho' ds \right)^2 \right] = -2r^{-\ell+1} \rho' \int_r^R s^{-\ell+1} \rho' ds \quad (\text{B.1.17})$$

il est possible de tirer profit d'une intégration par partie menant à l'expression :

$$\int_0^R r^2 \rho' \Phi' dr = -4\pi G \int_0^R \left(r^\ell \int_r^R s^{-\ell+1} \rho' ds \right)^2 dr, \quad (\text{B.1.18})$$

les termes intégrés s'annulant au centre en vertu de la condition de régularité $\rho' = \mathcal{O}(r^\ell)$. Il est également possible d'exploiter la décomposition (B.1.13) pour exprimer le terme de surface :

$$\rho(R) R^2 \xi_r \Phi' |_R = -\frac{4\pi G \rho(R)}{(2\ell+1)} \xi_r |_R R^{-\ell+1} \int_0^R r^{\ell+2} \rho' dr, \quad (\text{B.1.19})$$

ce que l'on réécrira en faisant intervenir la masse de l'étoile $M = 4\pi \int_0^R r^2 \rho dr$. Finalement la perturbation de l'énergie gravitationnelle dans son ensemble prend la forme suivante :

$$\mathcal{G}' = -4\pi G \int_0^R \left(r^\ell \int_r^R s^{-\ell+1} \rho' ds \right)^2 dr - GM \rho(R) R \xi_r |_R j_\ell. \quad (\text{B.1.20})$$

avec la définition :

$$j_\ell \equiv \frac{1}{2\ell+1} \frac{\int_0^1 x^{\ell+2} \rho' dx}{\int_0^1 x^2 \rho dx}. \quad (\text{B.1.21})$$

Compte tenu du terme de surface apparaissant lors du développement mené dans le paragraphe 2.5.2, on peut chercher à savoir comment se compare le second terme dérivé dans l'équation (B.1.20), ce qui pose la question de la comparaison entre Rj_ℓ et $\xi_r |_R$. En remarquant que les j_ℓ décroissent avec ℓ^2 et que :

$$j_0 = \frac{\int_0^1 x^2 \rho' dx}{\int_0^1 x^2 \rho dx} = -3 \frac{\xi_r |_R \rho(R)}{R \langle \rho \rangle} \quad (\text{B.1.22})$$

en intégrant par partie la relation (B.1.5) pour $\ell = 0$ avec $\langle \rho \rangle$ la densité moyenne dans l'étoile. Bien que l'on ait supposé $\rho(R) \neq 0$ dans le cas général, la densité en surface est d'ordinaire très faible comparée à la densité moyenne de l'étoile et l'on procédera donc à l'approximation

2. Cela est dû au fait que la fonction x^ℓ diminue en tout point sur $[0,1]$ si ℓ augmente (en plus du coefficient $1/(2\ell+1)$ en facteur).

$Rj_\ell \ll \xi_r|_R$ dans la partie principale. Sous cette hypothèse :

$$\mathcal{G}' = -4\pi G \int_0^R \left(r^\ell \int_r^R s^{-\ell+1} \rho' ds \right)^2 dr. \quad (\text{B.1.23})$$

B.2 Dérivation des kernels

On propose, dans ce paragraphe, de donner les détails quant à la dérivation de l'équation (2.5.32) et des noyaux d'inversion qui y figurent à partir de la forme intégrale (2.5.19). On ne fera pas apparaître explicitement les termes en $\delta_x \xi_r$ et $\delta_x \xi_h$, dont on en assurera de l'annulation par le principe variationnel. Commençons par remarquer qu'une intégrale sur r se perturbe de la façon suivante :

$$\begin{aligned} \delta \int_0^R A(r) dr &= (\delta R) \int_0^1 A(x) dx + R \int_0^1 \delta_x A(x) dx \\ &= \frac{\delta R}{R} \int_0^R A(r) dr + \int_0^R (\delta_x A)(r) dr. \end{aligned} \quad (\text{B.2.1})$$

Ainsi pour une expression s'écrivant comme

$$\int_0^R A(r) dr = \int_0^R B(r) dr + C(R), \quad (\text{B.2.2})$$

on obtiendra :

$$\int_0^R (\delta_x A)(r) dr = \int_0^R (\delta_x B)(r) dr + C(R) \left[\left. \frac{\delta C}{C} \right|_R - \frac{\delta R}{R} \right] \quad (\text{B.2.3})$$

en utilisant le fait que $\int_0^R B(r) dr - \int_0^R A(r) dr = -C(R)$. En remarquant que δ_x commute avec les opérateurs $r \frac{d}{dr}$ et $\frac{1}{r} \frac{d}{dr} r^2$ ainsi que les quantités ξ_r et $r\Lambda$, on voit que les grandeurs dont les perturbations restent à expliciter sont :

$$\delta_x(\rho r^2) = \rho r^2 \left(\frac{\delta_x \rho}{\rho} + 2 \frac{\delta R}{R} \right) \quad (\text{B.2.4})$$

$$\delta_x(\rho c^2) = \rho c^2 \left(\frac{\delta_x \rho}{\rho} + \frac{\delta_x c^2}{c^2} \right) \quad (\text{B.2.5})$$

$$\delta_x \left(r \frac{dp}{dr} \right) = \delta_x \left(\frac{-Gm\rho}{r} \right) = -\frac{G\rho}{r} \int_0^r 4\pi s^2 \rho \frac{\delta_x \rho}{\rho} ds + r \frac{dp}{dr} \left(\frac{\delta_x \rho}{\rho} + 2 \frac{\delta R}{R} \right) \quad (\text{B.2.6})$$

$$\delta_x \left(\frac{r}{\rho} \frac{d\rho}{dr} \right) = r \frac{d}{dr} \left(\frac{\delta_x \rho}{\rho} \right) \quad (\text{B.2.7})$$

$$\begin{aligned} \delta_x(r\rho') &= -\frac{1}{r} \frac{d}{dr} \left(r^2 \rho \xi_r \frac{\delta_x \rho}{\rho} \right) + \ell(\ell+1) \rho \xi_h \frac{\delta_x \rho}{\rho} \\ &= -r \left[\rho \Lambda + \frac{d\rho}{dr} \xi_r \right] \frac{\delta_x \rho}{\rho} - r \rho \xi_r \frac{d}{dr} \left(\frac{\delta_x \rho}{\rho} \right) \end{aligned} \quad (\text{B.2.8})$$

On voit que toutes les perturbations de structure peuvent s'exprimer uniquement en fonction de $\delta_x \rho / \rho$ et $\delta_x c^2 / c^2$ (en plus de $\delta R / R$). Pour que ces perturbations n'apparaissent plus dans une dérivée ou sous une intégrale, on pourra utiliser des intégrations par partie et des interversions

d'intégrales. À titre d'exemple, la manipulation :

$$\begin{aligned}
\int_0^R 2r\xi_r\Lambda\delta_x\left(r\frac{dp}{dr}\right)dr &= \int_0^R -2G\rho\xi_r\Lambda\int_0^r 4\pi s^2\rho\frac{\delta_x\rho}{\rho}dsdr \\
&+ \int_0^R 2r^2\xi_r\Lambda\frac{dp}{dr}\frac{\delta_x\rho}{\rho}dr + 2\frac{\delta R}{R}\int_0^R 2r^2\xi_r\Lambda\frac{dp}{dr}dr \\
&= \int_0^R \left[2r^2\xi_r\Lambda\frac{dp}{dr} - 4\pi G\rho r^2\int_r^R 2\rho\xi_r\Lambda ds\right]\frac{\delta_x\rho}{\rho}dr \\
&+ 2\frac{\delta R}{R}\int_0^R 2r^2\xi_r\Lambda\frac{dp}{dr}dr
\end{aligned} \tag{B.2.9}$$

permet de mettre en facteur $\delta_x\rho/\rho$. De la même façon et avec un peu plus de travail :

$$\begin{aligned}
\int_0^R \xi_r^2\delta_x\left(\frac{r^2}{\rho}\frac{dp}{dr}\frac{dp}{dr}\right)dr &= \int_0^R r^2\frac{dp}{dr}\xi_r^2\frac{d}{dr}\left(\frac{\delta_x\rho}{\rho}\right) - G\xi_r^2\frac{dp}{dr}\int_0^r 4\pi s^2\rho\frac{\delta_x\rho}{\rho}dsdr \\
&+ \int_0^R \frac{r^2}{\rho}\frac{dp}{dr}\frac{dp}{dr}\xi_r^2\frac{\delta_x\rho}{\rho}dr + 2\frac{\delta R}{R}\int_0^R \frac{r^2}{\rho}\frac{dp}{dr}\frac{dp}{dr}\xi_r^2dr \\
&= -\int_0^R \left[\frac{d}{dr}\left(r^2\frac{dp}{dr}\xi_r^2\right) - \frac{r^2}{\rho}\frac{dp}{dr}\frac{dp}{dr}\xi_r^2 + 4\pi G\rho r^2\int_r^R \xi_r^2\frac{dp}{ds}ds\right]\frac{\delta_x\rho}{\rho}dr \\
&- GM\xi_r^2\delta_x\rho|_R + 2\frac{\delta R}{R}\int_0^R \frac{r^2}{\rho}\frac{dp}{dr}\frac{dp}{dr}\xi_r^2dr \\
&= -\int_0^R \left[\rho\frac{d}{dr}\left(\frac{r^2}{\rho}\frac{dp}{dr}\xi_r^2\right) + 4\pi G\rho r^2\int_r^R \xi_r^2\frac{dp}{ds}ds\right]\frac{\delta_x\rho}{\rho}dr \\
&- GM\xi_r^2\delta\rho|_R + 2\frac{\delta R}{R}\int_0^R \frac{r^2}{\rho}\frac{dp}{dr}\frac{dp}{dr}\xi_r^2dr.
\end{aligned} \tag{B.2.10}$$

Le terme qui nécessite le plus de manipulations reste toujours la perturbation de l'énergie gravitationnelle. On procédera donc par étapes. Commençons par remarquer que :

$$\begin{aligned}
\int_r^R s^{-\ell}\delta_y(s\rho')ds &= -\int_r^R s^{-\ell+1}\left[\rho\Lambda + \frac{d\rho}{ds}\xi_r\right]\frac{\delta_y\rho}{\rho}ds - \int_r^R s^{-\ell+1}\rho\xi_r\frac{d}{ds}\left(\frac{\delta_y\rho}{\rho}\right)ds \\
&= \int_r^R \left[\frac{d}{ds}\left(s^{-\ell+1}\rho\xi_r\right) - s^{-\ell+1}\left(\rho\Lambda + \frac{d\rho}{ds}\xi_r\right)\right]\frac{\delta_y\rho}{\rho}ds - R^{-\ell+1}\xi_r\delta\rho|_R + r^{-\ell+1}\rho\xi_r\frac{\delta_x\rho}{\rho} \\
&= -(\ell+1)\int_r^R s^{-\ell}\rho(\xi_r + \ell\xi_h)\frac{\delta_y\rho}{\rho}ds + r^{-\ell+1}\rho\xi_r\frac{\delta_x\rho}{\rho} - R^{-\ell+1}\xi_r\delta\rho|_R
\end{aligned} \tag{B.2.11}$$

où $y \equiv s/R$ est l'analogue de $x \equiv r/R$. En introduisant la notation

$$Q_\ell(r) = r^\ell \int_r^R s^{-\ell+1}\rho' ds \tag{B.2.12}$$

pour faciliter le calcul, et puisque :

$$\delta_x Q_\ell = \frac{\delta R}{R}Q_\ell + r^\ell \int_r^R s^{-\ell}\delta_y(s\rho') ds, \tag{B.2.13}$$

on pourra écrire, à l'aide de l'équation (B.2.11) :

$$\begin{aligned}
\int_0^R \delta_x(Q_\ell^2) dr &= 2 \int_0^R r^\ell Q_\ell \left[-(\ell+1) \int_r^R s^{-\ell} \rho(\xi_r + \ell \xi_h) \frac{\delta_y \rho}{\rho} ds + r^{-\ell+1} \rho \xi_r \frac{\delta_x \rho}{\rho} \right] dr \\
&\quad - 2 \left(\int_0^R r^\ell Q_\ell dr \right) R^{-\ell+1} \xi_r \delta \rho|_R + 2 \left(\int_0^R Q_\ell^2 dr \right) \frac{\delta R}{R} \\
&= 2 \int_0^R \rho \left[r Q_\ell \xi_r - (\ell+1) r^{-\ell} (\xi_r + \ell \xi_h) \int_0^r s^\ell Q_\ell ds \right] \frac{\delta_x \rho}{\rho} dr \\
&\quad - 2 \left(\int_0^R r^\ell Q_\ell dr \right) R^{-\ell+1} \xi_r \delta \rho|_R + 2 \left(\int_0^R Q_\ell^2 dr \right) \frac{\delta R}{R}.
\end{aligned} \tag{B.2.14}$$

Intéressons-nous au premier terme de surface de cette équation en explicitant :

$$\begin{aligned}
\int_0^R r^\ell Q_\ell dr &= -\frac{1}{2\ell+1} \int_0^R r^{2\ell+1} \frac{dr^{-\ell} Q_\ell}{dr} dr \\
&= \frac{1}{2\ell+1} \int_0^R r^{\ell+2} \rho' dr = \frac{MR^\ell j_\ell}{4\pi}
\end{aligned} \tag{B.2.15}$$

à l'aide des définitions (B.2.12) de Q_ℓ et (B.1.21) de j_ℓ . Ainsi le terme de surface associé qui apparaîtra dans l'expression finale sera donné par :

$$-4\pi G \left(-2 \frac{MR^\ell j_\ell}{4\pi} \right) R^{-\ell+1} \xi_r \delta \rho|_R = 2GMR j_\ell \xi_r \delta \rho|_R. \tag{B.2.16}$$

L'importance de ce terme de surface est difficile à quantifier. Pour les mêmes arguments que ceux donnés dans le paragraphe 2.5.2, il est clair que ce dernier est négligeable devant un terme comme $-GM \xi_r^2 \delta \rho|_R$, apparaissant dans l'équation (B.2.10). Or, il se trouve que cet autre terme dominant en $\delta \rho|_R$ se compense exactement avec celui résultant de la perturbation du terme de surface $GM \rho \xi_r|_R$ de l'équation (2.5.19). Ainsi le résultat de l'équation (B.2.16) sera à comparer avec des termes de l'ordre $G\rho \xi_r^2|_R \delta M$, ce qui pose la question de savoir si :

$$R j_\ell \frac{\delta \rho}{\rho} \Big|_R \stackrel{?}{\ll} \xi_r|_R \frac{\delta M}{M}. \tag{B.2.17}$$

Si l'on a déjà montré que $R j_\ell \ll \xi_r|_R$, le fait que ρ atteigne des valeurs très proches de 0 en surface fait que l'on peut avoir des $\delta \rho/\rho|_R$ qui s'approche de l'unité, là où $\delta M/M \ll 1$ pour des étoiles raisonnablement similaires. À noter que si l'on avait conservé le second terme de surface (négligeable dans la structure à l'équilibre) de l'équation (B.1.20), il aurait contribué à hauteur de $-GMR j_\ell \xi_r \delta \rho|_R$, ce qui ne compense que la moitié de l'équation (B.2.16). On négligera par la suite ce terme qu'il est difficile de quantifier, bien qu'il soit utile de garder son existence à l'esprit. En introduisant la notation \mathcal{I} pour désigner l'inertie du mode

$$\mathcal{I} \equiv \int_0^R \rho r^2 (\xi_r^2 + \ell(\ell+1) \xi_h^2) dr, \tag{B.2.18}$$

les équations (B.2.9), (B.2.10) et (B.2.14) nous permettent d'expliciter complètement la perturbation de l'équation (2.5.19) :

$$\begin{aligned}
\delta\omega^2\mathcal{I} = & \int_0^R \rho c^2 r^2 \Lambda^2 \frac{\delta_x c^2}{c^2} + \left[\rho c^2 r^2 \Lambda^2 + 2r^2 \xi_r \Lambda \frac{dp}{dr} - 4\pi G \rho r^2 \int_r^R 2\rho \xi_r \Lambda ds \right] \frac{\delta_x \rho}{\rho} dr \\
& - \int_0^R \left[\rho \frac{d}{dr} \left(\frac{r^2}{\rho} \frac{dp}{dr} \xi_r^2 \right) + 4\pi G \rho r^2 \int_r^R \xi_r^2 \frac{d\rho}{ds} ds \right] \frac{\delta_x \rho}{\rho} dr \\
& - 8\pi G \int_0^R \rho \left[r^{\ell+1} \xi_r \int_r^R s^{-\ell+1} \rho' ds - (\ell+1)r^{-\ell} (\xi_r + \ell \xi_h) \int_0^r s^{2\ell} \int_s^R t^{-\ell+1} \rho' dt ds \right] \frac{\delta_x \rho}{\rho} dr \\
& - \omega^2 \int_0^R \rho r^2 (\xi_r^2 + \ell(\ell+1)\xi_h^2) \frac{\delta_x \rho}{\rho} dr + G\rho(R) \delta M \xi_r^2|_R \\
& + 2 \frac{\delta R}{R} \left[\int_0^R 2r^2 \xi_r \Lambda \frac{dp}{dr} + \frac{r^2}{\rho} \frac{d\rho}{dr} \frac{dp}{dr} \xi_r^2 - 4\pi G \left(r^\ell \int_r^R s^{-\ell+1} \rho' ds \right)^2 dr - \omega^2 \mathcal{I} \right] \\
& - \frac{\delta R}{R} GM\rho(R) \xi_r^2|_R
\end{aligned} \tag{B.2.19}$$

où le tout dernier terme provient de la propriété (B.2.3). Il est possible de procéder à quelques simplifications. Premièrement, en remarquant que $\frac{r^2}{\rho} \frac{dp}{dr} = -Gm$:

$$2r^2 \xi_r \Lambda \frac{dp}{dr} - \rho \frac{d}{dr} \left(\frac{r^2}{\rho} \frac{dp}{dr} \xi_r^2 \right) = -2\rho r^2 g \xi_r \left(\Lambda - \frac{d\xi_r}{dr} \right) + 4\pi G \rho^2 r^2 \xi_r^2, \tag{B.2.20}$$

ce qui rend l'ensemble des termes factorisables par ρr^2 . Notons ensuite que la plupart des termes ayant pour facteur $\delta R/R$ s'annulent en vertu de l'équation (2.5.19) elle-même. En effet :

$$\int_0^R 2r^2 \xi_r \Lambda \frac{dp}{dr} + \frac{r^2}{\rho} \frac{d\rho}{dr} \frac{dp}{dr} \xi_r^2 - 4\pi G \left(r^\ell \int_r^R s^{-\ell+1} \rho' ds \right)^2 dr - \omega^2 \mathcal{I} = -2 \int_0^R \rho c^2 r^2 \Lambda^2 dr - 2 GM\rho(R) \xi_r^2|_R \tag{B.2.21}$$

ne laissant que

$$GM\rho(R) \xi_r^2|_R \left[\frac{\delta M}{M} - 3 \frac{\delta R}{R} \right] - 2 \frac{\delta R}{R} \int_0^R \rho c^2 r^2 \Lambda^2 dr \tag{B.2.22}$$

pour ce qui est des termes de surfaces. On effectuera alors la manipulation suivante :

$$\begin{aligned}
\int_r^R 2\rho \xi_r \Lambda + \xi_r^2 \frac{d\rho}{ds} ds &= \int_r^R 2\rho \xi_r \left(\Lambda - \frac{d\xi_r}{ds} \right) ds - \int_r^R 2\rho \xi_r \frac{d\xi_r}{ds} + \xi_r^2 \frac{d\rho}{ds} ds \\
&= \int_r^R 2\rho \xi_r \left(\Lambda - \frac{d\xi_r}{ds} \right) ds + \rho \xi_r^2 - \rho(R) \xi_r|_R.
\end{aligned} \tag{B.2.23}$$

Puisque :

$$-4\pi G \rho(R) \xi_r|_R \int_0^R r^2 \delta_x \rho dr = -GM\rho(R) \xi_r^2|_R \left[\frac{\delta M}{M} - 3 \frac{\delta R}{R} \right], \tag{B.2.24}$$

le dernier terme apparaissant dans l'équation (B.2.23) annule la première partie des termes de surface (cf. équation (B.2.22)). Finalement, en combinant ces simplifications et en remarquant que $\delta\omega/\omega = (1/2)\delta\omega^2/\omega^2$, il est possible d'écrire l'équation (B.2.19) sous la forme suivante :

$$\frac{\delta\omega}{\omega} = \int_0^R \mathcal{K}_{c^2}(r, \omega) \frac{\delta_x c^2}{c^2} + \mathcal{K}_\rho(r, \omega) \frac{\delta_x \rho}{\rho} dr - \mathcal{S}_R(\omega) \frac{\delta R}{R} \tag{B.2.25}$$

avec :

$$\mathcal{K}_{c^2} = \frac{\rho r^2 c^2 \Lambda^2}{2\omega^2 \mathcal{I}}, \quad (\text{B.2.26})$$

$$\mathcal{K}_\rho = \frac{\rho r^2}{2\omega^2 \mathcal{I}} \left[c^2 \Lambda^2 - \omega^2 (\xi_r^2 + \ell(\ell+1)\xi_h^2) - 2g\xi_r \left(\Lambda - \frac{d\xi_r}{dr} \right) - 4\pi G \int_r^R 2\rho \xi_r \left(\Lambda - \frac{d\xi_r}{ds} \right) ds \right. \\ \left. + 8\pi G \left(\rho \xi_r^2 - r^{\ell-1} \xi_r \int_r^R s^{-\ell+1} \rho' ds + (\ell+1)r^{-\ell-2} (\xi_r + \ell \xi_h) \int_0^r s^{2\ell} \int_s^R t^{-\ell+1} \rho' dt ds \right) \right], \quad (\text{B.2.27})$$

$$\mathcal{S}_R = \frac{1}{\omega^2 \mathcal{I}} \int_0^R \rho r^2 \Lambda^2 c^2 dr, \quad (\text{B.2.28})$$

résultat que l'on retrouvera dans la partie principale du manuscrit.

B.3 Obtention de l'équation d'onde

L'objectif est ici d'obtenir l'équation régissant le comportement de la variable $c\Lambda$ (où la variable Λ désigne la divergence du champ de déplacement), apparaissant dans chacun des termes de l'équation (B.2.26) et liée à la perturbation lagrangienne de la pression δp par :

$$c\Lambda = -\frac{\delta p}{\rho c}. \quad (\text{B.3.1})$$

B.3.1 Cas général

On partira pour cela du système d'équations (2.4.87)-(2.4.88), réécrit ci-dessus en omettant le degré ℓ et l'ordre azimutal m auxquels est associée chacune des variables ξ_r , ξ_h et Φ' :

$$\omega^2 \xi_r = \frac{d}{dr} (\Phi' - c^2 \Lambda + g \xi_r) + \left(\frac{\Gamma_1}{\gamma} - 1 \right) g \Lambda \quad (\text{B.3.2})$$

$$\omega^2 \xi_h = \frac{1}{r} (\Phi' - c^2 \Lambda + g \xi_r). \quad (\text{B.3.3})$$

À noter que la perturbation du potentiel de gravitation, Φ' , vérifie une équation de Poisson dont on a donné une solution explicite en fonction de ρ' en équation (B.1.13). Le premier objectif va être de réécrire le système (B.3.2)-(B.3.3) en fonctions des trois variables ξ_r , $c\Lambda$ et Φ'/c . On éliminera donc la composante horizontale du déplacement à l'aide de l'expression de la divergence :

$$\frac{L^2 \xi_h}{r} = \frac{d\xi_r}{dr} + \frac{2\xi_r}{r} - \Lambda \quad (\text{B.3.4})$$

où l'on a défini $L \equiv \sqrt{\ell(\ell+1)}$. L'équation (B.3.3) pourra donc s'écrire comme :

$$\frac{d\xi_r}{dr} + \frac{2\xi_r}{r} - \Lambda = \left(\frac{L}{r\omega} \right)^2 (\Phi' - c^2 \Lambda + g \xi_r). \quad (\text{B.3.5})$$

Pour faire apparaître les variables recherchées, la solution naturelle est de multiplier l'équation par c , ce qui fait ressortir à l'occasion la variable canonique sur laquelle s'exprime l'équation d'onde, la profondeur acoustique τ :

$$\frac{d\xi_r}{d\tau} + \left[1 - \left(\frac{Lc}{r\omega} \right)^2 \right] c\Lambda + \left[\left(\frac{Lc}{r\omega} \right)^2 \frac{g}{c} - \frac{2c}{r} \right] \xi_r = - \left(\frac{Lc}{r\omega} \right)^2 \frac{\Phi'}{c}, \quad (\text{B.3.6})$$

en utilisant la propriété $cd\tau = -dr$. Le rapport de la fréquence de Lamb et de la fréquence propre apparaît à de nombreuses reprises, on en propose l'abréviation :

$$\mathcal{L} \equiv \frac{Lc}{r\omega}. \quad (\text{B.3.7})$$

Cette grandeur nous permet de définir ce à quoi correspond le comportement asymptotique des modes d'oscillation ; il s'agit de leur limite lorsque $\mathcal{L} \rightarrow 0$. L'équation (B.3.6) constitue la première des deux équations liant ξ_r à $c\Lambda$, la seconde provenant de (B.3.2). Commençons par remarquer que :

$$\frac{d\Phi'}{dr} = -\frac{d}{d\tau} \left(\frac{\Phi'}{c} \right) + \frac{dc}{dr} \frac{\Phi'}{c} \quad (\text{B.3.8})$$

$$\frac{dc^2\Lambda}{dr} = -\frac{dc\Lambda}{d\tau} + \frac{dc}{dr} c\Lambda \quad (\text{B.3.9})$$

$$\frac{dg\xi_r}{dr} = -\frac{g}{c} \frac{d\xi_r}{d\tau} + \frac{dg}{dr} \xi_r. \quad (\text{B.3.10})$$

On obtient ainsi :

$$\frac{dc\Lambda}{d\tau} + \left[\left(\frac{\Gamma_1}{\gamma} - 1 \right) \frac{g}{c} - \frac{dc}{dr} \right] c\Lambda - \frac{g}{c} \frac{d\xi_r}{d\tau} + \left[\frac{dg}{dr} - \omega^2 \right] \xi_r = \frac{d}{d\tau} \left(\frac{\Phi'}{c} \right) - \frac{dc}{dr} \frac{\Phi'}{c}. \quad (\text{B.3.11})$$

On peut alors éliminer $d\xi_r/d\tau$ par son expression (B.3.6), menant à l'équation :

$$\frac{dc\Lambda}{d\tau} + \left[\left(\frac{\Gamma_1}{\gamma} - \mathcal{L}^2 \right) \frac{g}{c} - \frac{dc}{dr} \right] c\Lambda + \left[\frac{dg}{dr} - \frac{2g}{r} + \left(\frac{g}{c} \right)^2 \mathcal{L}^2 - \omega^2 \right] \xi_r = \frac{d}{d\tau} \left(\frac{\Phi'}{c} \right) - \left[\frac{g}{c} \mathcal{L}^2 + \frac{dc}{dr} \right] \frac{\Phi'}{c}. \quad (\text{B.3.12})$$

Sans même parler de la résoudre, donner une forme explicite à l'équation que suit $c\Lambda$ à partir du système (B.3.12)-(B.3.6) est loin d'être chose aisée. On passera donc en notation symbolique pour le développement qui suit. Remarquons d'abord que le système est écrit sous la forme :

$$y' + g_1y + g_2z = u' + d_1u, \quad (\text{B.3.13})$$

$$z' + g_3y + g_4z = d_2u, \quad (\text{B.3.14})$$

où y, z et u désignent nos trois variables $c\Lambda, \xi_r$ et Φ'/c ; les termes de gauche sont définis par :

$$g_1 \equiv \left(\frac{\Gamma_1}{\gamma} - \mathcal{L}^2 \right) \frac{g}{c} - \frac{dc}{dr}, \quad g_2 \equiv \frac{dg}{dr} - \frac{2g}{r} + \left(\frac{g}{c} \right)^2 \mathcal{L}^2 - \omega^2, \quad (\text{B.3.15})$$

$$g_3 \equiv 1 - \mathcal{L}^2, \quad g_4 \equiv \frac{g}{c} \mathcal{L}^2 - \frac{2c}{r}, \quad (\text{B.3.16})$$

et ceux de droite par :

$$d_1 \equiv - \left[\frac{g}{c} \mathcal{L}^2 + \frac{dc}{dr} \right], \quad d_2 \equiv -\mathcal{L}^2. \quad (\text{B.3.17})$$

Commençons par dériver la relation (B.3.13) :

$$y'' + g_1y' + g_1'y + g_2z' + g_2'z = u'' + d_1u' + d_1'u. \quad (\text{B.3.18})$$

En remplaçant y' et z' à l'aide de leurs expressions (B.3.13) et (B.3.14) :

$$y'' + (g_1' - g_1^2 - g_2g_3)y + (g_2' - g_1g_2 - g_2g_4)z = u'' + (d_1 - g_1)u' + (d_1' - g_1d_1 - g_2d_2)u. \quad (\text{B.3.19})$$

On peut éliminer z de cette équation en remarquant que :

$$\begin{aligned} (g_2' - g_1 g_2 - g_2 g_4)z &= (\ln' g_2 - g_1 - g_4)g_2 z \\ &= (\ln' g_2 - g_1 - g_4)(-y' - g_1 y + u' + d_1 u) \end{aligned} \quad (\text{B.3.20})$$

avec $\ln' g_2 \equiv g_2'/g_2$ et où l'on a remplacé $g_2 z$ à l'aide de l'équation (B.3.13). Cette substitution conduit à une équation liant directement y à u :

$$\begin{aligned} y'' + (g_1 + g_4 - \ln' g_2)y' + (g_1' + g_1 g_4 - g_2 g_3 - g_1 \ln' g_2)y \\ = u'' + (d_1 + g_4 - \ln' g_2)u' + (d_1' + g_4 d_1 - g_2 d_2 - d_1 \ln' g_2)u. \end{aligned} \quad (\text{B.3.21})$$

On pourra en donner une expression un peu plus compacte en définissant $\text{Det}_G = g_1 g_4 - g_2 g_3$, $\text{Det}_D = g_4 d_1 - g_2 d_2$ et en remarquant que $g_1' - g_1 \ln' g_2 = g_2(g_1/g_2)'$:

$$y'' + (g_1 + g_4 - \ln' g_2)y' + (\text{Det}_G + g_2(g_1/g_2)')y = u'' + (d_1 + g_4 - \ln' g_2)u' + (\text{Det}_D + g_2(d_1/g_2)')u. \quad (\text{B.3.22})$$

La prochaine étape sera de proposer un changement de variable permettant de se débarrasser de la dérivée première y' (on souhaite se rapprocher le plus possible de la forme standard de l'équation de Schrödinger). En revenant aux définitions de g_1 et g_4 , on peut remarquer que :

$$g_1 + g_4 = \frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} - \frac{2c}{r} = \ln' \rho + \ln' c + \ln' r^2 = \ln' \lambda^2 \quad (\text{B.3.23})$$

où λ est défini par l'équation :

$$\lambda^2 \equiv -\frac{dm}{d\tau} = 4\pi r^2 \rho c. \quad (\text{B.3.24})$$

De façon assez remarquable, il est alors possible d'écrire :

$$\begin{aligned} y'' + (g_1 + g_4 - \ln' g_2)y' &= y'' + 2 \ln'(\lambda/\sqrt{g_2})y' \\ &= \frac{\sqrt{g_2}}{\lambda} \left[\frac{\lambda}{\sqrt{g_2}} y'' + 2 \left(\frac{\lambda}{\sqrt{g_2}} \right)' y' \right] \\ &= \frac{\sqrt{g_2}}{\lambda} \left[\Xi'' - \frac{\sqrt{g_2}}{\lambda} \left(\frac{\lambda}{\sqrt{g_2}} \right)'' \Xi \right] \end{aligned} \quad (\text{B.3.25})$$

en ayant préalablement défini la nouvelle variable :

$$\Xi \equiv \frac{\lambda y}{\sqrt{g_2}} = \left[\frac{4\pi r^2 \rho c}{\frac{dg}{dr} - \frac{2g}{r} + \left(\frac{g}{c}\right)^2 \mathcal{L}^2 - \omega^2} \right]^{1/2} c \Lambda. \quad (\text{B.3.26})$$

En multipliant l'équation (B.3.22) par $\lambda/\sqrt{g_2}$, on pourra alors écrire :

$$\Xi'' + \left[\text{Det}_G + g_2 \left(\frac{g_1}{g_2} \right)' - \frac{\sqrt{g_2}}{\lambda} \left(\frac{\lambda}{\sqrt{g_2}} \right)'' \right] \Xi = \frac{\lambda}{\sqrt{g_2}} [u'' + (d_1 + g_4 - \ln' g_2)u' + (\text{Det}_D + g_2(d_1/g_2)')u]. \quad (\text{B.3.27})$$

On pourra accessoirement remarquer qu'il est possible de procéder de la même façon à droite. En effet :

$$d_1 + g_4 = -\frac{dc}{dr} - \frac{2c}{r} = \ln' \rho c. \quad (\text{B.3.28})$$

En procédant de manière analogue, on voit donc que :

$$u'' + (d_1 + g_4 - \ln' g_2)u' = \frac{r\sqrt{g_2}}{\lambda} \left[\Psi'' - \frac{r\sqrt{g_2}}{\lambda} \left(\frac{\lambda}{r\sqrt{g_2}} \right)'' \Psi \right], \quad (\text{B.3.29})$$

avec :

$$\Psi \equiv \frac{\lambda u}{r\sqrt{g_2}} = \left[\frac{4\pi\rho c}{\frac{dg}{dr} - \frac{2g}{r} + \left(\frac{g}{c}\right)^2 \mathcal{L}^2 - \omega^2} \right]^{1/2} \frac{\Phi'}{c}. \quad (\text{B.3.30})$$

L'équation régissant le comportement de l'onde est alors donnée par une équation agréable-ment symétrique :

$$\Xi'' + \left[\text{Det}_G + g_2 \left(\frac{g_1}{g_2} \right)' - \frac{\sqrt{g_2}}{\lambda} \left(\frac{\lambda}{\sqrt{g_2}} \right)'' \right] \Xi = r\Psi'' + \left[\text{Det}_D + g_2 \left(\frac{d_1}{g_2} \right)' - \frac{r\sqrt{g_2}}{\lambda} \left(\frac{\lambda}{r\sqrt{g_2}} \right)'' \right] r\Psi. \quad (\text{B.3.31})$$

B.3.2 Approximation de Cowling

En l'état, l'équation (B.3.31) est extrêmement complexe à étudier. L'approximation de Cowling consiste à simplifier cette relation dans la limite où $\Psi \rightarrow 0$:

$$\Xi'' + \left[\text{Det}_G + g_2 \left(\frac{g_1}{g_2} \right)' - \frac{\sqrt{g_2}}{\lambda} \left(\frac{\lambda}{\sqrt{g_2}} \right)'' \right] \Xi = 0. \quad (\text{B.3.32})$$

L'équation obtenue est alors une équation homogène du second ordre ne faisant intervenir que la dérivée seconde de Ξ . En revenant aux définitions des g_i , on voit que cette relation reste difficile à expliciter malgré sa forme symbolique (B.3.32) relativement concise. Par exemple :

$$\begin{aligned} \text{Det}_G &\equiv g_1 g_4 - g_3 g_2 \\ &= \left[\left(\frac{\Gamma_1}{\gamma} - \mathcal{L}^2 \right) \frac{g}{c} - \frac{dc}{dr} \right] \left[\frac{g}{c} \mathcal{L}^2 - \frac{2c}{r} \right] - (1 - \mathcal{L}^2) \left[\frac{dg}{dr} - \frac{2g}{r} + \left(\frac{g}{c} \right)^2 \mathcal{L}^2 - \omega^2 \right] \\ &= \left[\omega^2 - (\mathcal{L}\omega)^2 - \frac{dg}{dr} + \frac{2g}{r} - \frac{2c}{r} \left(\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} \right) \right] + \mathcal{L}^2 \left[\frac{dg}{dr} - \frac{g}{c} \frac{dc}{dr} + \left(\frac{\Gamma_1}{\gamma} - 1 \right) \left(\frac{g}{c} \right)^2 \right] \\ &= \left[\omega^2 - S_\ell^2 - \left(\frac{\Gamma_1}{\gamma} - 1 \right) \frac{2g}{r} - \frac{dg}{dr} + \frac{2c}{r} \frac{dc}{dr} \right] + \mathcal{L}^2 \left[\frac{dg}{dr} - \frac{g}{c} \frac{dc}{dr} + \mathcal{N}^2 \right] \end{aligned} \quad (\text{B.3.33})$$

avec $S_\ell \equiv \mathcal{L}\omega$ la fréquence de Lamb, $\mathcal{N}^2 \equiv (\Gamma_1/\gamma - 1)(g/c)^2$ la fréquence de Brunt-Väisälä au carré et où les termes en \mathcal{L}^4 se sont compensés. Pour obtenir la forme générale de l'équation d'onde on effectuera les remarques qui suivent. Tout d'abord :

$$g_2 \left(\frac{g_1}{g_2} \right)' = g_1' - \frac{g_1 g_2'}{g_2} \quad (\text{B.3.34})$$

et

$$\frac{\sqrt{g_2}}{\lambda} \left(\frac{\lambda}{\sqrt{g_2}} \right)'' = \frac{\lambda''}{\lambda} - \frac{\lambda' g_2'}{\lambda g_2} + \frac{3}{4} \left(\frac{g_2'}{g_2} \right)^2 - \frac{1}{2} \frac{g_2''}{g_2}, \quad (\text{B.3.35})$$

ce qui nous permet de constater que seules les dérivées de g_2 apparaissent au numérateur de ces deux termes. Or, g_2 est le seul terme qui fait apparaître la fréquence propre ω^2 au numérateur (elle intervient au dénominateur des autres termes, via \mathcal{L}^2), celle-ci disparaît dans ses dérivées. En considérant que la seule occurrence de la fréquence est donnée par l'équation (B.3.33), on pourra alors réécrire la relation (B.3.32) sous la forme d'une équation de Schrödinger stationnaire :

$$\Xi'' + (\omega^2 - V)\Xi = 0, \quad (\text{B.3.36})$$

où $V(\tau)$ désigne le *potentiel acoustique* perçu par l'onde. On supposera que ce potentiel n'est pas fonction de la fréquence propre ω^2 , bien qu'il en soit faiblement dépendant (via les interventions de ω^2 au dénominateur). Si le potentiel acoustique possède une expression non triviale, on peut

chercher à en donner une approximation dans le régime asymptotique ($\mathcal{L} \rightarrow 0$). Sous cette limite :

$$g_1 \sim \frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} \quad \text{et} \quad g_2 \sim \frac{dg}{dr} - \frac{2g}{r} - \omega^2. \quad (\text{B.3.37})$$

En remarquant que :

$$\begin{aligned} \frac{\sqrt{g_2}}{\lambda} \left(\frac{\lambda}{\sqrt{g_2}} \right)'' &= \frac{1}{4} (\ln' \lambda^2 - \ln' g_2)^2 + \frac{1}{2} (\ln' \lambda^2 - \ln' g_2)' \\ &= \frac{1}{4} \left[\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} - \frac{2c}{r} - \ln' \left(\frac{dg}{dr} - \frac{2g}{r} - \omega^2 \right) \right]^2 \\ &\quad + \frac{1}{2} \left[\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} - \frac{2c}{r} - \ln' \left(\frac{dg}{dr} - \frac{2g}{r} - \omega^2 \right) \right]', \end{aligned} \quad (\text{B.3.38})$$

on obtient l'approximation qui suit pour le potentiel acoustique :

$$\begin{aligned} V(\tau) \underset{\mathcal{L} \rightarrow 0}{\sim} & \left(\frac{\Gamma_1}{\gamma} - 1 \right) \frac{2g}{r} + \frac{dg}{dr} - \frac{2c}{r} \frac{dc}{dr} - \frac{d}{d\tau} \left(\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} \right) \\ & + \left(\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} \right) \left(\frac{dg}{dr} - \frac{2g}{r} - \omega^2 \right)^{-1} \frac{d}{d\tau} \left(\frac{dg}{dr} - \frac{2g}{r} \right) \\ & + \frac{1}{4} \left[\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} - \frac{2c}{r} - \left(\frac{dg}{dr} - \frac{2g}{r} - \omega^2 \right)^{-1} \frac{d}{d\tau} \left(\frac{dg}{dr} - \frac{2g}{r} \right) \right]^2 \\ & + \frac{1}{2} \frac{d}{d\tau} \left[\frac{\Gamma_1 g}{\gamma c} - \frac{dc}{dr} - \frac{2c}{r} - \left(\frac{dg}{dr} - \frac{2g}{r} - \omega^2 \right)^{-1} \frac{d}{d\tau} \left(\frac{dg}{dr} - \frac{2g}{r} \right) \right], \end{aligned} \quad (\text{B.3.39})$$

laquelle est réemployée dans le paragraphe 2.4.10.

Annexe C

Compléments probabilistes

Contenu de l'annexe

C.1	Contexte et notations	189
C.2	L'inférence bayésienne	190
C.3	Les processus gaussiens	199

C.1 Contexte et notations

Cette annexe est à considérer en complément de l'approche probabiliste mise en œuvre dans l'article [HOUDAYER, REESE et GOUPIL, 2022](#), inclus dans la partie 5.1 et qui se penche en particulier sur les méthodes d'inférence bayésienne. Le cœur du formalisme cherche à mettre en relation les notions théorique de *modèle* et plus concrète de *donnée* et l'on s'y attellera à l'aide d'un formalisme suffisamment général (en grande partie emprunté à [RASMUSSEN et WILLIAMS, 2005](#)). Ainsi, on notera y une observable quelconque (température d'une pièce, luminosité d'un astre, écart d'une fréquence d'oscillation, ...) observée au point x (potentiellement multidimensionnel : position dans la pièce (3D), moment de l'observation (1D), fréquence d'oscillation concernée (1D), ...) lesquels peuvent être respectivement agrégés en un vecteur \mathbf{y} et une matrice X dans le cas d'observations multiples. On notera alors de façon compacte $\mathcal{D} \equiv \{\mathbf{y}, X\}$ le jeu de données. On peut définir en parallèle les objets auxquels on confie la tâche de faire le lien entre variables et observables, les modèles \mathcal{M}_i , que l'on indexera symboliquement par i et qui seront munis de leurs *paramètres* et *hyperparamètres*. Si les notions de modèle et de paramètre sont assez communes, celle d'hyperparamètre l'est moins et l'on l'illustrera par un simple exemple. Imaginons un jeu de données que l'on souhaite modéliser à l'aide d'un modèle polynomial :

$$\mathcal{M}(x; \mathbf{a}, N) = \sum_{i=0}^N a_i x^i = \mathbf{a}^T \boldsymbol{\phi}_N(x) \quad (\text{C.1.1})$$

où $\boldsymbol{\phi}_N$ désigne la fonction vectorielle qui à tout x renvoie un vecteur contenant ses puissances successives jusqu'à x^N , $[1 \ x \ x^2 \ \dots \ x^N]^T$ avec \cdot^T l'opération de transposition. Dans une telle situation, il est courant de chercher à ajuster "au mieux" le vecteur de paramètres \mathbf{a} pour une valeur de N donnée : le cas $N = 1$ correspond à l'ajustement linéaire, $N = 2$ à l'ajustement quadratique, etc... . Il est en revanche moins courant de chercher à ajuster N en même temps que \mathbf{a} , ce qui transformerait un problème d'optimisation linéaire en problème hautement non-linéaire¹. Dans cet exemple, N peut être considéré comme un hyperparamètre dans le sens où il agit à niveau différent des paramètres au sein du modèle : il définit des sous-modèles (linéaire, quadratique, ...) et est ainsi ajusté dans un second temps². Sur un échelon encore supérieur,

1. Même s'il est évidemment souhaitable de s'interroger préalablement quant à la valeur à adopter pour N !

2. On reviendra sur la notion d'hyperparamètre lors de l'étude des processus gaussiens, cadre où elle acquiert un sens légèrement différent.

il serait possible de définir des modèles qui diffèrent du modèle polynomial que ce soit en considérant une autre fonction pour $\Phi(x)$ ³ ou bien une expression intégralement différente pour l'équation (C.1.1) (modèle exponentiel, modèle hyperbolique, processus gaussien, ...).

Le formalisme probabiliste qui s'appuie sur ces notions, l'inférence bayésienne, permet d'aborder de nombreux sujets : le choix de paramètres et d'hyperparamètres pour un modèle donné, la comparaison de modèles et de paramètres entre différents modèles, l'anticipation de données futures disposant de mesures passées, le contenu informationnel des modèles et données (conjointement avec la théorie de l'information de SHANNON, 1948), la maximisation d'une utilité espérée (qui recoupe la théorie de la décision, cf. BERGER, 1985) et beaucoup d'autres. Il s'agira ici d'en étudier quelques propriétés avant de s'attarder sur le formalisme propre aux processus gaussiens.

C.2 L'inférence bayésienne

C.2.1 Loi de probabilité, loi conditionnelle, loi marginale et indépendance

Commençons par considérer une loi de probabilité p , et qui vérifie donc par définition :

$$\int_{y \in \mathcal{A}} p(y) dy = 1 \quad (\text{C.2.1})$$

$$\int_{y \in \cup_i \mathcal{A}_i} p(y) dy = \sum_i \int_{y \in \mathcal{A}_i} p(y) dy, \quad (\text{C.2.2})$$

si les évènements composant les ensembles \mathcal{A}_i sont disjoints, où $p(y)dy$ désigne de façon succincte la probabilité que la variable aléatoire Y se trouve dans l'intervalle $[y, y + dy]$ ⁴ et où \mathcal{A} contient tous les évènements de l'univers Ω . Les propriétés sus-citées se généralisent sans problème à la définition d'une loi p définie sur un espace de dimension supérieur (cf. paragraphe 2.1.2 par exemple) afin de manipuler des quantités telles que $p(\mathbf{y})$. En dénotant $p(\mathbf{y}, \mathbf{z})$ la loi jointe associée à l'évènement $(Y = \mathbf{y}) \cap (Z = \mathbf{z})$ (avec Y et Z deux variables aléatoires), on définira l'objet central du formalisme, la probabilité conditionnelle $p(\mathbf{y}|\mathbf{z})$ par :

$$p(\mathbf{y}, \mathbf{z}) = p(\mathbf{y}|\mathbf{z})p(\mathbf{z}). \quad (\text{C.2.3})$$

Notons que puisque $p(\mathbf{z})$ est obtenue de la loi jointe par marginalisation sur toutes les possibilités vis-à-vis de \mathbf{y} :⁵

$$p(\mathbf{y}|\mathbf{z}) = \frac{p(\mathbf{y}, \mathbf{z})}{\int_{\mathbf{y} \in \mathcal{A}_y} p(\mathbf{y}, \mathbf{z}) d\mathbf{y}'} \quad (\text{C.2.4})$$

la loi conditionnelle $p(\mathbf{y}|\mathbf{z})$ définie bien une loi de probabilité. De plus, la symétrie de la relation (C.2.3) nous fournit la loi de Bayes :

$$p(\mathbf{y}|\mathbf{z}) = \frac{p(\mathbf{z}|\mathbf{y})p(\mathbf{y})}{p(\mathbf{z})}, \quad (\text{C.2.5})$$

qui lie les deux lois conditionnelles. On dira alors que les deux variables aléatoires sont indépendantes si elles vérifient :

$$p(\mathbf{y}|\mathbf{z}) = p(\mathbf{y}), \quad (\text{C.2.6})$$

ou autrement dit, si la valeur de \mathbf{z} n'influe aucunement la loi que suit \mathbf{y} . D'après la loi de Bayes, on voit immédiatement que cette propriété est mutuelle et que $p(\mathbf{z}|\mathbf{y}) = p(\mathbf{z})$ dans un tel cas.

3. Auquel cas le problème d'optimisation reste linéaire.

4. On s'intéressera le plus souvent à des variables aléatoires continues, même si les définitions données plus haut s'adaptent très bien à des variables discrètes en remplaçant les intégrales par des sommes.

5. La notation $d\mathbf{y}$ nous sert ici d'écriture compacte pour le produit $\prod_i dy_i$.

C.2.2 Formalisme de l'inférence

Sous le prisme de l'inférence bayésienne, les propriétés et définitions évoquées plus haut peuvent être appliquées aux notions introduites dans le paragraphe C.1. Par exemple, on s'intéressera sur la plausibilité⁶ de telle ou telle valeur θ adoptée par un jeu de paramètres dans un modèle \mathcal{M} et pour un jeu de données $\mathcal{D} = \{\mathbf{y}, X\}$. On nommera la loi de probabilité correspondante $p(\theta|\mathcal{D}, \mathcal{M})$ et la formule de Bayes nous donne la relation⁷ :

$$p(\theta|\mathcal{D}, \mathcal{M}) = \frac{p(\mathbf{y}|X, \theta, \mathcal{M})p(\theta|\mathcal{M})}{p(\mathbf{y}|X, \mathcal{M})}. \quad (\text{C.2.7})$$

Dans cette équation, on réfère au terme de gauche, $p(\theta|\mathcal{D}, \mathcal{M})$ par la *plausibilité à posteriori* du jeu de paramètres θ après la découverte des données \mathcal{D} , par opposition à la *plausibilité à priori*, $p(\theta|\mathcal{M})$ de ce jeu de paramètres avant la découverte de \mathcal{D} . On notera que ce terme devrait formellement s'écrire $p(\theta|X, \mathcal{M})$ en appliquant la formule de Bayes mais la plausibilité à priori des paramètres est évidemment indépendante du choix des seuls points de mesures (avant que ces mesures ne soient effectuées). Le terme $p(\mathbf{y}|X, \theta, \mathcal{M})$ désigne la *vraisemblance* des mesures \mathbf{y} aux points X selon le modèle \mathcal{M} paramétré par θ et le dénominateur la *vraisemblance marginale* dans le sens où elle résulte de la marginalisation sur toutes les possibilités de paramètres dans la vraisemblance :

$$p(\mathbf{y}|X, \mathcal{M}) = \int_{\theta \in \Theta_{\mathcal{M}}} p(\mathbf{y}|X, \theta, \mathcal{M})p(\theta|\mathcal{M}) d\theta \quad (\text{C.2.8})$$

avec $\Theta_{\mathcal{M}}$ l'espace des paramètres de \mathcal{M} dans lequel évolue θ . Terme le plus complexe à calculer en pratique, son indépendance vis-à-vis d'un jeu de paramètres particulier conduit souvent à la considérer comme une simple constante de normalisation pour la loi à posteriori exprimée sur θ :

$$p(\theta|\mathcal{D}, \mathcal{M}) \propto p(\mathbf{y}|X, \theta, \mathcal{M})p(\theta|\mathcal{M}). \quad (\text{C.2.9})$$

Cette quantité recèle cependant de précieuses informations (comme on le verra par la suite) ce qui en fait en particulier un outil central dans l'étude des processus gaussiens.

C.2.3 La plausibilité à priori

Au vu des équations (C.2.7) et (C.2.8), les quantités qui doivent être explicitées dans le but d'évaluer le jeu de paramètres le plus plausible sont les plausibilité des paramètres à priori $p(\theta|\mathcal{M})$ et vraisemblance des données $p(\mathbf{y}|X, \theta, \mathcal{M})$. Commençons par nous attarder sur cette notion de plausibilité à priori. La question à se poser est : en l'absence des données \mathcal{D} , quelles sont les valeurs les plus plausibles de θ pour paramétrer \mathcal{M} ? Il y a alors autant de façons d'y répondre que de contextes distincts. Si l'on a déjà été exposé à des données \mathcal{D}_0 avant de disposer du jeu \mathcal{D} , cette plausibilité à priori peut très bien être assimilée à la plausibilité à posteriori connaissant \mathcal{D}_0 , auquel cas :

$$p(\theta|\mathcal{M}) = p(\theta|\mathcal{D}_0, \mathcal{M}). \quad (\text{C.2.10})$$

Dans le cas contraire, la loi à priori peut être choisie de façon à refléter des impossibilités mathématiques sur les paramètres. Par exemple, si $\theta_k \in \theta$ désigne une échelle de temps, celle-ci devra vérifier $\theta_k > 0$ et l'on pourra assigner une probabilité nulle aux zones exclues : $p(\theta_k \leq 0) = 0$. Il convient néanmoins d'être prudent avant d'imposer une probabilité nulle à une région de l'espace des paramètres. En effet, de par l'équation (C.2.7), cette "région interdite" se transmettra

6. J'emploie ici le terme de "plausibilité" plutôt que de "probabilité" pour souligner l'aspect subjectif de cette métrique (i.e., dans un monde décrit par le modèle \mathcal{M} et disposant des données \mathcal{D}) qui caractérise davantage un degré de connaissance qu'une réelle probabilité d'occurrence d'une variable aléatoire (on peut alors parler de *probabilité épistémique*), voir la discussion de FINETTI, 1992.

7. Dans toutes les relations qui suivront, on notera les variables "connaissant le jeu de données \mathcal{D} " avec l'écriture $|\mathcal{D}$ dès lors que les points de mesure X et les mesures en ces points \mathbf{y} sont connues. Il s'agit donc d'une notation un peu plus compacte pour l'écriture $|X, \mathbf{y}$, qui lui est totalement équivalente.

automatiquement à la plausibilité à posteriori et ce quelque soit les données \mathcal{D} observées. Il est donc essentiel de ne pas définir comme impossible des régions qui ne seraient seulement que très peu plausibles à priori, cet avis pouvant être remis en cause par l'acquisition de données inattendues.

Un autre moyen (assez courant) de définir des lois à priori consiste à choisir d'abord une loi paramétrée pour le jeu de paramètres, par exemple une loi gaussienne :

$$p(\boldsymbol{\theta}|\boldsymbol{\mu}, \Sigma, \mathcal{M}) = \mathcal{G}(\boldsymbol{\theta}, \boldsymbol{\mu}, \Sigma) = \frac{1}{(2\pi)^{N/2}|\Sigma|^{1/2}} e^{-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\mu})^\top \Sigma^{-1}(\boldsymbol{\theta}-\boldsymbol{\mu})} \quad (\text{C.2.11})$$

avec N la longueur de $\boldsymbol{\theta}$, $\boldsymbol{\mu}$ la moyenne de la loi gaussienne et Σ sa matrice de covariance (et $|\Sigma|$ son déterminant). Le choix de la loi à priori se résume alors à la spécification des $N(N+3)/2$ composantes des vecteur $\boldsymbol{\mu}$ et matrice (symétrique) Σ , laquelle pourra se faire à l'aide de raisonnements physiques. Les composantes de $\boldsymbol{\mu}$ peuvent être vues comme des "first guesses" et celles de Σ comme un reflet de notre incertitude quant à ces estimations.

Si aucune information n'est connue quant à $\boldsymbol{\theta}$, une dernière possibilité reste le choix d'une *loi non-informative* sur les paramètres. Une option populaire est la loi à priori de Jeffreys (JEFFREYS, 1946) qui se déduit d'un groupe de transformations laissant le problème invariant. Cette loi est alors définie de façon à satisfaire l'invariance sous l'action de ce groupe. Les problèmes invariants par translation laissent alors apparaître une loi uniforme $p(\theta_k) = \text{cste}$, alors que l'invariance par changement d'échelle induit plutôt des lois hyperboliques $p(\theta_k) \propto 1/\theta_k$. Une seconde option élégante réside dans la maximisation de l'entropie contenue dans la loi à priori. Définie par :

$$\mathcal{H}[p(\cdot|\mathcal{M})] \equiv \mathbb{E}[-\log p(\boldsymbol{\theta}|\mathcal{M})]_{p(\boldsymbol{\theta}|\mathcal{M})} = - \int_{\boldsymbol{\theta} \in \Theta_{\mathcal{M}}} p(\boldsymbol{\theta}|\mathcal{M}) \log p(\boldsymbol{\theta}|\mathcal{M}) d\boldsymbol{\theta} \quad (\text{C.2.12})$$

avec $\mathbb{E}[f(y)]_{p(y)}$ l'espérance de la quantité $f(y)$ suivant la distribution $p(y)$, l'entropie de Shannon peut être pensée comme une fonctionnelle assignant un nombre à une loi de probabilité. Ce nombre reflète la quantité d'information que contient la loi et est d'autant plus faible que la l'information contenue est grande. Une façon d'interpréter la maximisation de l'entropie est donc la minimisation de l'information injectée à priori. L'objectif implicite est de laisser un maximum de poids à l'information apportée par les données \mathcal{D} dans la plausibilité à posteriori $p(\boldsymbol{\theta}|\mathcal{D}, \mathcal{M})$ par rapport à "ce que l'on aurait pu introduire subjectivement à priori". Enfin, cette approche se couple très bien avec les méthodes informatives évoquées plus-haut. Il est ainsi courant de commencer par injecter l'information contenue dans les contraintes physiques et impossibilités mathématiques avant de chercher la loi à priori la moins informative vérifiant ces contraintes par maximisation de l'entropie.

C.2.4 La vraisemblance

Dans l'équation (C.2.7), on voit que toutes les probabilités sont évaluées dans le cadre défini par le modèle \mathcal{M} . Si cela tombe sous le sens lorsque l'on pense aux plausibilités (puisque $\boldsymbol{\theta}$ contient les paramètres du modèle), il est néanmoins utile de l'indiquer explicitement si l'on souhaite interpréter de façon adéquate les vraisemblances. En particulier, une exigence qui transparaît de la formule de Bayes est la capacité du modèle \mathcal{M} à prédire des données. Le calcul de cette formule exige en effet l'expression de $p(\mathbf{y}|X, \boldsymbol{\theta}, \mathcal{M})$ et donc la détermination d'une loi de probabilité pour la distribution des mesures \mathbf{y} aux points X selon le modèle \mathcal{M} paramétré par $\boldsymbol{\theta}$. Cette remarque nous permet de définir plus clairement ce que l'on considère être ou non un modèle dans le formalisme bayésien. Par exemple, un modèle stellaire en tant que tel ne prédit rien quant aux données photométriques et spectroscopiques que nous fournit notre instrument. C'est seulement muni de nombreux modèles annexes (modèle atmosphérique de l'étoile, modèle de notre propre atmosphère, modèle de l'instrument lui-même, ...) et hypothèses (quant à la position de l'étoile ou au milieu interstellaire par exemple) que ce modèle stellaire devient prédictif quand à la lumière qui nous parvient, et c'est donc la plausibilité de cette machinerie complexe

que la formule de Bayes va évaluer. Ainsi, une plausibilité faible du modèle dans son ensemble ne permettra pas d'affirmer que le modèle stellaire est à remettre en cause plus qu'une autre partie de cet ensemble, et le formalisme nous pousse à répartir cette incertitude sur chacun des constituants du modèle.

C.2.5 La hiérarchie bayésienne

On s'est penché jusqu'ici au cadre à définir avant de chercher à quantifier la plausibilité d'une valeur θ des paramètres du modèle \mathcal{M} ayant à disposition les données \mathcal{D} . Le formalisme introduit permet cependant d'aborder de nombreuses autres questions, à commencer par le choix du modèle lui-même. Plaçons-nous dans un cadre où l'on compte un ensemble de modèles $\{\mathcal{M}_i\}_{i \in \mathcal{I}}$ où l'on cherche à sélectionner le modèle le plus à même d'expliquer les données \mathcal{D} . La formule de Bayes nous permet d'évaluer les plausibilités à posteriori de chaque modèle :

$$p(\mathcal{M}_i|\mathcal{D}) = \frac{p(\mathbf{y}|X, \mathcal{M}_i)p(\mathcal{M}_i)}{p(\mathcal{D})}, \quad (\text{C.2.13})$$

ce que l'on peut réécrire en explicitant la vraisemblance marginale $p(\mathcal{D})$ comme :

$$p(\mathcal{M}_i|\mathcal{D}) = \frac{p(\mathbf{y}|X, \mathcal{M}_i)p(\mathcal{M}_i)}{\sum_{i \in \mathcal{I}} p(\mathbf{y}|X, \mathcal{M}_i)p(\mathcal{M}_i)}. \quad (\text{C.2.14})$$

On remarque qu'il est alors nécessaire de spécifier comment se répartit notre confiance sur l'ensemble des modèles avant d'être exposé aux données \mathcal{D} , ce que symbolise la distribution à priori $p(\mathcal{M}_i)$. En plus de ce facteur, il est intéressant de noter la réapparition d'un terme présent à l'échelon inférieur (lors du choix des paramètres pour un modèle donné) au travers de la vraisemblance des données $p(\mathbf{y}|X, \mathcal{M}_i)$ (qui était alors la vraisemblance marginale, cf. relation C.2.8). Cette hiérarchie permet de faire apparaître clairement l'intérêt de calculer la vraisemblance marginale par l'intégrale (C.2.8) plutôt que de l'interpréter comme une simple constante de normalisation. En effet, regardée sous l'angle d'une fonction de \mathcal{M}_i , cette intégrale nous renseigne sur la capacité du modèle à expliquer les données \mathcal{D} . Dans le cas particulier où l'on ne juge aucun modèle plus plausible qu'un autre à priori (i.e. si $p(\mathcal{M}_i) = \text{cste}$), le modèle le plus plausible à posteriori, \mathcal{M}_i^* , se déduit même directement de la maximisation de cette intégrale :

$$\mathcal{M}_i^* = \arg \max_{\mathcal{M}_i} p(\mathcal{M}_i|\mathcal{D}) = \arg \max_{\mathcal{M}_i} \int_{\theta \in \Theta_{\mathcal{M}_i}} p(\mathbf{y}|X, \theta, \mathcal{M}_i)p(\theta|\mathcal{M}_i) d\theta. \quad (\text{C.2.15})$$

La hiérarchie bayésienne peut assembler davantage d'échelons. En reprenant les notations de RASMUSSEN et WILLIAMS, 2005, on peut supposer disposer d'un ensemble de modèles $\{\mathcal{M}_i\}_{i \in \mathcal{I}}$ dont les hyperparamètres θ définissent des sous-modèles, eux-mêmes paramétrés par les poids \mathbf{w} . La hiérarchie d'équations prendrait alors la forme suivante :

$$p(\mathbf{w}|\mathcal{D}, \theta, \mathcal{M}_i) \propto_{\mathbf{w}} p(\mathbf{y}|X, \theta, \mathbf{w}, \mathcal{M}_i)p(\mathbf{w}|\theta, \mathcal{M}_i) \quad (\text{C.2.16})$$

$$p(\mathbf{y}|X, \theta, \mathcal{M}_i) = \int_{\mathbf{w} \in \mathbb{W}_{\mathcal{M}_i(\theta)}} p(\mathbf{y}|X, \mathbf{w}, \theta, \mathcal{M}_i)p(\mathbf{w}|\theta, \mathcal{M}_i) d\mathbf{w} \quad (\text{C.2.17})$$

$$p(\theta|\mathcal{D}, \mathcal{M}_i) \propto_{\theta} p(\mathbf{y}|X, \theta, \mathcal{M}_i)p(\theta|\mathcal{M}_i) \quad (\text{C.2.18})$$

$$p(\mathbf{y}|X, \mathcal{M}_i) = \int_{\theta \in \Theta_{\mathcal{M}_i}} p(\mathbf{y}|X, \theta, \mathcal{M}_i)p(\theta|\mathcal{M}_i) d\theta \quad (\text{C.2.19})$$

$$p(\mathcal{M}_i|\mathcal{D}) \propto_{\mathcal{M}_i} p(\mathbf{y}|X, \mathcal{M}_i)p(\mathcal{M}_i) \quad (\text{C.2.20})$$

où $\mathbb{W}_{\mathcal{M}_i(\theta)}$ désigne l'espace des paramètres du sous-modèle $\mathcal{M}_i(\theta)$. Les différents échelons et leur dépendance vis-à-vis de ce que l'on considère être variable ou non soulignent alors la nécessité de définir clairement l'objectif que l'on se fixe avant de réaliser l'inférence pour vérifier que

l'on ne néglige pas un aspect du problème (ou à l'inverse que l'on n'effectue pas de calcul inutile). Par exemple, quelqu'un qui souhaiterait seulement obtenir le jeu de paramètres \mathbf{w} le plus plausible d'un sous-modèle $\mathcal{M}_i(\boldsymbol{\theta})$ particulier pourrait se contenter de spécifier les lois des variables $\mathbf{y}|X, \boldsymbol{\theta}, \mathbf{w}, \mathcal{M}_i$ et $\mathbf{w}|\boldsymbol{\theta}, \mathcal{M}_i$ et ne calculer que l'équation (C.2.16). À l'inverse, quelqu'un qui voudrait déterminer le modèle \mathcal{M}_i le plus crédible au vu des données \mathcal{D} se devrait de fournir, plus de la vraisemblance $p(\mathbf{y}|X, \boldsymbol{\theta}, \mathbf{w}, \mathcal{M}_i)$, les trois lois à priori : $p(\mathbf{w}|\boldsymbol{\theta}, \mathcal{M}_i)$, $p(\boldsymbol{\theta}|\mathcal{M}_i)$ et $p(\mathcal{M}_i)$. Il serait également nécessaire de passer par le calcul complet des deux intégrales (C.2.17) et (C.2.19) avant de pouvoir estimer $p(\mathcal{M}_i|\mathcal{D})$ pour chaque modèle via la dernière équation (C.2.20).

C.2.6 La forêt de modèles

Le formalisme offre également une réponse élégante au problème de la détermination d'une grandeur qui serait commune à plusieurs modèles. Si l'on peut intuitivement penser que la solution optimale serait de sélectionner le modèle le plus performant (à l'aide de l'équation C.2.20) avant de chercher la valeur la plus plausible des paramètres avec la formule (C.2.18), la réponse apportée par l'inférence bayésienne est tout autre. Imaginons cette fois-ci que l'on possède un ensemble de modèles $\{\mathcal{M}_i\}_{i \in \mathcal{I}}$, chacun paramétré par un vecteur de paramètres qui lui est propre, $\boldsymbol{\theta}_i$. Supposons maintenant qu'il existe une partie commune, $\boldsymbol{\theta}_\cap$, à tous ces vecteurs :

$$\boldsymbol{\theta}_\cap = \bigcap_{i \in \mathcal{I}} \boldsymbol{\theta}_i \quad (\text{C.2.21})$$

ainsi donc qu'une partie propre à chaque modèle \mathcal{M}_i , $\bar{\boldsymbol{\theta}}_i$, telle que :

$$\boldsymbol{\theta}_i = (\boldsymbol{\theta}_\cap) \cup (\bar{\boldsymbol{\theta}}_i). \quad (\text{C.2.22})$$

On aimerait maintenant savoir ce qu'il est possible de dire à propos des paramètres communs à tous nos modèles après l'observation des données \mathcal{D} . En supposant que l'ensemble d'indices \mathcal{I} permet bien de dénombrer tous les modèles possibles⁸, il s'agit alors d'évaluer la loi $p(\boldsymbol{\theta}_\cap|\mathcal{D})$. La formule des probabilités composées permet de la décomposer comme :

$$p(\boldsymbol{\theta}_\cap|\mathcal{D}) = \sum_{i \in \mathcal{I}} p(\boldsymbol{\theta}_\cap|\mathcal{M}_i, \mathcal{D})p(\mathcal{M}_i|\mathcal{D}). \quad (\text{C.2.23})$$

Le premier terme résulte de la marginalisation de $p(\boldsymbol{\theta}_i|\mathcal{M}_i, \mathcal{D})$ sur $\bar{\boldsymbol{\theta}}_i$ et la formule de Bayes permet d'écrire :

$$p(\boldsymbol{\theta}_\cap|\mathcal{M}_i, \mathcal{D}) = \int_{\bar{\boldsymbol{\theta}}_i \in \bar{\Theta}_{\mathcal{M}_i}} p(\boldsymbol{\theta}_i|\mathcal{M}_i, \mathcal{D}) d\bar{\boldsymbol{\theta}}_i = \frac{1}{p(\mathbf{y}|X, \mathcal{M}_i)} \int_{\bar{\boldsymbol{\theta}}_i \in \bar{\Theta}_{\mathcal{M}_i}} p(\mathbf{y}|X, \boldsymbol{\theta}_i, \mathcal{M}_i)p(\boldsymbol{\theta}_i|\mathcal{M}_i) d\bar{\boldsymbol{\theta}}_i. \quad (\text{C.2.24})$$

où $\bar{\Theta}_{\mathcal{M}_i}$ désigne l'espace regroupant les paramètres propres à \mathcal{M}_i . Le facteur $p(\mathcal{M}_i|\mathcal{D})$ peut être quant à lui directement exprimé avec la relation de proportionnalité (C.2.20) pour finalement donner :

$$p(\boldsymbol{\theta}_\cap|\mathcal{D}) \propto_{\boldsymbol{\theta}_\cap} \sum_{i \in \mathcal{I}} p(\mathcal{M}_i) \int_{\bar{\boldsymbol{\theta}}_i \in \bar{\Theta}_{\mathcal{M}_i}} p(\mathbf{y}|X, \boldsymbol{\theta}_i, \mathcal{M}_i)p(\boldsymbol{\theta}_i|\mathcal{M}_i) d\bar{\boldsymbol{\theta}}_i. \quad (\text{C.2.25})$$

L'enseignement à tirer de cette formule est intéressant. Loin de ne se reposer que sur le modèle le plus performant, l'équation (C.2.25) nous pousse à utiliser les propositions de l'ensemble des modèles, une *forêt de modèles*⁹, pour déterminer les paramètres communs, $\boldsymbol{\theta}_\cap$, les plus plausibles. On voit également que "l'opinion" de chaque modèle sur la décision finale est pondérée par sa crédibilité à priori $p(\mathcal{M}_i)$.

8. Et donc que $p(\mathcal{M}_i)$ désigne bien une loi de probabilité.

9. Le terme de forêt de modèles est ici inspirée de la dénomination *random forest* introduite par Ho, 1995 dans le contexte du machine learning. L'idée était de combiner l'avis des plusieurs modèles, en l'occurrence des arbres de classification, lesquels forment ensemble la forêt.

Prenons un exemple concret pour illustrer cette façon de procéder. On supposera que l'on cherche à estimer la température effective d'une étoile, T_{eff} , en s'appuyant sur des données nous renseignant sur la lumière qui nous parvient, \mathcal{D} . On dispose pour cela de deux modèles : un modèle photométrique, \mathcal{P} , s'appuyant sur l'intensité lumineuse reçue et un modèle spectroscopique, \mathcal{S} , qui repose le spectre lumineux obtenu. Ces deux modèles ont pour paramètre T_{eff} (on supposera même qu'il s'agit de leur seul paramètre pour simplifier) mais n'ont à priori aucune raison de coïncider quant à leur loi de distribution à posteriori $p(T_{\text{eff}}|\mathcal{P}, \mathcal{D})$ et $p(T_{\text{eff}}|\mathcal{S}, \mathcal{D})$. La formule (C.2.25) nous propose alors une estimation de la loi $p(T_{\text{eff}}|\mathcal{D})$, s'affranchissant du point de vue particulier d'un modèle mais prenant leurs deux estimations en considération :

$$p(T_{\text{eff}}|\mathcal{D}) \propto_{T_{\text{eff}}} p(\mathcal{P})p(\mathbf{y}|\mathcal{X}, T_{\text{eff}}, \mathcal{P})p(T_{\text{eff}}|\mathcal{P}) + p(\mathcal{S})p(\mathbf{y}|\mathcal{X}, T_{\text{eff}}, \mathcal{S})p(T_{\text{eff}}|\mathcal{S}). \quad (\text{C.2.26})$$

Une illustration de cet exemple simple est proposée en FIGURE C.1.

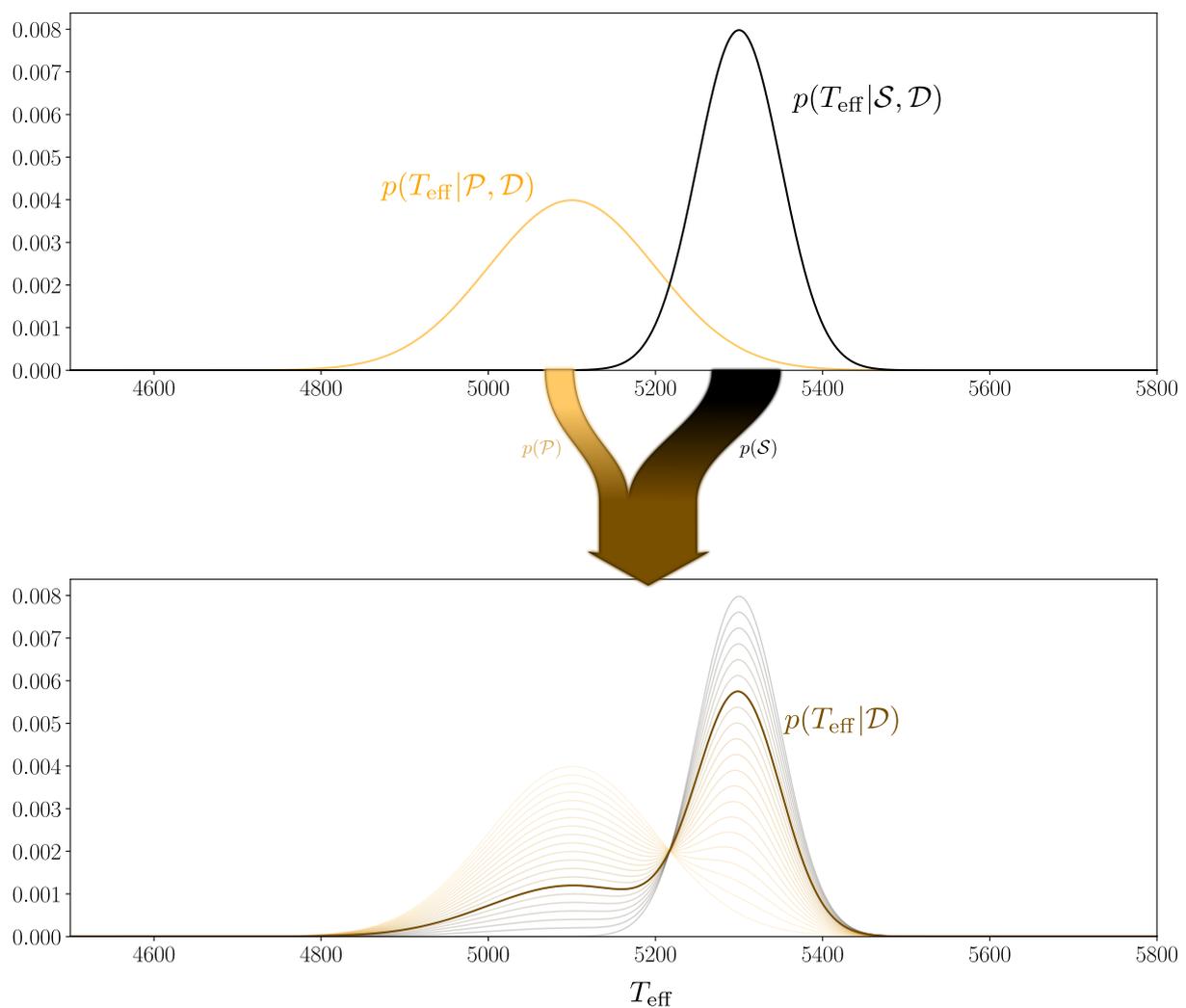


FIGURE C.1 – Illustration schématique de la distribution sur la température effective obtenue à partir des modèles photométrique (\mathcal{P}) et spectroscopique (\mathcal{S}). Le panneau du haut fournit les lois à posteriori de chacun des modèles sur T_{eff} tandis que celui du bas indique la loi commune $p(T_{\text{eff}}|\mathcal{D})$ obtenue pour une confiance à priori plus élevée dans le modèle spectroscopique $p(\mathcal{S}) = 0.7$ (et donc $p(\mathcal{P}) = 0.3$).

De façon un peu détournée, les forêts de modèles nous poussent aussi à nous interroger sur la signification qu'ont les paramètres en dehors du modèle qui les emploie. En revenant à l'exemple

donné plus haut, on pourrait se demander si cela a réellement un sens de parler de température effective en dehors d'un modèle photométrique ou spectroscopique, ou autrement formulé : la température effective a-t-elle *exactement* la même signification dans les modèles photométriques et spectroscopiques ? Il est tentant de répondre par l'affirmative à cette question puisque, dans les deux cas, la quantité T_{eff} se veut correspondre à une grandeur réelle et bien définie, intrinsèque à l'étoile, qui caractérise sa couleur.

Je me permettrai d'élargir cette réflexion en apportant un autre exemple, bien connu des physiciens stellaires et qui soulève à tous les coups des remarques lorsqu'il apparaît sur une slide, celui de la masse du soleil. Il n'est curieusement pas rare, lorsque l'on essaie d'ajuster la masse du soleil à l'aide d'un modèle stellaire et utilisant des observables spectroscopiques ou sismiques, de ne pas exactement trouver $1M_{\odot}$ mais des valeurs légèrement différentes comme $0.99M_{\odot}$ ou $1.01M_{\odot}$. Entendons-nous bien, la masse solaire est calibrée pour représenter la masse nécessaire à l'explication des mouvements des astres qui nous entourent et le soleil devrait bien peser $1M_{\odot}$. On se trouve de nouveau dans le cas d'une comparaison de terminologie commune (la masse du soleil) à deux modèles (le modèle stellaire et le modèle d'attraction gravitationnelle) et à l'instar de la température effective on peut se poser la question : ces deux modèles parlent-ils vraiment de la même chose ? Et sinon, quel sens donner au concept de masse stellaire en dehors d'un modèle stellaire ?

Ces considérations sont abstraites et touchent aux liens entre réalité et langage employé à sa description, si bien qu'il est difficile de fournir une réponse claire aux questions soulevées. Dans ce cas précis de la masse du soleil, on peut penser que le modèle d'attraction gravitationnel s'approche davantage de la description d'une réalité physique dans le sens où elle repose sur une mesure plus directe de ce que l'on entend le plus communément par masse (c'est à dire une résistance à la force). Par opposition, la masse apparaissant dans le modèle stellaire s'interprète plus facilement comme un "réglage", un curseur à modifier pour faire correspondre observables issues du modèle et données, la mesure de cette grandeur par la lumière qui nous parvient étant bien moins directe que dans l'autre cas.

De manière générale, on peut noter que le concept de "distance" à la grandeur observée, que l'on peut lier à celui (mieux défini) de *complexité du modèle*, semble être un critère important dans la signification que l'on prête au paramètre correspondant. On s'entendra plus facilement, par exemple, sur le sens physique à attribuer à la luminosité d'une étoile (et donc sur son sens en dehors de tout modèle stellaire) qu'à celui de longueur de mélange qui, bien que pourvu de sens dans le modèle MLT (cf. paragraphe 2.3.8), apparaît davantage comme un réglage que comme la description d'une propriété intrinsèque à l'étoile.

De façon intéressante, la forêt de modèles fait ressortir un autre point important lorsqu'il s'agit de juger du sens d'un paramètre commun. En plus de la complexité des modèles employés, la formule (C.2.25) insiste sur la nécessité pour l'ensemble de modèles $\{\mathcal{M}_i\}_{i \in \mathcal{I}}$ de pouvoir utiliser des données communes, \mathcal{D} , et en particulier d'être prédictif quant à ces données (via le terme $p(\mathbf{y}|\mathbf{X}, \theta_i, \mathcal{M}_i)$). Le sens à attribuer à un paramètre serait alors relatif à l'ensemble de modèles considéré : la masse stellaire aurait bien un sens dans un cadre où l'on ne pense qu'à des modèles stellaires (puisque ces derniers sont prédictifs vis-à-vis des observables spectroscopiques et sismiques) mais le perdrait dès lors que l'on envisage la comparaison avec un modèle gravitationnel (ce dernier n'étant pas plus prédictif vis-à-vis des observables sismiques que le modèle stellaire ne l'est quant aux trajectoires des astres qui l'entourent). Dans ce dernier cas, le conflit entre les différentes significations prêtées à la même notion favoriserait alors le modèle le moins complexe (et donc le plus "proche" de l'observation) au détriment des autres où l'abstraction du paramètre se ferait plus saillante et où il serait donc davantage perçu comme un moyen d'ajuster les modèles.

C.2.7 La prédiction de données

On s'est restreint jusqu'ici à l'étude des plausibilités de paramètres et de modèles mais le formalisme de l'inférence bayésienne permet en réalité d'aborder des questions bien plus générales. On étudiera ici la possibilité d'anticiper des données futures, $\mathcal{D}_* = \{\mathbf{y}_*, X_*\}$, à partir d'observations passées \mathcal{D} . L'objectif se résume alors à estimer la plausibilité $p(\mathbf{y}_*|X_*, \mathcal{D})$ des mesures futures \mathbf{y}_* aux points X_* et l'on supposera avoir à notre disposition l'ensemble de modèles $\{\mathcal{M}_i\}_{i \in \mathcal{I}}$. Commençons par décomposer la probabilité jointe des mesures passées et futures :

$$p(\mathbf{y}_*, \mathbf{y}, \mathcal{M}_i | X_*, X) = p(\mathbf{y}_* | X_*, \mathcal{D}, \mathcal{M}_i) p(\mathcal{M}_i | \mathcal{D}) p(\mathbf{y} | X) \quad (\text{C.2.27})$$

et remarquons qu'à modèle \mathcal{M}_i imposé, les données passées n'influent aucunement l'anticipation des données futures : $p(\mathbf{y}_* | X_*, \mathcal{D}, \mathcal{M}_i) = p(\mathbf{y}_* | X_*, \mathcal{M}_i)$. D'autre part, la loi de probabilité conditionnelle $p(\mathbf{y}_* | X_*, \mathcal{D})$ s'exprime comme :

$$p(\mathbf{y}_* | X_*, \mathcal{D}) = \frac{p(\mathbf{y}_*, \mathbf{y} | X_*, X)}{p(\mathbf{y} | X)} = \frac{1}{p(\mathbf{y} | X)} \sum_{i \in \mathcal{I}} p(\mathbf{y}_*, \mathbf{y}, \mathcal{M}_i | X_*, X) \quad (\text{C.2.28})$$

et la combinaison des deux expressions donne :

$$p(\mathbf{y}_* | X_*, \mathcal{D}) = \sum_{i \in \mathcal{I}} p(\mathbf{y}_* | X_*, \mathcal{M}_i) p(\mathcal{M}_i | \mathcal{D}). \quad (\text{C.2.29})$$

La formule obtenue est une nouvelle fois assez instructive : on voit que l'anticipation des données futures combine une nouvelle fois les anticipations propres à chaque modèle (avec le terme $p(\mathbf{y}_* | X_*, \mathcal{M}_i)$). Ces avis sont cependant pondérés par la crédibilité attribuée par les données passées à chaque modèle (terme en $p(\mathcal{M}_i | \mathcal{D})$), traduisant le fait qu'un modèle peu pertinent par le passé n'a que peu de chance d'anticiper correctement l'avenir.

C.2.8 La théorie bayésienne de la décision

Si l'on a parcouru un certain nombre de propriétés sur les lois de probabilité jusqu'ici, on ne s'est pour l'instant pas attardé sur la façon de choisir un modèle (ou un jeu de paramètres) connaissant $p(\mathcal{M}_i | \mathcal{D})$ (respectivement $p(\theta | \mathcal{M}_i, \mathcal{D})$). Le formalisme bayésien a cependant pour propriété de s'agencer de façon élégante avec la théorie de la décision pour former la théorie bayésienne de la décision (BERGER, 1985) et apporter des réponses à ces questions. En prenant le cas du choix de paramètres, la théorie nous propose alors de définir le coût $\mathcal{L}(\theta^{\text{réel}}, \theta)$ engendré par le choix de θ si la valeur réelle des paramètres est $\theta^{\text{réel}}$. Notons que si $\theta^{\text{réel}}$ est connu avec certitude, le problème est alors trivial et se résume à la minimisation de $\mathcal{L}(\theta^{\text{réel}}, \theta)$, objectif généralement atteint par le choix $\theta_* = \theta^{\text{réel}}$. Lorsque l'on dispose plutôt d'une distribution $p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D})$ estimée à partir des données \mathcal{D} , la théorie nous propose de choisir θ de façon à minimiser le coût espéré à partir de cette distribution :

$$\begin{aligned} \theta_* &= \arg \min_{\theta} \mathbb{E} \left[\mathcal{L}(\theta^{\text{réel}}, \theta) \right]_{p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D})} \\ &= \arg \min_{\theta} \int_{\theta^{\text{réel}} \in \Theta_{\mathcal{M}_i}} \mathcal{L}(\theta^{\text{réel}}, \theta) p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D}) d\theta^{\text{réel}}. \end{aligned} \quad (\text{C.2.30})$$

On donnera ici deux exemples classiques d'une telle minimisation. Supposons d'abord que le coût d'une erreur soit tel que $\mathcal{L}(\theta^{\text{réel}}, \theta) \propto -\delta(\theta^{\text{réel}} - \theta)$ avec $\delta(x)$ la fonction de Dirac. Dans un tel scénario, le coût de la moindre erreur est infini et l'intégrale s'obtient immédiatement :

$$\int_{\theta^{\text{réel}} \in \Theta_{\mathcal{M}_i}} \mathcal{L}(\theta^{\text{réel}}, \theta) p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D}) d\theta^{\text{réel}} \propto -p(\theta | \mathcal{M}_i, \mathcal{D}). \quad (\text{C.2.31})$$

La valeur minimisant ce coût espéré est donc le vecteur θ qui maximise la plausibilité des paramètres à posteriori, que l'on appelle l'estimateur MAP (pour maximum à posteriori) :

$$\theta^{\text{MAP}} \equiv \arg \max_{\theta} p(\theta | \mathcal{M}_i, \mathcal{D}) \quad (\text{C.2.32})$$

auquel on fait régulièrement référence dans l'article [HOUDAYER, REESE et GOUPIL, 2022](#).

Imaginons maintenant un second cas où le coût de l'erreur est quadratique : $\mathcal{L}(\theta^{\text{réel}}, \theta) \propto |\theta^{\text{réel}} - \theta|^2$. On introduira également la moyenne de la distribution $p(\theta | \mathcal{M}_i, \mathcal{D})$ afin d'alléger les notations :

$$\mu \equiv \mathbb{E}[\theta]_{p(\theta | \mathcal{M}_i, \mathcal{D})} = \int_{\theta \in \Theta_{\mathcal{M}_i}} \theta p(\theta | \mathcal{M}_i, \mathcal{D}) d\theta. \quad (\text{C.2.33})$$

Notons maintenant que la fonction $|\theta^{\text{réel}} - \theta|^2$ se décompose de la façon suivante :

$$\begin{aligned} |\theta^{\text{réel}} - \theta|^2 &= |\theta^{\text{réel}} - \mu + \mu - \theta|^2 \\ &= |\theta^{\text{réel}} - \mu|^2 + 2(\mu - \theta)^T (\theta^{\text{réel}} - \mu) + |\mu - \theta|^2, \end{aligned} \quad (\text{C.2.34})$$

ce qui amène l'apparition de trois intégrales distinctes :

$$\begin{aligned} \int_{\theta^{\text{réel}} \in \Theta_{\mathcal{M}_i}} \mathcal{L}(\theta^{\text{réel}}, \theta) p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D}) d\theta^{\text{réel}} &\propto \int_{\theta^{\text{réel}} \in \Theta_{\mathcal{M}_i}} |\theta^{\text{réel}} - \theta|^2 p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D}) d\theta^{\text{réel}}, \\ &\propto \int_{\theta^{\text{réel}} \in \Theta_{\mathcal{M}_i}} |\theta^{\text{réel}} - \mu|^2 p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D}) d\theta^{\text{réel}} \\ &\quad + 2(\mu - \theta)^T \int_{\theta^{\text{réel}} \in \Theta_{\mathcal{M}_i}} (\theta^{\text{réel}} - \mu) p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D}) d\theta^{\text{réel}} \\ &\quad + |\mu - \theta|^2 \int_{\theta^{\text{réel}} \in \Theta_{\mathcal{M}_i}} p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D}) d\theta^{\text{réel}}, \\ &\propto \text{Tr}(\Sigma) + |\mu - \theta|^2 \end{aligned} \quad (\text{C.2.35})$$

où l'on a utilisé la définition de la moyenne (C.2.33) pour éliminer la seconde intégrale et introduit la matrice de covariance de la distribution :

$$\Sigma \equiv \mathbb{E}[(\theta - \mu)(\theta - \mu)^T]_{p(\theta | \mathcal{M}_i, \mathcal{D})} = \int_{\theta \in \Theta_{\mathcal{M}_i}} (\theta - \mu)(\theta - \mu)^T p(\theta | \mathcal{M}_i, \mathcal{D}) d\theta. \quad (\text{C.2.36})$$

Cette matrice ne dépend pas du choix des paramètres et il vient immédiatement que la valeur de θ minimisant le coût espéré est donné par :

$$\theta_* = \mu. \quad (\text{C.2.37})$$

Notons qu'il est possible de modifier toutes les formules énoncées plus haut en cherchant à minimiser le coût espéré tout en prenant en compte l'avis de l'ensemble des modèles (plutôt que celui de \mathcal{M}_i en particulier). Il suffira alors de remplacer $p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D})$ par la formule (C.2.25) :

$$\theta_* = \arg \min_{\theta} \int_{\theta^{\text{réel}} \in \Theta_{\cap}} \mathcal{L}(\theta^{\text{réel}}, \theta) \left[\sum_{i \in \mathcal{I}} p(\mathcal{M}_i) \int_{\bar{\theta}_i \in \bar{\Theta}_{\mathcal{M}_i}} p(\mathbf{y} | \mathcal{X}, \theta_i, \mathcal{M}_i) p(\theta_i | \mathcal{M}_i) d\bar{\theta}_i \right] d\theta^{\text{réel}} \quad (\text{C.2.38})$$

avec $\Theta_{\cap} = \bigcap_{i \in \mathcal{I}} \Theta_{\mathcal{M}_i}$ l'espace de paramètres communs.

Plus généralement, plutôt que le seul choix d'une valeur θ dont on connaîtrait la distribution, le formalisme de la théorie bayésienne de la décision s'étend au choix d'une action, \mathbf{a} , dont on

connaît le coût, $\mathcal{L}(\theta^{\text{réel}}, \mathbf{a})$, en fonction des paramètres sous-jacents¹⁰, $\theta^{\text{réel}}$. Ici, \mathbf{a} peut référer à peu près à n'importe quoi, que ce soit une variable discrète (*ex* : Choisir entre a et $\neg a$ comme dénoncer ou non un collègue qui se trouverait dans une cellule voisine, scénario d'ailleurs étonnamment fréquent dans l'esprit d'un théoricien des jeux), continue (*ex* : Choisir a heures pour réviser le prochain examen) ou même un vecteur (*ex* : Répartir a_i minutes de son temps sur N activités). Encore une fois, la solution proposée par le formalisme est de choisir l'action qui minimise son coût espéré d'après la distribution $p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D})$:

$$\mathbf{a}_* = \arg \min_{\mathbf{a}} \int_{\theta^{\text{réel}} \in \Theta_{\mathcal{M}_i}} \mathcal{L}(\theta^{\text{réel}}, \mathbf{a}) p(\theta^{\text{réel}} | \mathcal{M}_i, \mathcal{D}) d\theta^{\text{réel}} \quad (\text{C.2.39})$$

ou bien $p(\theta^{\text{réel}} | \mathcal{D})$:

$$\mathbf{a}_* = \arg \min_{\mathbf{a}} \int_{\theta^{\text{réel}} \in \Theta_{\cap}} \mathcal{L}(\theta^{\text{réel}}, \mathbf{a}) \left[\sum_{i \in \mathcal{I}} p(\mathcal{M}_i) \int_{\theta_i \in \Theta_{\mathcal{M}_i}} p(\mathbf{y} | X, \theta_i, \mathcal{M}_i) p(\theta_i | \mathcal{M}_i) d\bar{\theta}_i \right] d\theta^{\text{réel}}. \quad (\text{C.2.40})$$

C.3 Les processus gaussiens

C.3.1 Les processus

Les lois de probabilité étudiées dans le paragraphe précédent s'appliquaient soit à une variable discrète (la collection de modèles $\{\mathcal{M}_i\}_{i \in \mathcal{I}}$ par exemple) ou continue (l'ensemble des paramètres du modèle \mathcal{M}_i : $\theta \in \Theta_{\mathcal{M}_i}$). On appelle *processus* la généralisation du caractère stochastique de ces variables à l'espace des fonctions continues. Formellement, si l'on pouvait considérer une variable aléatoire $Y(\omega)$ comme une fonction associant une valeur \mathbf{y} à tout évènement de l'univers $\omega \in \Omega$ (information qui était contenue dans la loi $p(Y = \mathbf{y})$), le processus $F(\mathbf{x}, \omega)$ peut se voir comme l'association à tout évènement ω d'une fonction $f(\mathbf{x})$. La loi de probabilité décrivant le processus gaussien se définit par analogie comme $p(F(\mathbf{x}) = f(\mathbf{x}) | \mathbf{x})$, ce que l'on abrégera par $p(f(\mathbf{x}) | \mathbf{x})$, et qui s'interprète comme la probabilité que le processus aboutisse à la valeur $f(\mathbf{x})$ en \mathbf{x} . Notons que l'on s'interrogera plus souvent sur la probabilité associée à un ensemble de valeurs $\{f(\mathbf{x}_i)\}_{i \in \{1, \dots, N\}}$ et il sera alors utile de rassembler les \mathbf{x}_i en une matrice, X , et les valeurs de $f(\mathbf{x}_i)$ en un vecteur, \mathbf{f} . La loi de probabilité associée pourra alors être écrite de façon compacte comme $p(\mathbf{f} | X)$.

À l'instar des variables aléatoires classiques, le processus $f(\mathbf{x})$ peut-être caractérisé par un certain nombre de propriétés à commencer par sa moyenne :

$$m(\mathbf{x}) \equiv \mathbb{E} [f(\mathbf{x})]_{p(f(\mathbf{x}) | \mathbf{x})} = \int_{f(\mathbf{x}) \in \mathbb{F}_x} f(\mathbf{x}) p(f(\mathbf{x}) | \mathbf{x}) d[f(\mathbf{x})], \quad (\text{C.3.1})$$

avec \mathbb{F}_x l'espace des valeurs du processus à \mathbf{x} fixé. Remarquons qu'à \mathbf{x} fixé, cette définition de la moyenne correspond à celle d'une variable aléatoire classique. C'est cependant la possibilité de faire varier continûment la valeur de \mathbf{x} qui fait la particularité des processus ; la moyenne $m(\mathbf{x})$ apparaît alors comme une fonction continue qui peut être évaluée en tout point de l'espace de définition du processus. On définira de façon analogue, non plus une matrice de covariance mais un *kernel de covariance*¹¹, $k(\mathbf{x}, \mathbf{x}')$:

$$\begin{aligned} k(\mathbf{x}, \mathbf{x}') &\equiv \mathbb{E} [(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))]_{p(\mathbf{f} | X)} \\ &= \int_{\mathbf{f} \in \mathbb{F}_{\mathbf{x}, \mathbf{x}'}} (f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}')) p(\mathbf{f} | X) d\mathbf{f} \end{aligned} \quad (\text{C.3.2})$$

10. On emploie généralement la terminologie d'"état des choses" (*state of nature* [en]) (BERGER, 1985) dans ce contexte.

11. Appelé des fois fonction de covariance ou kernel tout court.

qui assigne la covariance du processus à tout couple de points $(\mathbf{x}, \mathbf{x}')$, avec X la matrice composée des vecteurs \mathbf{x} et \mathbf{x}' et \mathbf{f} le vecteur composé de $f(\mathbf{x})$ et $f(\mathbf{x}')$.

C.3.2 Propriétés des processus gaussiens

Les processus gaussiens forment une classe de processus qui a ceci de particulier qu'il suffit de spécifier une moyenne, $m(\mathbf{x})$, ainsi qu'un kernel $k(\mathbf{x}, \mathbf{x}')$ pour les définir complètement. On dira que la fonction f suit un processus gaussien de la façon suivante :

$$f(\mathbf{x}) \sim \mathcal{GP}(m(\mathbf{x}), k(\mathbf{x}, \mathbf{x}')). \quad (\text{C.3.3})$$

Plus précisément, les processus gaussiens se caractérisent par le fait que toute évaluation de f sur un ensemble de points X suit une loi normale multivariée de moyenne $\mathbf{m} = m(X)$ et de matrice de covariance $K = k(X, X)$ ¹² :

$$f(X) = \mathbf{f}|X \sim \mathcal{N}(\mathbf{m}, K). \quad (\text{C.3.4})$$

On peut déjà intuitiver que cette propriété va imposer des restrictions sur ce qui peut être considéré comme un processus gaussien ou non. En effet, si l'on considère par exemple deux ensembles de points X_1 et X_2 et $\mathbf{f}_1, \mathbf{f}_2$ les évaluations respectives du processus sur ces points, on voit que :

$$\mathbf{f}_1|X_1 \sim \mathcal{N}(\mathbf{m}_1, K_{11}) \quad (\text{C.3.5})$$

$$\mathbf{f}_2|X_2 \sim \mathcal{N}(\mathbf{m}_2, K_{22}). \quad (\text{C.3.6})$$

Cependant, l'évaluation de f sur la totalité des points composant à la fois X_1 et X_2 doit également suivre une loi normale d'après la propriété (C.3.4). On notera alors :

$$\begin{bmatrix} \mathbf{f}_1|X_1 \\ \mathbf{f}_2|X_2 \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix}, \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \right). \quad (\text{C.3.7})$$

En effet, les moyennes et matrices de covariances sur X_1 et X_2 sont déjà imposées par les lois (C.3.5) et (C.3.6), ne laissant qu'une indétermination au niveau des matrices de covariance K_{12} et K_{21} , lesquelles doivent par ailleurs vérifier :

$$K_{21} = K_{12}^\top. \quad (\text{C.3.8})$$

Prise telle quelle, la formule (C.3.7) spécifie intégralement la loi de probabilité jointe suivie par les deux vecteurs : $p(\mathbf{f}_1, \mathbf{f}_2|X_1, X_2)$. Une question que l'on peut alors se poser (et qui se trouve au cœur de l'immense majorité des propriétés attribuées aux processus gaussiens) est quelle est la loi suivie par la variable $\mathbf{f}_1|\mathbf{f}_2, X_1, X_2$? Ou autrement formulé : connaissant les valeurs prises par le processus en X_2 , quelle est la loi suivie par le processus en X_1 ?

La réponse à cette question est loin d'être triviale, et la démonstration qui suit s'inspire en grande partie de celle proposée par **VON MISES, 1964**. L'astuce est de commencer par considérer les variables intermédiaires \mathbf{g}_1 et \mathbf{g}_2 , que l'on définira par :

$$\mathbf{g}_1|X_1, X_2 = \mathbf{f}_1|X_1 - K_{12}K_{22}^{-1}\mathbf{f}_2|X_2 \quad (\text{C.3.9})$$

$$\mathbf{g}_2|X_1, X_2 = \mathbf{f}_2|X_2. \quad (\text{C.3.10})$$

On introduira la matrice associée à cette transformation linéaire :

$$T \equiv \begin{bmatrix} I & -K_{12}K_{22}^{-1} \\ 0 & I \end{bmatrix}. \quad (\text{C.3.11})$$

12. On entend par cette écriture que $K_{ij} = k(\mathbf{x}_i, \mathbf{x}_j)$ pour \mathbf{x}_i et \mathbf{x}_j respectivement les vecteurs i et j formant la matrice X .

Puisque le vecteur $[\mathbf{g}_1|X_1, X_2 \quad \mathbf{g}_2|X_1, X_2]^T$ résulte d'une combinaison linéaire de variables suivant une loi normale, ce dernier suit donc lui-même une loi normale de moyenne :

$$\mathbf{T}\mathbf{m} = \begin{bmatrix} \mathbf{I} & -\mathbf{K}_{12}\mathbf{K}_{22}^{-1} \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{m}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{m}_2 \\ \mathbf{m}_2 \end{bmatrix} \quad (\text{C.3.12})$$

et de covariance :

$$\mathbf{TKT}^T = \begin{bmatrix} \mathbf{I} & -\mathbf{K}_{12}\mathbf{K}_{22}^{-1} \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{K}_{11} & \mathbf{K}_{12} \\ \mathbf{K}_{21} & \mathbf{K}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{I} & 0 \\ -\mathbf{K}_{22}^{-1}\mathbf{K}_{21} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{K}_{11} - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21} & 0 \\ 0 & \mathbf{K}_{22} \end{bmatrix}. \quad (\text{C.3.13})$$

La matrice de covariance étant diagonale par bloc, on en déduit que les variables \mathbf{g}_1 et \mathbf{g}_2 sont indépendantes et donc que leur loi jointe est donnée par le produit de leurs lois marginales :

$$p(\mathbf{g}_1, \mathbf{g}_2|X_1, X_2) = \mathcal{G}(\mathbf{g}_1, \mathbf{m}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{m}_2, \mathbf{K}_{11} - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21})\mathcal{G}(\mathbf{g}_2, \mathbf{m}_2, \mathbf{K}_{22}). \quad (\text{C.3.14})$$

Puisque le jacobien de la transformation linéaire vaut 1, la loi jointe entre \mathbf{f}_1 et \mathbf{f}_2 s'obtient simplement par substitution de \mathbf{g}_1 par $\mathbf{f}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{f}_2$ et de \mathbf{g}_2 par \mathbf{f}_2 :

$$p(\mathbf{f}_1, \mathbf{f}_2|X_1, X_2) = \mathcal{G}(\mathbf{f}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{f}_2, \mathbf{m}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{m}_2, \mathbf{K}_{11} - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21})\mathcal{G}(\mathbf{f}_2, \mathbf{m}_2, \mathbf{K}_{22}). \quad (\text{C.3.15})$$

Cependant, la loi marginale de \mathbf{f}_2 est précisément $\mathcal{G}(\mathbf{f}_2, \mathbf{m}_2, \mathbf{K}_{22})$ par hypothèse et puisque la formule des lois conditionnelles donne $p(\mathbf{f}_1|\mathbf{f}_2, X_1, X_2) = p(\mathbf{f}_1, \mathbf{f}_2|X_1, X_2) / p(\mathbf{f}_2|X_1, X_2)$, on obtient :

$$p(\mathbf{f}_1|\mathbf{f}_2, X_1, X_2) = \mathcal{G}(\mathbf{f}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{f}_2, \mathbf{m}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{m}_2, \mathbf{K}_{11} - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21}). \quad (\text{C.3.16})$$

Notons que puisque la fonction gaussienne est uniquement fonction de $\mathbf{x} - \boldsymbol{\mu}$, le changement de variable $\mathbf{x} \rightarrow \mathbf{x} + \mathbf{x}_0$, $\boldsymbol{\mu} \rightarrow \boldsymbol{\mu} + \mathbf{x}_0$ laisse la fonction invariante et l'on pourra écrire plus simplement cette loi comme seule fonction de \mathbf{f}_1 :

$$p(\mathbf{f}_1|\mathbf{f}_2, X_1, X_2) = \mathcal{G}(\mathbf{f}_1, \mathbf{m}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}(\mathbf{m}_2 - \mathbf{f}_2), \mathbf{K}_{11} - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21}). \quad (\text{C.3.17})$$

On vient donc de démontrer que la variable conditionnelle $\mathbf{f}_1|\mathbf{f}_2, X_1, X_2$ suit également une loi normale :

$$\mathbf{f}_1|\mathbf{f}_2, X_1, X_2 \sim \mathcal{N}(\mathbf{m}_1 - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}(\mathbf{m}_2 - \mathbf{f}_2), \mathbf{K}_{11} - \mathbf{K}_{12}\mathbf{K}_{22}^{-1}\mathbf{K}_{21}) \quad (\text{C.3.18})$$

de moyenne et de matrice de covariance connues dès lors que l'on connaît la loi jointe. Cette propriété est centrale pour les formules qui vont suivre et dont l'objectif sera souvent d'estimer une loi en un point \mathbf{x}_* pour des données $\mathcal{D} = \{\mathbf{y}, X\}$ connues.

C.3.3 Application au formalisme de l'inférence

Commençons par supposer que l'on a observé les valeurs du processus \mathbf{f} sur les points X et que l'on cherche à estimer la plausibilité de la valeur $f_* = f(\mathbf{x}_*)$ en \mathbf{x}_* . Connaissant la moyenne $m(\mathbf{x})$ et le kernel $k(\mathbf{x}, \mathbf{x}')$ du processus, on sait que la loi jointe suit une loi normale :

$$\begin{bmatrix} f_*|\mathbf{x}_* \\ \mathbf{f}|X \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} m_* \\ \mathbf{m} \end{bmatrix}, \begin{bmatrix} k_{**} & \mathbf{k}_*^T \\ \mathbf{k}_* & \mathbf{K} \end{bmatrix} \right) \quad (\text{C.3.19})$$

avec $m_* = m(\mathbf{x}_*)$, $k_{**} = k(\mathbf{x}_*, \mathbf{x}_*)$ et $\mathbf{k}_* = k(X, \mathbf{x}_*)$. L'équation (C.3.18) nous fournit alors immédiatement la loi conditionnelle souhaitée :

$$f_*|\mathbf{f}, \mathbf{x}_*, X \sim \mathcal{N}(m_* - \mathbf{k}_*^T \mathbf{K}^{-1}(\mathbf{m} - \mathbf{f}), k_{**} - \mathbf{k}_*^T \mathbf{K}^{-1} \mathbf{k}_*). \quad (\text{C.3.20})$$

Il paraît cependant peu vraisemblable en pratique de mesurer directement les valeurs \mathbf{f} du processus sous-jacent. On aura en général seulement accès à des valeurs \mathbf{y} entachées d'un bruit,

\mathbf{n} , dont on suppose en général la distribution normale (et non biaisée) :

$$\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \Sigma_n). \quad (\text{C.3.21})$$

La loi suivie par les mesures \mathbf{y} se voit alors modifiée par rapport au cas précédent :

$$\mathbf{y}|X = \mathbf{f}|X + \mathbf{n}|X \sim \mathcal{N}(\mathbf{m}, \mathbf{K} + \Sigma_n), \quad (\text{C.3.22})$$

ce que l'on peut introduire dans la loi jointe :

$$\begin{bmatrix} f_*|\mathbf{x}_* \\ \mathbf{y}|X \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} m_* \\ \mathbf{m} \end{bmatrix}, \begin{bmatrix} k_{**} & \mathbf{k}_*^\top \\ \mathbf{k}_* & \mathbf{K} + \Sigma_n \end{bmatrix} \right). \quad (\text{C.3.23})$$

On obtient alors une nouvelle distribution conditionnelle pour $f_*|\mathbf{x}_*, \mathcal{D}$:

$$f_*|\mathbf{x}_*, \mathcal{D} \sim \mathcal{N}(m_* - \mathbf{k}_*^\top (\mathbf{K} + \Sigma_n)^{-1} (\mathbf{m} - \mathbf{y}), k_{**} - \mathbf{k}_*^\top (\mathbf{K} + \Sigma_n)^{-1} \mathbf{k}_*) \quad (\text{C.3.24})$$

et dans le cas particulier où la moyenne du processus est nulle (ce qu'il est courant de supposer à priori), il vient :

$$f_*|\mathbf{x}_*, \mathcal{D} \sim \mathcal{N}(\mathbf{k}_*^\top (\mathbf{K} + \Sigma_n)^{-1} \mathbf{y}, k_{**} - \mathbf{k}_*^\top (\mathbf{K} + \Sigma_n)^{-1} \mathbf{k}_*). \quad (\text{C.3.25})$$

Ainsi, si l'on souhaite faire une prédiction de la valeur du processus en \mathbf{x}_* , on l'effectuera généralement en utilisant la moyenne de la distribution¹³ (C.3.25) :

$$\overline{f_*|\mathbf{x}_*, \mathcal{D}} = \mathbf{k}_*^\top (\mathbf{K} + \Sigma_n)^{-1} \mathbf{y}, \quad (\text{C.3.26})$$

propriété que l'on a utilisée dans [HOUDAYER, REESE et GOUPIL, 2022](#) pour l'estimation des différences de fréquences (et qui est automatiquement calculée dans à peu près tous les programmes d'inférence par processus gaussien).

À noter que si l'on souhaite effectuer le calcul en pratique pour un ensemble de points X_* , il peut être avantageux de tirer profit du produit matriciel et de calculer directement

$$\overline{\mathbf{f}_*|X_*, \mathcal{D}} = \mathbf{K}_*^\top (\mathbf{K} + \Sigma_n)^{-1} \mathbf{y} \quad (\text{C.3.27})$$

avec $\mathbf{K}_* = k(X, X_*)$ plutôt que de calculer un à un les f_* . La matrice de covariance de cette estimation est alors déduite du produit matriciel :

$$\text{Cov}[\mathbf{f}_*|X_*, \mathcal{D}] = \mathbf{K}_{**} - \mathbf{K}_*^\top (\mathbf{K} + \Sigma_n)^{-1} \mathbf{K}_* \quad (\text{C.3.28})$$

avec $\mathbf{K}_{**} = k(X_*, X_*)$. Une visualisation des estimations proposées par les formules (C.3.27) et (C.3.28) pour un jeu de données fictif est donnée en [FIGURE C.2](#). En particulier, on peut voir que l'estimation moyenne amène une explication crédible à la distribution des données observées.

C.3.4 Kernels, "paramètres" et hyperparamètres d'un processus gaussien

Dans le cas d'un choix de moyenne nulle à priori pour le processus f , on peut voir que sa définition complète résulte du seul choix du kernel $k(\mathbf{x}, \mathbf{x}')$ (cf. équation C.3.3). Par définition, cette fonction impose la corrélation entre les valeurs de f prises en \mathbf{x} et \mathbf{x}' . Ainsi, une forte valeur de $k(\mathbf{x}, \mathbf{x}')$ conduira à des comportements similaires du processus sur ces deux points qui seront alors fortement corrélés alors qu'une valeur proche de zéro tendra à rendre indépendantes les valeurs de $f(\mathbf{x})$ et $f(\mathbf{x}')$. Bien que la fonction k doive satisfaire un certain nombre de contraintes (voir [RASMUSSEN et WILLIAMS, 2005](#)), les choix potentiels sur sa forme sont quasi-illimités et peuvent mener à des résultats très différents. Il est donc crucial de s'interroger à

13. Puisqu'il s'agit à la fois de la moyenne, du mode et de la médiane de la distribution gaussienne.

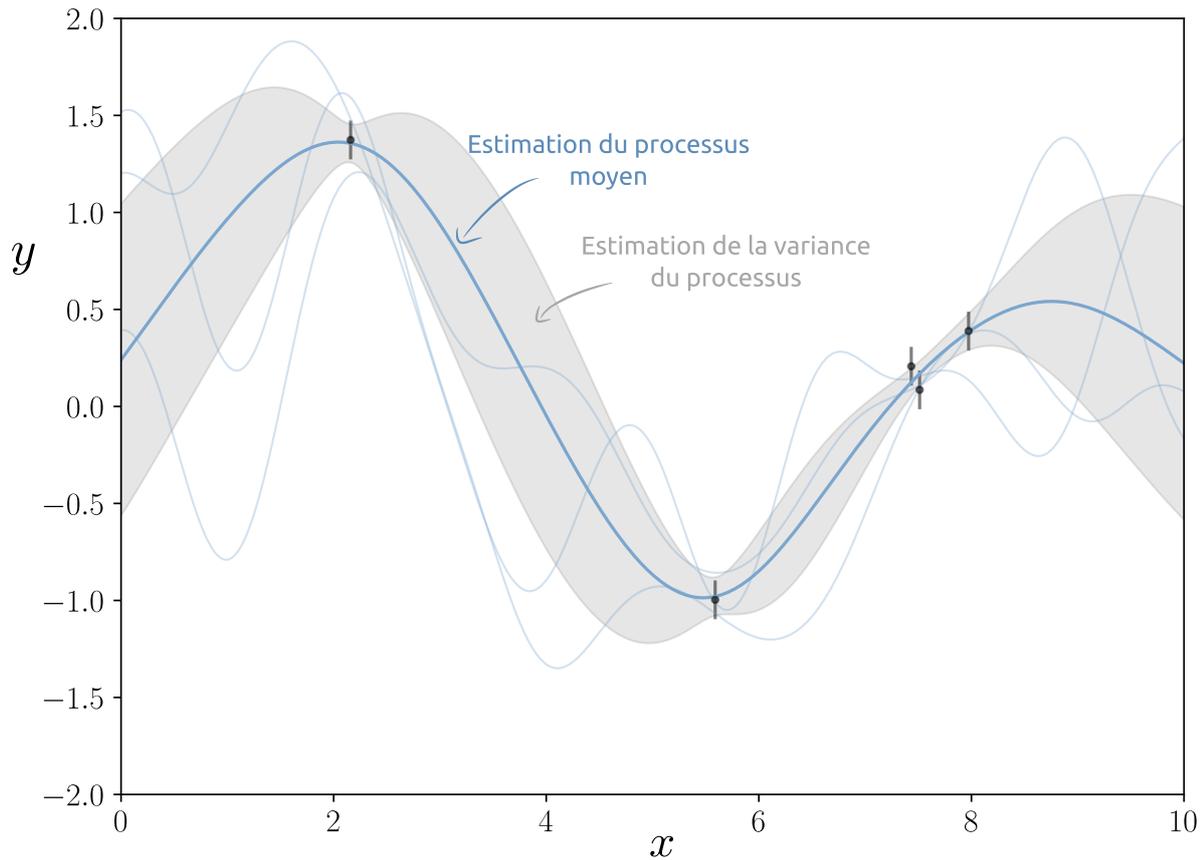


FIGURE C.2 – Illustration de la loi suivie par le processus gaussien f , pour un jeu de données fictif, $\mathcal{D} = \{X, y\}$ (en noir), entaché d'un bruit non corrélé ($\Sigma_n = 0.1^2 \cdot I$). Les trois courbes les plus claires ont été tirées suivant la loi (C.3.25) pour un grand nombre de points x_* , la courbe foncée donne l'estimation de la moyenne de f connaissant les données \mathcal{D} à l'aide de la formule (C.3.27) tandis que la zone grise illustre la variance du processus issue des termes diagonaux de l'expression (C.3.28) autour de cette moyenne.

priori sur la forme adéquate à donner au processus, ce que l'on déduira généralement des caractéristiques physiques du problème. Par exemple, on remarquera que des valeurs élevées de $k(x, x')$ autour des points vérifiant $x = x'$ tend à faire apparaître des processus "lisses" là où des kernels diminuant rapidement avec $|x - x'|$ conduisent à des courbes beaucoup plus erratiques (de type "cours boursier"). En effet, dans ce dernier cas, deux points voisins peuvent être relativement peu corrélés et donc faire apparaître des brisures de pentes s'apparentant à du bruit. De l'autre côté du spectre, des covariances restant très élevées pour sur une grande distance $|x - x'|$ induisent des processus ne variant que sur les très grandes échelles (le cas limite étant le choix $k(x, x') = \text{cste}$ qui donne lieu à un processus constant). Un compromis alors courant pour obtenir un processus gaussien continu mais variable est celui de la fonction exponentielle au carré :

$$k(x, x') = e^{-|x-x'|^2/2}. \quad (\text{C.3.29})$$

On voit ici que la covariance entre x et x' décroît de façon monotone avec la distance entre ces points tout en formant un plateau autour de $|x - x'| = 0$, ce qui assure sa dérivabilité en tout point. Notons que plutôt que de choisir une fonction fixe pour k comme avec l'équation (C.3.29), il est courant de définir une classe de fonctions paramétrées : $k(x, x', \theta)$. Par exemple pour la fonction exponentielle au carré, le choix d'une longueur caractéristique de corrélation, ℓ , peut

amener à définir la classe de fonctions :

$$k(\mathbf{x}, \mathbf{x}', \ell) = \exp \left[-\frac{|\mathbf{x} - \mathbf{x}'|^2}{2\ell^2} \right]. \quad (\text{C.3.30})$$

On voit que des valeurs de ℓ faibles tendent à faire apparaître des processus très variables et inversement. Une seconde classe de fonctions très utilisée¹⁴ est la suivante :

$$k(\mathbf{x}, \mathbf{x}', P) = \exp \left[-\sin^2 \left(\pi \frac{|\mathbf{x} - \mathbf{x}'|}{P} \right) \right]. \quad (\text{C.3.31})$$

Les propriétés de l'exponentielle assurent toujours la dérivabilité lorsque le sinus tend vers 0 (et donc que $|\mathbf{x} - \mathbf{x}'| \rightarrow 0$), à ceci près que deux points distants d'un multiple de P seront à nouveau très fortement corrélés. Ces fonctions sont donc utilisées pour générer des processus périodiques ou quasi-périodiques (c'est d'ailleurs le nom qu'on leur prête) de période P , particulièrement adaptés pour reproduire des schéma répétitifs dans les données.

Il existe ainsi de très nombreuses classes de fonctions, lesquelles peuvent être combinées pour modéliser des comportements particuliers. En effet, une propriété fondamentale des kernels est la suivante : si $f_1(\mathbf{x})$ et $f_2(\mathbf{x})$ sont deux processus gaussiens de kernels respectifs $k_1(\mathbf{x}, \mathbf{x}')$ et $k_2(\mathbf{x}, \mathbf{x}')$, alors $f_1(\mathbf{x}) + f_2(\mathbf{x})$ et $f_1(\mathbf{x})f_2(\mathbf{x})$ sont également des processus gaussiens ayant pour kernels respectifs $k_1(\mathbf{x}, \mathbf{x}') + k_2(\mathbf{x}, \mathbf{x}')$ et $k_1(\mathbf{x}, \mathbf{x}')k_2(\mathbf{x}, \mathbf{x}')$ (RASMUSSEN et WILLIAMS, 2005)! Ainsi, la classe de fonctions utilisée pour le processus gaussien apparaissant en FIGURE C.2 est la suivante :

$$k(\mathbf{x}, \mathbf{x}', \theta) = \exp \left[-\frac{|\mathbf{x} - \mathbf{x}'|^2}{2\ell^2} \right] + \alpha \exp \left[-\sin^2 \left(\pi \frac{|\mathbf{x} - \mathbf{x}'|}{P} \right) \right]. \quad (\text{C.3.32})$$

avec $\theta = [P \ \ell \ \alpha]^T$. Le processus engendré varie alors sur une échelle de l'ordre de ℓ tout en possédant un certain de degré α de périodicité sur une période P . Ce caractère est particulièrement flagrant à la vue de la fonction $k(\mathbf{x}, \mathbf{x}', \theta)$ représentée en fonction de \mathbf{x} et \mathbf{x}' après ajustement (cf. FIGURE C.3), qui correspond à la matrice de covariance à posteriori issue de l'équation (C.3.28) pour la valeur de θ choisie :

$$\text{Cov} [\mathbf{f}_* | X_*, \mathcal{D}, \theta] = \mathbf{K}_{**}(\theta) - \mathbf{K}_*(\theta)^T (\mathbf{K}(\theta) + \Sigma_n)^{-1} \mathbf{K}_*(\theta) \quad (\text{C.3.33})$$

avec $\mathbf{K}(\theta) = k(X, X, \theta)$, $\mathbf{K}_*(\theta) = k(X, X_*, \theta)$ et $\mathbf{K}_{**}(\theta) = k(X_*, X_*, \theta)$.

On peut alors s'interroger sur la façon de choisir ces paramètres et sur le sens à donner au vecteur θ dans la hiérarchie du modèle que constitue le processus gaussien. S'il l'on peut intuitivement penser que θ désigne des paramètres du modèle, l'analogie avec la hiérarchie définie dans le paragraphe C.2.5 les assimile davantage à des hyperparamètres. En effet, après avoir pris connaissance des données \mathcal{D} , la loi que suivra le processus pour un choix de classe de fonctions $k(\mathbf{x}, \mathbf{x}', \theta)$ donnée sera :

$$\mathbf{f}_* | X_*, \mathcal{D}, \theta \sim \mathcal{N} \left(\mathbf{K}_*(\theta)^T (\mathbf{K}(\theta) + \Sigma_n)^{-1} \mathbf{y}, \mathbf{K}_{**}(\theta) - \mathbf{K}_*(\theta)^T (\mathbf{K}(\theta) + \Sigma_n)^{-1} \mathbf{K}_*(\theta) \right). \quad (\text{C.3.34})$$

Par comparaison avec la loi $\mathbf{w} | \mathcal{D}, \theta, \mathcal{M}_i$ ¹⁵, il apparaît que l'analogie du jeu de paramètre, \mathbf{w} , est plutôt le vecteur \mathbf{f}_* des valeurs du processus aux points X_* . La sélection de la valeur de θ se font donc à un échelon supérieur, en essayant de maximiser la loi $p(\theta | X_*, \mathcal{D}) = p(\theta | \mathbf{y}, X)$ (puisque X_* n'influe aucunement le choix de θ , qui ne dépend que de $\mathcal{D} = \{\mathbf{y}, X\}$). La formule (C.2.18) nous donne alors :

$$p(\theta | \mathbf{y}, X) \propto p(\mathbf{y} | \theta, X) p(\theta), \quad (\text{C.3.35})$$

14. Des exemples d'applications concrètes peuvent être trouvées dans les travaux de AIGRAIN, PARVIAINEN et POPE, 2016; ANGUS et al., 2018.

15. Notons que le choix du modèle \mathcal{M}_i se résume à celui de la classe de fonctions $k(\mathbf{x}, \mathbf{x}', \theta)$ pour un processus gaussien, choix qui n'apparaît pas explicitement dans la loi $\mathbf{f}_* | X_*, \mathcal{D}, \theta$ pour éviter d'alourdir les notations.

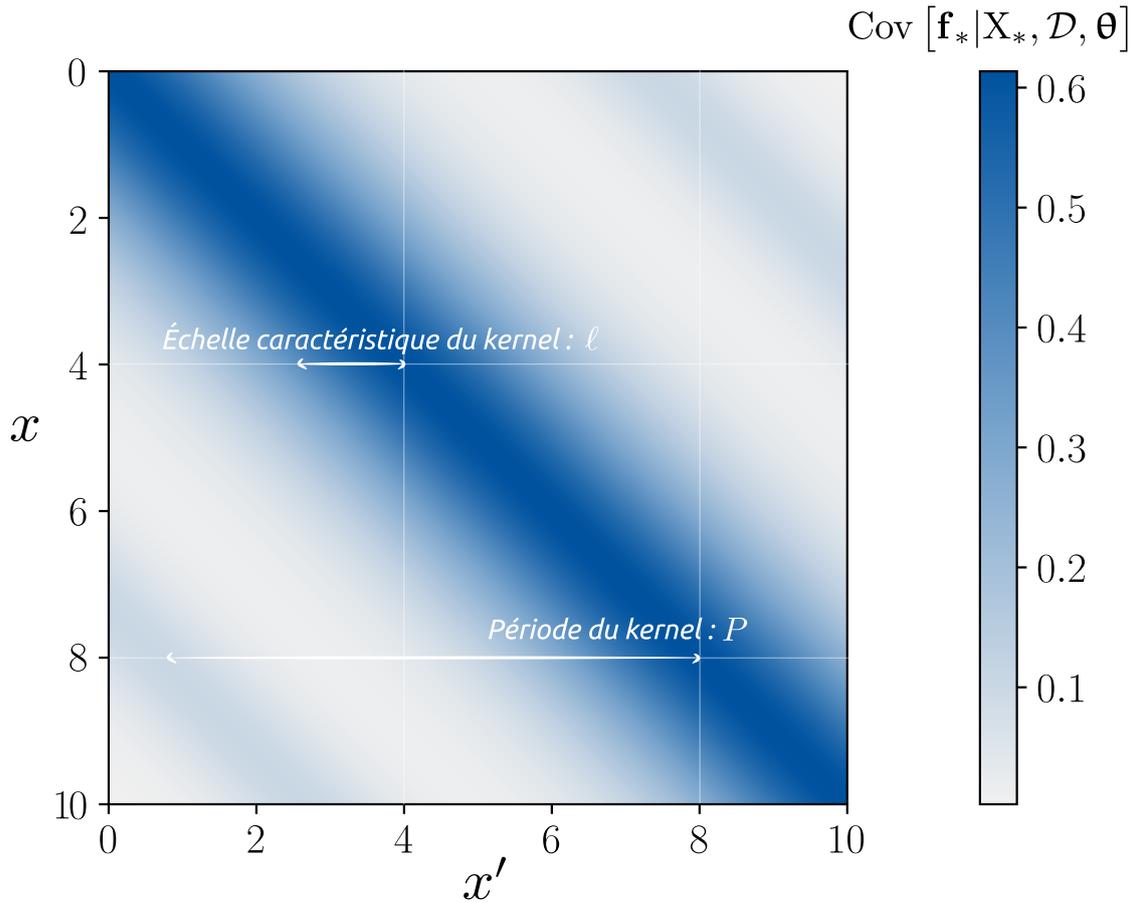


FIGURE C.3 – Représentation de la matrice de covariance du processus gaussien associé à la classe de fonctions (C.3.32) après l’ajustement réalisé en FIGURE C.2. L’échelle caractéristique de variation, ℓ , contrôle la distance sur laquelle la covariance entre deux points chute d’un facteur e tandis que la période du kernel permet un regain de covariance périodique sur une échelle P .

ce qui nécessite donc de spécifier la loi à priori sur les hyperparamètres de la classe de fonction. En pratique cependant, cette distribution est rarement intégrée à l’inférence du vecteur $\boldsymbol{\theta}$ (ou alors supposée uniforme), si bien que le choix des hyperparamètres se résume généralement à la maximisation de l’intégrale :

$$p(\mathbf{y} | \boldsymbol{\theta}, \mathbf{X}) = \int_{\mathbf{f} \in \mathbb{F}_{\mathbf{X}}} p(\mathbf{y} | \mathbf{f}, \boldsymbol{\theta}, \mathbf{X}) p(\mathbf{f} | \boldsymbol{\theta}, \mathbf{X}) d\mathbf{f}. \quad (\text{C.3.36})$$

Or, les mesures \mathbf{y} sont obtenues par simple ajout d’un bruit non biaisé sur \mathbf{f} (cf. équation C.3.22), si bien que sa distribution se résume à :

$$\mathbf{y} | \mathbf{f}, \boldsymbol{\theta}, \mathbf{X} \sim \mathcal{N}(\mathbf{f}, \Sigma_n) \quad (\text{C.3.37})$$

comme fonction de \mathbf{y} (montrant au passage qu’elle ne dépend pas de $\boldsymbol{\theta}$) et donc à :

$$\mathbf{y} | \mathbf{f}, \boldsymbol{\theta}, \mathbf{X} \sim \mathcal{N}(\mathbf{y}, \Sigma_n) \quad (\text{C.3.38})$$

comme fonction de \mathbf{f} (variable sur laquelle porte l'intégrale). De plus, la loi de \mathbf{f} à priori résulte directement de la définition (C.3.4) du processus gaussien :

$$\mathbf{f}|\boldsymbol{\theta}, X \sim \mathcal{N}(\mathbf{0}, \mathbf{K}(\boldsymbol{\theta})), \quad (\text{C.3.39})$$

soulignant que l'intégrand de l'équation (C.3.36) est le produit de deux distributions gaussiennes sur \mathbf{f} et donc lui-même proportionnel à une gaussienne. On a en effet (ABRAMOWITZ et STEGUN, 1964) :

$$\mathcal{G}(\mathbf{x}, \boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1) \mathcal{G}(\mathbf{x}, \boldsymbol{\mu}_2, \boldsymbol{\Sigma}_2) \propto \mathcal{G}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (\text{C.3.40})$$

avec :

$$\boldsymbol{\Sigma} = (\boldsymbol{\Sigma}_1^{-1} + \boldsymbol{\Sigma}_2^{-1})^{-1} = \boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)^{-1}\boldsymbol{\Sigma}_2 \quad (\text{C.3.41})$$

$$\boldsymbol{\mu} = \boldsymbol{\Sigma}(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_2^{-1}\boldsymbol{\mu}_2) = \boldsymbol{\Sigma}_2(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)^{-1}\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2)^{-1}\boldsymbol{\mu}_2, \quad (\text{C.3.42})$$

ce qui, dans notre cas, donne :

$$p(\mathbf{y}|\mathbf{f}, \boldsymbol{\theta}, X)p(\mathbf{f}|\boldsymbol{\theta}, X) \propto \mathcal{G}\left(\mathbf{K}(\boldsymbol{\theta})(\mathbf{K}(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_n)^{-1}\mathbf{y}, \boldsymbol{\Sigma}_n(\mathbf{K}(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_n)^{-1}\mathbf{K}(\boldsymbol{\theta})\right) \quad (\text{C.3.43})$$

et puisque l'intégrale d'une gaussienne vaut tout simplement 1 :

$$p(\mathbf{y}|\boldsymbol{\theta}, X) = z(\mathbf{y}, \boldsymbol{\theta}) \quad (\text{C.3.44})$$

avec z la constante de proportionnalité de l'équation (C.3.43). La détermination de cette constante (par rapport à \mathbf{f}) est cependant importante puisque l'on cherche à maximiser $p(\mathbf{y}|\boldsymbol{\theta}, X)$ comme une fonction de $\boldsymbol{\theta}$. Celle-ci est elle-même gaussienne (en tant que fonction de \mathbf{y} , cf. RASMUSSEN et WILLIAMS, 2005) :

$$p(\mathbf{y}|\boldsymbol{\theta}, X) = z(\mathbf{y}, \boldsymbol{\theta}) = \frac{1}{\sqrt{(2\pi)^N |\mathbf{K}(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_n|}} e^{-\mathbf{y}^\top (\mathbf{K}(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_n)^{-1} \mathbf{y} / 2}. \quad (\text{C.3.45})$$

La valeur de $\boldsymbol{\theta}$ la plus plausible à posteriori, $\boldsymbol{\theta}_*$ minimise donc la fonction suivante¹⁶ :

$$\boldsymbol{\theta}_* = \arg \min_{\boldsymbol{\theta}} \left\{ \mathbf{y}^\top (\mathbf{K}(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_n)^{-1} \mathbf{y} + \ln |\mathbf{K}(\boldsymbol{\theta}) + \boldsymbol{\Sigma}_n| \right\}, \quad (\text{C.3.46})$$

opération qui se mène numériquement en pratique à l'aide du gradient de cette fonction.

C.3.5 Ajout d'une base de fonctions à un processus gaussien

Un dernier point que j'aimerais aborder dans cette annexe concerne l'ajout d'une base de fonctions à un processus gaussien f . Cette question a été abordée en particulier dans l'article HOUDAYER, REESE et GOUPIL, 2022, où l'on disposait à la fois d'un modèle (déterministe) de décomposition linéaire sur une base de fonction :

$$g(X, \boldsymbol{\theta}) \equiv \mathbf{g}(\boldsymbol{\theta}) = \mathbf{G}(X)^\top \boldsymbol{\theta} \quad (\text{C.3.47})$$

et d'un processus gaussien :

$$f(X, \boldsymbol{\varphi}) \equiv \mathbf{f}|\boldsymbol{\varphi}, X \sim \mathcal{N}(\mathbf{0}, \mathbf{K}(\boldsymbol{\varphi})). \quad (\text{C.3.48})$$

16. Il est courant de recourir au logarithme pour ce genre de minimisation.

Les notations de l'article, qui traite du cas spécifique de l'inférence à l'aide du glitch sismique, peuvent être retrouvées en appliquant les transformations :

$$f, g \longrightarrow \delta\sigma_{\text{surf}}^2, \delta\sigma_{\text{mod}}^2 \quad (\text{C.3.49})$$

$$X \longrightarrow \sigma \quad (\text{C.3.50})$$

$$K \longrightarrow K_{\text{surf}} \quad (\text{C.3.51})$$

$$G \longrightarrow \nabla_{\theta} \delta\sigma_{\text{mod}}^2 \quad (\text{C.3.52})$$

$$\theta, \varphi \longrightarrow \delta\theta, \mu. \quad (\text{C.3.53})$$

Dans HOUDAYER, REESE et GOUPIL, 2022, on avait donc assimilé le modèle total à un processus gaussien hyperparamétré par θ et φ :

$$\mathbf{h}|\theta, \varphi, X \equiv \mathbf{f}|\varphi, X + \mathbf{g}(\theta) \sim \mathcal{N}(\mathbf{g}(\theta), K(\varphi)), \quad (\text{C.3.54})$$

$\mathbf{g}(\theta)$ constituant alors une moyenne à priori déterministe pour le processus total h . La distribution des mesures \mathbf{y} s'en déduit alors par ajout d'un bruit non biaisé \mathbf{n} :

$$\mathbf{y}|\theta, \varphi, X \sim \mathcal{N}(\mathbf{g}(\theta), K(\varphi) + \Sigma_n). \quad (\text{C.3.55})$$

En appliquant directement le résultat démontré en (C.3.46) à la variable $\mathbf{y} - \mathbf{g}(\theta)$ (de moyenne nulle), les paramètres optimaux du modèle, θ_* , ainsi que du processus gaussien, φ_* , se déduisent alors de la minimisation de la fonction suivante :

$$\theta_*, \varphi_* = \arg \min_{\theta, \varphi} \left\{ (\mathbf{y} - \mathbf{g}(\theta))^T (K(\varphi) + \Sigma_n)^{-1} (\mathbf{y} - \mathbf{g}(\theta)) + \ln |K(\varphi) + \Sigma_n| \right\}. \quad (\text{C.3.56})$$

Si la procédure décrite ci-dessus est tout à fait valide, elle n'est néanmoins pas optimale, dans le sens où il peut être coûteux d'effectuer une minimisation à la fois sur θ et φ suivant la taille de ces vecteurs. On va montrer ci-dessus comment il est possible de considérer h comme seulement hyperparamétré par φ , le vecteur θ jouant alors le rôle de simples paramètres et donc automatiquement ajusté par l'obtention de la loi à posteriori (cf. équation C.3.34). On commencera pour cela par supposer que le vecteur θ suit une loi normale :

$$\theta \sim \mathcal{N}(\mu_{\theta}, \Sigma_{\theta}). \quad (\text{C.3.57})$$

Le vecteur $\mathbf{g}|X = G(X)^T \theta$ est donc une transformation linéaire de θ et suit lui-même une loi normale :

$$\mathbf{g}|X \sim \mathcal{N}\left(G(X)^T \mu_{\theta}, G(X)^T \Sigma_{\theta} G(X)\right). \quad (\text{C.3.58})$$

Or, puisque la variable $\mathbf{g}|X$ n'est autre que l'application du modèle g sur l'ensemble des points composant X , on voit que le modèle g peut être lui-même considéré comme un processus gaussien¹⁷ :

$$g(\mathbf{x}) \sim \mathcal{GP}\left(G(\mathbf{x})^T \mu_{\theta}, G(\mathbf{x})^T \Sigma_{\theta} G(\mathbf{x}')\right). \quad (\text{C.3.59})$$

Formulé de cette manière, le processus total h , résultant de la somme de f et g , suit une loi légèrement différente de celle énoncée plus haut et, en particulier, ne dépend plus de θ (qui est maintenant une variable aléatoire) :

$$h(\mathbf{x}, \varphi) \equiv f(\mathbf{x}, \varphi) + g(\mathbf{x}) \sim \mathcal{GP}\left(G(\mathbf{x})^T \mu_{\theta}, k(\mathbf{x}, \mathbf{x}', \varphi) + G(\mathbf{x})^T \Sigma_{\theta} G(\mathbf{x}')\right) \quad (\text{C.3.60})$$

¹⁷. On notera que la fonction G , appliquée au vecteur \mathbf{x} et non plus à l'ensemble de vecteurs contenus dans X , aboutit à un vecteur et non plus une matrice. Les fonctions $G(\mathbf{x})^T \mu$ et $G(\mathbf{x})^T \Sigma_{\theta} G(\mathbf{x}')$ désignent ainsi bel et bien des fonctions scalaires (respectivement la moyenne et le kernel du processus g).

avec $k(X, X, \boldsymbol{\varphi}) = \mathbf{K}(\boldsymbol{\varphi})$. Évalué sur X , on écrira :

$$\mathbf{h}|\boldsymbol{\varphi}, X \equiv \mathbf{f}|\boldsymbol{\varphi}, X + \mathbf{g}|X \sim \mathcal{N}(\mathbf{G}^\top \boldsymbol{\mu}_\theta, \mathbf{K}(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G}), \quad (\text{C.3.61})$$

$$\text{et } \mathbf{y}|\boldsymbol{\varphi}, X = \mathbf{h}|\boldsymbol{\varphi}, X + \mathbf{n}|X \sim \mathcal{N}(\mathbf{G}^\top \boldsymbol{\mu}_\theta, \mathbf{K}(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G} + \boldsymbol{\Sigma}_n) \quad (\text{C.3.62})$$

après avoir abrégé la matrice $\mathbf{G}(X) \rightarrow \mathbf{G}$, et l'expression (C.3.61) peut-être comparée avec l'équation (C.3.54).

Après observation des données \mathcal{D} , l'équation (C.3.24) peut être utilisée pour définir la moyenne du processus total h en X_* à posteriori :

$$\begin{aligned} \overline{\mathbf{h}_*|\mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}} &= \mathbf{G}_*^\top \boldsymbol{\mu}_\theta - \left(\mathbf{K}_*(\boldsymbol{\varphi}) + \mathbf{G}_*^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \right)^\top \left(\mathbf{K}(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G} + \boldsymbol{\Sigma}_n \right)^{-1} \left(\mathbf{G}^\top \boldsymbol{\mu}_\theta - \mathbf{y} \right) \\ &= \mathbf{G}_*^\top \boldsymbol{\mu}_\theta + \left(\mathbf{K}_*(\boldsymbol{\varphi})^\top + \mathbf{G}_*^\top \boldsymbol{\Sigma}_\theta \mathbf{G} \right) \left(\mathbf{K}_y(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G} \right)^{-1} \left(\mathbf{y} - \mathbf{G}^\top \boldsymbol{\mu}_\theta \right) \end{aligned} \quad (\text{C.3.63})$$

avec $\mathbf{G}_* = \mathbf{G}(X_*)$, $\mathbf{K}_*(\boldsymbol{\varphi}) = k(X, X_*, \boldsymbol{\varphi})$ et $\mathbf{K}_y(\boldsymbol{\varphi}) = \mathbf{K}(\boldsymbol{\varphi}) + \boldsymbol{\Sigma}_n$, la matrice de covariance à priori de $\mathbf{f}|X, \boldsymbol{\varphi} + \mathbf{n}|X$. L'expression obtenue peut être réarrangée en utilisant la formule de Woodbury (WOODBURY et STATISTICS, 1950) :

$$\left(\mathbf{K}_y(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G} \right)^{-1} = \mathbf{K}_y(\boldsymbol{\varphi})^{-1} - \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1}. \quad (\text{C.3.64})$$

L'injection de cette expression dans l'équation (C.3.63) fait apparaître de nombreux termes :

$$\begin{aligned} \overline{\mathbf{h}_*|\mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}} &= \mathbf{G}_*^\top \boldsymbol{\mu}_\theta + \overline{\mathbf{f}_*|\mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}} - \mathbf{K}_*(\boldsymbol{\varphi})^\top \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \boldsymbol{\mu}_\theta \\ &\quad - \mathbf{K}_*(\boldsymbol{\varphi})^\top \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{y} - \mathbf{G}^\top \boldsymbol{\mu}_\theta \right) \\ &\quad + \mathbf{G}_*^\top \boldsymbol{\Sigma}_\theta \mathbf{G} \left[\mathbf{K}_y(\boldsymbol{\varphi})^{-1} - \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \right] \left(\mathbf{y} - \mathbf{G}^\top \boldsymbol{\mu}_\theta \right) \end{aligned} \quad (\text{C.3.65})$$

dont $\overline{\mathbf{f}_*|\mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}} = \mathbf{K}_*(\boldsymbol{\varphi})^\top \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{y}$, la moyenne de f à posteriori. En factorisant les termes par \mathbf{G}_*^\top ou $\mathbf{K}_*(\boldsymbol{\varphi})^\top \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top$:

$$\begin{aligned} \overline{\mathbf{h}_*|\mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}} &= \overline{\mathbf{f}_*|\mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}} - \mathbf{K}_*(\boldsymbol{\varphi})^\top \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \overbrace{\left[\boldsymbol{\mu}_\theta + \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{y} - \mathbf{G}^\top \boldsymbol{\mu}_\theta \right) \right]}^{\text{A}} \\ &\quad + \mathbf{G}_*^\top \underbrace{\left[\boldsymbol{\mu}_\theta + \boldsymbol{\Sigma}_\theta \mathbf{G} \left[\mathbf{K}_y(\boldsymbol{\varphi})^{-1} - \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \right] \left(\mathbf{y} - \mathbf{G}^\top \boldsymbol{\mu}_\theta \right) \right]}_{\text{B}} \end{aligned} \quad (\text{C.3.66})$$

Or le terme A se développe comme :

$$\begin{aligned} \text{A} &= \boldsymbol{\mu}_\theta + \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \left(\mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{y} - \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \boldsymbol{\mu}_\theta \right) \\ &= \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \left(\boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{\mu}_\theta + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{y} \right) \\ &\equiv \bar{\boldsymbol{\theta}}(\boldsymbol{\varphi}) \end{aligned} \quad (\text{C.3.67})$$

où $\bar{\boldsymbol{\theta}}(\boldsymbol{\varphi})$ désigne la moyenne du produit des gaussiennes $\mathcal{G}(\boldsymbol{\theta}, \boldsymbol{\mu}_\theta, \boldsymbol{\Sigma}_\theta)$ et $\mathcal{G}(\boldsymbol{\theta}, (\mathbf{G}^\top)^{-1} \mathbf{y}, (\mathbf{G}^\top)^{-1} \mathbf{K}_y(\boldsymbol{\varphi}) \mathbf{G}^{-1})$ (cf. équation C.3.42). Elle peut s'interpréter comme une moyenne pondérée entre à priori $\boldsymbol{\mu}_\theta$ et influence des données $(\mathbf{G}^\top)^{-1} \mathbf{y}$ et donne la valeur des paramètres $\boldsymbol{\theta}$ la plus plausible selon le processus h lors de l'ajustement sur les données. Cette grandeur possède donc une expression analytique en fonction de $\boldsymbol{\varphi}$ et ne résulte plus de la minimisation numérique (C.3.56). D'autre

part :

$$\begin{aligned}
\mathbf{B} &= \boldsymbol{\mu}_\theta + \boldsymbol{\Sigma}_\theta \left[\mathbf{I} - \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \right] \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{y} - \mathbf{G}^\top \boldsymbol{\mu}_\theta \right) \\
&= \boldsymbol{\mu}_\theta + \boldsymbol{\Sigma}_\theta \left[\boldsymbol{\Sigma}_\theta^{-1} \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \right] \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{y} - \mathbf{G}^\top \boldsymbol{\mu}_\theta \right) \\
&= \mathbf{A} \\
&\equiv \bar{\boldsymbol{\theta}}(\boldsymbol{\varphi}).
\end{aligned} \tag{C.3.68}$$

On obtient donc finalement :

$$\overline{\mathbf{h}_* | \mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}} = \overline{\mathbf{f}_* | \mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}} + \tilde{\mathbf{G}}_*(\boldsymbol{\varphi})^\top \bar{\boldsymbol{\theta}}(\boldsymbol{\varphi}), \tag{C.3.69}$$

où l'on a défini :

$$\tilde{\mathbf{G}}_*(\boldsymbol{\varphi}) = \mathbf{G}_* - \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{K}_*(\boldsymbol{\varphi}) \tag{C.3.70}$$

qui résulte de l'influence sur la base de fonctions G du kernel k définissant le processus f . L'équation (C.3.69) nous fournit ainsi un ajustement analytique du modèle h sur les données sans qu'il ne soit nécessaire d'ajuster le vecteur $\boldsymbol{\theta}$.

On pourra, de la même façon, déterminer la covariance du processus h par rapport à l'estimation (C.3.69) en utilisant l'équation (C.3.24) :

$$\begin{aligned}
\text{Cov} [\mathbf{h}_* | \mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}] &= \mathbf{K}_{**}(\boldsymbol{\varphi}) + \mathbf{G}_*^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \\
&\quad - \left(\mathbf{K}_*(\boldsymbol{\varphi})^\top + \mathbf{G}_*^\top \boldsymbol{\Sigma}_\theta \mathbf{G} \right) \left(\mathbf{K}_y(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G} \right)^{-1} \left(\mathbf{K}_*(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \right)
\end{aligned} \tag{C.3.71}$$

avec $\mathbf{K}_{**}(\boldsymbol{\varphi}) = k(\mathcal{X}_*, \mathcal{X}_*, \boldsymbol{\varphi})$. En utilisant de nouveau la formule de Woodbury (C.3.64) :

$$\begin{aligned}
\text{Cov} [\mathbf{h}_* | \mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}] &= \text{Cov} [\mathbf{f}_* | \mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}] + \mathbf{G}_*^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* - \mathbf{K}_*(\boldsymbol{\varphi})^\top \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \\
&\quad - \mathbf{G}_*^\top \boldsymbol{\Sigma}_\theta \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{K}_*(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \right) \\
&\quad + \left(\mathbf{K}_*(\boldsymbol{\varphi})^\top + \mathbf{G}_*^\top \boldsymbol{\Sigma}_\theta \mathbf{G} \right) \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \\
&\quad \times \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{K}_*(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \right)
\end{aligned} \tag{C.3.72}$$

où l'on aura reconnu $\text{Cov} [\mathbf{f}_* | \mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}] = \mathbf{K}_{**}(\boldsymbol{\varphi}) + \mathbf{K}_*(\boldsymbol{\varphi})^\top \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{K}_*(\boldsymbol{\varphi})$. On a cette fois des termes se factorisant en \mathbf{G}_*^\top ou en $\mathbf{K}_*(\boldsymbol{\varphi})^\top \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top = \left(\mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{K}_*(\boldsymbol{\varphi}) \right)^\top$:

$$\begin{aligned}
\text{Cov} [\mathbf{h}_* | \mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}] &= \text{Cov} [\mathbf{f}_* | \mathcal{X}_*, \mathcal{D}, \boldsymbol{\varphi}] \\
&\quad + \mathbf{G}_*^\top \left[\boldsymbol{\Sigma}_\theta \mathbf{G}_* - \boldsymbol{\Sigma}_\theta \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{K}_*(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \right) \right. \\
&\quad \quad \left. + \boldsymbol{\Sigma}_\theta \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{K}_*(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \right) \right] \\
&\quad - \left(\mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{K}_*(\boldsymbol{\varphi}) \right)^\top \left[\boldsymbol{\Sigma}_\theta \mathbf{G}_* - \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \mathbf{G}^\top \right)^{-1} \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi})^{-1} \left(\mathbf{K}_*(\boldsymbol{\varphi}) + \mathbf{G}^\top \boldsymbol{\Sigma}_\theta \mathbf{G}_* \right) \right].
\end{aligned} \tag{C.3.73}$$

Du côté droit maintenant, les termes se factorisent également par G_* ou $G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi})$:

$$\begin{aligned}
\text{Cov} [\mathbf{h}_* | X_*, \mathcal{D}, \boldsymbol{\varphi}] &= \text{Cov} [\mathbf{f}_* | X_*, \mathcal{D}, \boldsymbol{\varphi}] \\
&+ G_*^\top \left[\Sigma_\theta - \Sigma_\theta G K_y(\boldsymbol{\varphi})^{-1} G^\top \Sigma_\theta \right. \\
&\quad \left. + \Sigma_\theta G K_y(\boldsymbol{\varphi})^{-1} G^\top \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G K_y(\boldsymbol{\varphi})^{-1} G^\top \Sigma_\theta \right] G_* \\
&- G_*^\top \left[\Sigma_\theta - \Sigma_\theta G K_y(\boldsymbol{\varphi})^{-1} G^\top \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} \right] G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi}) \\
&- \left(G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi}) \right)^\top \left[\Sigma_\theta - \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G K_y(\boldsymbol{\varphi})^{-1} G^\top \Sigma_\theta \right] G_* \\
&+ \left(G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi}) \right)^\top \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi}).
\end{aligned} \tag{C.3.74}$$

Intéressons-nous maintenant aux trois termes entre crochets. Le dernier se simplifie de la façon suivante :

$$\begin{aligned}
\Sigma_\theta - \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G K_y(\boldsymbol{\varphi})^{-1} G^\top \Sigma_\theta &= \left[I - \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G K_y(\boldsymbol{\varphi})^{-1} G^\top \right] \Sigma_\theta \\
&= \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} \Sigma_\theta^{-1} \Sigma_\theta \\
&= \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1}.
\end{aligned} \tag{C.3.75}$$

Le second n'est que le transposé du terme que l'on vient calculer et lui est donc égal puisqu'il est symétrique. Enfin, on remarquera que le dernier est égal à :

$$\Sigma_\theta - \Sigma_\theta G K_y(\boldsymbol{\varphi})^{-1} G^\top \left[\dots \right]$$

où le crochet désigne le second terme. Or, on a montré plus haut qu'il vaut $\left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1}$ et les deux premiers termes sont donc égaux. On vient donc de montrer que :

$$\begin{aligned}
\text{Cov} [\mathbf{h}_* | X_*, \mathcal{D}, \boldsymbol{\varphi}] &= \text{Cov} [\mathbf{f}_* | X_*, \mathcal{D}, \boldsymbol{\varphi}] \\
&+ G_*^\top \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G_* \\
&- G_*^\top \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi}) \\
&- \left(G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi}) \right)^\top \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G_* \\
&+ \left(G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi}) \right)^\top \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} G K_y(\boldsymbol{\varphi})^{-1} K_*(\boldsymbol{\varphi}).
\end{aligned} \tag{C.3.76}$$

En reconnaissant une forme quadratique en $\tilde{G}_*(\boldsymbol{\varphi})$ (cf. définition C.3.70), on écrira finalement :

$$\text{Cov} [\mathbf{h}_* | X_*, \mathcal{D}, \boldsymbol{\varphi}] = \text{Cov} [\mathbf{f}_* | X_*, \mathcal{D}, \boldsymbol{\varphi}] + \tilde{G}_*(\boldsymbol{\varphi})^\top \left(\Sigma_\theta^{-1} + G K_y(\boldsymbol{\varphi})^{-1} G^\top \right)^{-1} \tilde{G}_*(\boldsymbol{\varphi}). \tag{C.3.77}$$

On voit donc que la covariance du processus total est la somme de la covariance de f et d'une incertitude apportée par le modèle linéaire qui sera toujours positive.

La recherche des paramètres $\boldsymbol{\varphi}$ les plus plausibles peut alors être effectuée numériquement. Puisque la loi de $\mathbf{y}|\boldsymbol{\varphi}, \mathbf{X}$ est connue (cf. équation C.3.62), il s'agira de minimiser la fonction suivante :

$$\boldsymbol{\varphi}_* = \arg \min_{\boldsymbol{\varphi}} \left\{ \left(\mathbf{y} - \mathbf{G}^T \boldsymbol{\mu}_\theta \right)^T \left(\mathbf{K}_y(\boldsymbol{\varphi}) + \mathbf{G}^T \boldsymbol{\Sigma}_\theta \mathbf{G} \right)^{-1} \left(\mathbf{y} - \mathbf{G}^T \boldsymbol{\mu}_\theta \right) + \ln \left| \mathbf{K}_y(\boldsymbol{\varphi}) + \mathbf{G}^T \boldsymbol{\Sigma}_\theta \mathbf{G} \right| \right\}. \quad (\text{C.3.78})$$

Par comparaison avec l'équation (C.3.56), la minimisation sur $\boldsymbol{\theta}$ et $\boldsymbol{\varphi}$ s'est transformée en une minimisation sur $\boldsymbol{\varphi}$ seulement. Concrètement, l'application de cette méthode dans [HOUDAYER, REESE et GOUPIL, 2022](#) réduirait l'étude des hyperparamètres à celle d'un espace de dimension 2 (plutôt que 5). On ajoutera que le gain sur la minimisation des hyperparamètres peut être encore bien plus conséquent, les bases de fonctions contenant généralement plus d'éléments que le processus gaussien de paramètres. Enfin, l'estimation la plus plausible de $\boldsymbol{\theta}$ est donnée par l'application de la relation $\bar{\boldsymbol{\theta}}(\boldsymbol{\varphi})$ en $\boldsymbol{\varphi}_*$:

$$\boldsymbol{\theta}_* = \bar{\boldsymbol{\theta}}(\boldsymbol{\varphi}_*) = \left(\boldsymbol{\Sigma}_\theta^{-1} + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi}_*)^{-1} \mathbf{G}^T \right)^{-1} \left(\boldsymbol{\Sigma}_\theta^{-1} \boldsymbol{\mu}_\theta + \mathbf{G} \mathbf{K}_y(\boldsymbol{\varphi}_*)^{-1} \mathbf{y} \right). \quad (\text{C.3.79})$$

Bibliographie

1. ABE, R. (août 1959). « Giant Cluster Expansion Theory and Its Application to High Temperature Plasma ». In : *Progress of Theoretical Physics* 22.2, p. 213-226. DOI : [10.1143/PTP.22.213](https://doi.org/10.1143/PTP.22.213).
2. ABRAMOWITZ, Milton et Irene A. STEGUN (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. ninth Dover printing, tenth GPO printing. New York : Dover.
3. AIGRAIN, S., H. PARVIAINEN et B. J. S. POPE (juill. 2016). « K2SC : flexible systematics correction and detrending of K2 light curves using Gaussian process regression ». In : *Monthly Notices of the Royal Astronomical Society* 459.3, p. 2408-2419. DOI : [10.1093/mnras/stw706](https://doi.org/10.1093/mnras/stw706).
4. AIZENMAN, M. L. et P. SMEYERS (mai 1977). « An Analysis of the Linear, Adiabatic Oscillations of a Star in Terms of Potential Fields ». In : *Astrophysics and Space Science* 48.1, p. 123-136. DOI : [10.1007/BF00643044](https://doi.org/10.1007/BF00643044).
5. ALLER, L. (jan. 1961). « Solar and stellar abundances of the elements ». In : *Physics and Chemistry of the Earth* 4, p. 1-24. DOI : [10.1016/0079-1946\(61\)90002-7](https://doi.org/10.1016/0079-1946(61)90002-7).
6. AMARD, L. et al. (nov. 2019). « First grids of low-mass stellar models and isochrones with self-consistent treatment of rotation. From 0.2 to 1.5 M_{\odot} at seven metallicities from PMS to TAMS ». In : *Astronomy & Astrophysics* 631, A77, A77. DOI : [10.1051/0004-6361/201935160](https://doi.org/10.1051/0004-6361/201935160).
7. ANGULO, C. et al. (août 1999). « A compilation of charged-particle induced thermonuclear reaction rates ». In : *Nuclear Physics A* 656.1, p. 3-183. DOI : [10.1016/S0375-9474\(99\)00030-5](https://doi.org/10.1016/S0375-9474(99)00030-5).
8. ANGUS, Ruth et al. (fév. 2018). « Inferring probabilistic stellar rotation periods using Gaussian processes ». In : *Monthly Notices of the Royal Astronomical Society* 474.2, p. 2094-2108. DOI : [10.1093/mnras/stx2109](https://doi.org/10.1093/mnras/stx2109).
9. ANTIA, H. M., Sarbani BASU et S. M. CHITRE (août 1998). « Solar internal rotation rate and the latitudinal variation of the tachocline ». In : *Monthly Notices of the Royal Astronomical Society* 298.2, p. 543-556. DOI : [10.1046/j.1365-8711.1998.01635.x](https://doi.org/10.1046/j.1365-8711.1998.01635.x).
10. ARIS, R. (1990). *Vectors, Tensors and the Basic Equations of Fluid Mechanics*. Dover Books on Mathematics. Dover Publications.
11. ASPLUND, Martin et al. (sept. 2009). « The Chemical Composition of the Sun ». In : *Annual Review of Astronomy and Astrophysics* 47.1, p. 481-522. DOI : [10.1146/annurev.astro.46.060407.145222](https://doi.org/10.1146/annurev.astro.46.060407.145222).
12. BACKUS, George et Freeman GILBERT (oct. 1968). « The Resolving Power of Gross Earth Data ». In : *Geophysical Journal* 16.2, p. 169-205. DOI : [10.1111/j.1365-246X.1968.tb00216.x](https://doi.org/10.1111/j.1365-246X.1968.tb00216.x).
13. BAGLIN, A. et al. (jan. 2002). « COROT : asteroseismology and planet finding ». In : *Stellar Structure and Habitable Planet Finding*. Sous la dir. de B. BATTRICK et al. T. 485. ESA Special Publication, p. 17-24.
14. BAGLIN, Annie et al. (fév. 2009). « CoRoT : Description of the Mission and Early Results ». In : *Transiting Planets*. Sous la dir. de Frédéric PONT, Dimitar SASSELOV et Matthew J. HOLMAN. T. 253, p. 71-81. DOI : [10.1017/S1743921308026252](https://doi.org/10.1017/S1743921308026252).
15. BALLARD, Sarah et al. (juill. 2014). « Kepler-93b : A Terrestrial World Measured to within 120 km, and a Test Case for a New Spitzer Observing Mode ». In : *The Astrophysical Journal* 790.1, 12, p. 12. DOI : [10.1088/0004-637X/790/1/12](https://doi.org/10.1088/0004-637X/790/1/12).

16. BALLOT, J., S. TURCK-CHIÈZE et R. A. GARCÍA (sept. 2004). « Seismic extraction of the convective extent in solar-like stars. The observational point of view ». In : *Astronomy and Astrophysics* 423, p. 1051-1061. DOI : [10.1051/0004-6361:20035898](https://doi.org/10.1051/0004-6361:20035898).
17. BASU, S. et H. M. ANTIA (août 1994a). « Effects of Diffusion on the Extent of Overshoot Below the Solar Convection Zone ». In : *Monthly Notices of the Royal Astronomical Society* 269, p. 1137. DOI : [10.1093/mnras/269.4.1137](https://doi.org/10.1093/mnras/269.4.1137).
18. BASU, Sarbani et H. M. ANTIA (juin 1994b). « Seismology of the solar convection zone ». In : *Journal of Astrophysics and Astronomy* 15.2, p. 143-156. DOI : [10.1007/BF03012174](https://doi.org/10.1007/BF03012174).
19. — (mai 1997). « Seismic measurement of the depth of the solar convection zone ». In : *Monthly Notices of the Royal Astronomical Society* 287.1, p. 189-198. DOI : [10.1093/mnras/287.1.189](https://doi.org/10.1093/mnras/287.1.189).
20. — (mai 2004). « Constraining Solar Abundances Using Helioseismology ». In : *ApJ Letters* 606.1, p. L85-L88. DOI : [10.1086/421110](https://doi.org/10.1086/421110).
21. — (mars 2008). « Helioseismology and solar abundances ». In : *Physics Reports* 457.5-6, p. 217-283. DOI : [10.1016/j.physrep.2007.12.002](https://doi.org/10.1016/j.physrep.2007.12.002).
22. BASU, Sarbani, H. M. ANTIA et D. NARASIMHA (mars 1994). « Helioseismic measurement of the extent of overshoot below the solar convection zone ». In : *Monthly Notices of the Royal Astronomical Society* 267, p. 209. DOI : [10.1093/mnras/267.1.209](https://doi.org/10.1093/mnras/267.1.209).
23. BASU, Sarbani et al. (déc. 1997). « Solar internal sound speed as inferred from combined BiSON and LOWL oscillation frequencies ». In : *Monthly Notices of the Royal Astronomical Society* 292.2, p. 243-251. DOI : [10.1093/mnras/292.2.243](https://doi.org/10.1093/mnras/292.2.243).
24. BASU, Sarbani et al. (mai 2004). « Asteroseismic determination of helium abundance in stellar envelopes ». In : *Monthly Notices of the Royal Astronomical Society* 350.1, p. 277-286. DOI : [10.1111/j.1365-2966.2004.07644.x](https://doi.org/10.1111/j.1365-2966.2004.07644.x).
25. BELKACEM, K. et al. (juin 2011). « The underlying physical meaning of the $v_{max} - v_c$ relation ». In : *Astronomy and Astrophysics* 530, A142, A142. DOI : [10.1051/0004-6361/201116490](https://doi.org/10.1051/0004-6361/201116490).
26. BENDER, C. M. et S. A. ORSZAG (1978). *Advanced Mathematical Methods for Scientists and Engineers*.
27. BERGER, James O. (1985). *Statistical decision theory and Bayesian analysis*. New York : Springer-Verlag.
28. BIRD, G. A. (1994). *Molecular Gas Dynamics And The Direct Simulation Of Gas Flows*.
29. BOLTZMANN, L. (1872). *Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen*. k. und k. Hof- und Staatsdr.
30. BORN, M. et H. S. GREEN (déc. 1946). « A General Kinetic Theory of Liquids. I. The Molecular Distribution Functions ». In : *Proceedings of the Royal Society of London Series A* 188.1012, p. 10-18. DOI : [10.1098/rspa.1946.0093](https://doi.org/10.1098/rspa.1946.0093).
31. BOSSINI, D. et al. (août 2017). « Kepler red-clump stars in the field and in open clusters : constraints on core mixing ». In : *Monthly Notices of the Royal Astronomical Society* 469.4, p. 4718-4725. DOI : [10.1093/mnras/stx1135](https://doi.org/10.1093/mnras/stx1135).
32. BOSSINI, Diego et al. (nov. 2015). « Uncertainties on near-core mixing in red-clump stars : effects on the period spacing and on the luminosity of the AGB bump ». In : *Monthly Notices of the Royal Astronomical Society* 453.3, p. 2290-2301. DOI : [10.1093/mnras/stv1738](https://doi.org/10.1093/mnras/stv1738).
33. BRAUS, Michael J. (2019). « 'The theory of electrolytes. I. Freezing point depression and related phenomena' (Debye & Huckel, 1923) ». In.
34. BRILLOUIN, L. (1926). « Remarques sur la mécanique ondulatoire ». In : *J. Phys. Radium* 7.12, p. 353-368. DOI : [10.1051/jphysrad:01926007012035300](https://doi.org/10.1051/jphysrad:01926007012035300).
35. BROOMHALL, A. M. et al. (mai 2014). « Prospects for asteroseismic inference on the envelope helium abundance in red giant stars ». In : *Monthly Notices of the Royal Astronomical Society* 440.2, p. 1828-1843. DOI : [10.1093/mnras/stu393](https://doi.org/10.1093/mnras/stu393).

36. BROWN, Timothy M. et Ronald L. GILLILAND (jan. 1994). « Asteroseismology ». In : *Annual Review of Astronomy and Astrophysics* 32, p. 37-82. DOI : [10.1146/annurev.aa.32.090194.000345](https://doi.org/10.1146/annurev.aa.32.090194.000345).
37. BROWN, Timothy M. et al. (juin 1994). « The Effectiveness of Oscillation Frequencies in Constraining Stellar Model Parameters ». In : *The Astrophysical Journal* 427, p. 1013. DOI : [10.1086/174208](https://doi.org/10.1086/174208).
38. BUCHLER, J.Robert (1979). « Radiation hydrodynamics in the fluid frame ». In : *Journal of Quantitative Spectroscopy and Radiative Transfer* 22.3, p. 293-300. DOI : [https://doi.org/10.1016/0022-4073\(79\)90119-5](https://doi.org/10.1016/0022-4073(79)90119-5).
39. BULDGEN, G., D. R. REESE et M. A. DUPRET (nov. 2015). « Using seismic inversions to obtain an indicator of internal mixing processes in main-sequence solar-like stars ». In : *Astronomy and Astrophysics* 583, A62, A62. DOI : [10.1051/0004-6361/201526390](https://doi.org/10.1051/0004-6361/201526390).
40. BULDGEN, G. et al. (nov. 2017a). « Determining the metallicity of the solar envelope using seismic inversion techniques ». In : *Monthly Notices of the Royal Astronomical Society* 472.1, p. 751-764. DOI : [10.1093/mnras/stx2057](https://doi.org/10.1093/mnras/stx2057).
41. BULDGEN, G. et al. (nov. 2017b). « Seismic inversion of the solar entropy. A case for improving the standard solar model ». In : *Astronomy and Astrophysics* 607, A58, A58. DOI : [10.1051/0004-6361/201731354](https://doi.org/10.1051/0004-6361/201731354).
42. BULDGEN, G. et al. (jan. 2019a). « Combining multiple structural inversions to constrain the solar modelling problem ». In : *Astronomy and Astrophysics* 621, A33, A33. DOI : [10.1051/0004-6361/201833971](https://doi.org/10.1051/0004-6361/201833971).
43. BULDGEN, Gaël et al. (jan. 2019b). « Mean density inversions for red giants and red clump stars ». In : *Monthly Notices of the Royal Astronomical Society* 482.2, p. 2305-2319. DOI : [10.1093/mnras/sty2346](https://doi.org/10.1093/mnras/sty2346).
44. CANUTO, V. M. (juin 1997). « Compressible Turbulence ». In : *The Astrophysical Journal* 482.2, p. 827-851. DOI : [10.1086/304175](https://doi.org/10.1086/304175).
45. CANUTO, V. M., I. GOLDMAN et I. MAZZITELLI (déc. 1996). « Stellar Turbulent Convection : A Self-consistent Model ». In : *Astrophysical Journal* 473, p. 550. DOI : [10.1086/178166](https://doi.org/10.1086/178166).
46. CANUTO, V. M. et I. MAZZITELLI (mars 1991). « Stellar Turbulent Convection : A New Model and Applications ». In : *Astrophysical Journal* 370, p. 295. DOI : [10.1086/169815](https://doi.org/10.1086/169815).
47. CASTOR, John I. (déc. 1972). « Radiative Transfer in Spherically Symmetric Flows ». In : *Astrophysical Journal* 178, p. 779-792. DOI : [10.1086/151834](https://doi.org/10.1086/151834).
48. CHABOYER, B. et J. P. ZAHN (jan. 1992). « Effect of horizontal turbulent diffusion on transport by meridional circulation. » In : *Astronomy & Astrophysics* 253, p. 173-177.
49. CHANDRASEKHAR, S. (fév. 1964). « A General Variational Principle Governing the Radial and the Non-Radial Oscillations of Gaseous Masses. » In : *Astrophysical Journal* 139, p. 664. DOI : [10.1086/147792](https://doi.org/10.1086/147792).
50. CHANDRASEKHAR, S. et P. C. KENDALL (sept. 1957). « On Force-Free Magnetic Fields. » In : *Astrophysical Journal* 126, p. 457. DOI : [10.1086/146413](https://doi.org/10.1086/146413).
51. CHANDRASEKHAR, Subrahmanyan (1939). *An introduction to the study of stellar structure*.
52. CHAPLIN, William J. et al. (sept. 1996). « BiSON Performance ». In : *Solar Physics* 168.1, p. 1-18. DOI : [10.1007/BF00145821](https://doi.org/10.1007/BF00145821).
53. CHAPMAN, Sydeny et T. G. COWLING (1970). *The mathematical theory of non-uniform gases. an account of the kinetic theory of viscosity, thermal conduction and diffusion in gases*.
54. CHARBONNEL, C. et J. P. ZAHN (mai 2007). « Thermohaline mixing : a physical mechanism governing the photospheric composition of low-mass giants ». In : *Astronomy and Astrophysics* 467.1, p. L15-L18. DOI : [10.1051/0004-6361:20077274](https://doi.org/10.1051/0004-6361:20077274).

55. CHRISTENSEN-DALSGAARD, J. et M. J. THOMPSON (jan. 1997). « On solar p-mode frequency shifts caused by near-surface model changes ». In : *Monthly Notices of the Royal Astronomical Society* 284.3, p. 527-540. DOI : [10.1093/mnras/284.3.527](https://doi.org/10.1093/mnras/284.3.527).
56. CHRISTENSEN-DALSGAARD, J. et al. (mai 1985). « Speed of sound in the solar interior ». In : *Nature* 315.6018, p. 378-382. DOI : [10.1038/315378a0](https://doi.org/10.1038/315378a0).
57. CHRISTENSEN-DALSGAARD, J. et al. (mai 1996). « The Current State of Solar Modeling ». In : *Science* 272.5266, p. 1286-1292. DOI : [10.1126/science.272.5266.1286](https://doi.org/10.1126/science.272.5266.1286).
58. CHRISTENSEN-DALSGAARD, J. et al. (juin 2011). « A more realistic representation of overshoot at the base of the solar convective envelope as seen by helioseismology ». In : *Monthly Notices of the Royal Astronomical Society* 414.2, p. 1158-1174. DOI : [10.1111/j.1365-2966.2011.18460.x](https://doi.org/10.1111/j.1365-2966.2011.18460.x).
59. CHRISTENSEN-DALSGAARD, Jørgen (nov. 2002). « Helioseismology ». In : *Reviews of Modern Physics* 74.4, p. 1073-1129. DOI : [10.1103/RevModPhys.74.1073](https://doi.org/10.1103/RevModPhys.74.1073).
60. CHRISTENSEN-DALSGAARD, Jørgen, Mario J. P. F. G. MONTEIRO et Michael J. THOMPSON (sept. 1995). « Helioseismic estimation of convective overshoot in the Sun ». In : *Monthly Notices of the Royal Astronomical Society* 276, p. 283-292. DOI : [10.1093/mnras/276.1.283](https://doi.org/10.1093/mnras/276.1.283).
61. COWLING, T. G. (jan. 1941). « The non-radial oscillations of polytropic stars ». In : *Monthly Notices of the Royal Astronomical Society* 101, p. 367. DOI : [10.1093/mnras/101.8.367](https://doi.org/10.1093/mnras/101.8.367).
62. COX, J. P. et R. T. GIULI (1968). *Principles of stellar structure*.
63. CUNHA, M. S. et al. (juin 2015). « Structural Glitches near the Cores of Red Giants Revealed by Oscillations in g-mode Period Spacings from Stellar Models ». In : *The Astrophysical Journal* 805.2, 127, p. 127. DOI : [10.1088/0004-637X/805/2/127](https://doi.org/10.1088/0004-637X/805/2/127).
64. DAEPFEN, Werner et al. (sept. 1988). « The Equation of State for Stellar Envelopes. III. Thermodynamic Quantities ». In : *The Astrophysical Journal* 332, p. 261. DOI : [10.1086/166650](https://doi.org/10.1086/166650).
65. DAVIES, G. R. et A. MIGLIO (sept. 2016). « Asteroseismology of red giants : From analysing light curves to estimating ages ». In : *Astronomische Nachrichten* 337.8-9, p. 774. DOI : [10.1002/asna.201612371](https://doi.org/10.1002/asna.201612371).
66. DEAL, M. et al. (oct. 2018). « Impacts of radiative accelerations on solar-like oscillating main-sequence stars ». In : *Astronomy & Astrophysics* 618, A10, A10. DOI : [10.1051/0004-6361/201833361](https://doi.org/10.1051/0004-6361/201833361).
67. EDDINGTON, A. S. (mars 1925). « Circulating currents in rotating stars ». In : *The Observatory* 48, p. 73-75.
68. — (1926). *The Internal Constitution of the Stars*.
69. EKSTRÖM, S. et al. (jan. 2012). « Grids of stellar models with rotation. I. Models from 0.8 to 120 M_{\odot} at solar metallicity ($Z = 0.014$) ». In : *Astronomy & Astrophysics* 537, A146, A146. DOI : [10.1051/0004-6361/201117751](https://doi.org/10.1051/0004-6361/201117751).
70. ENDAL, A. S. et S. SOFIA (nov. 1976). « The evolution of rotating stars. I. Method and exploratory calculations for a 7 M_{\odot} star ». In : *Astrophysical Journal* 210, p. 184-198. DOI : [10.1086/154817](https://doi.org/10.1086/154817).
71. — (fév. 1978). « The evolution of rotating stars. II. Calculations with time-dependent redistribution of angular momentum for 7 and 10 M_{\odot} stars ». In : *Astrophysical Journal* 220, p. 279-290. DOI : [10.1086/155904](https://doi.org/10.1086/155904).
72. FARNIR, M. et al. (fév. 2019). « Comprehensive stellar seismic analysis. New method exploiting the glitches information in solar-like pulsators ». In : *Astronomy and Astrophysics* 622, A98, A98. DOI : [10.1051/0004-6361/201834044](https://doi.org/10.1051/0004-6361/201834044).
73. FARNIR, M. et al. (déc. 2020). « Thorough characterisation of the 16 Cygni system. I. Forward seismic modelling with WhoSGLAd ». In : *Astronomy and Astrophysics* 644, A37, A37. DOI : [10.1051/0004-6361/202038522](https://doi.org/10.1051/0004-6361/202038522).

74. FINETTI, Bruno de (1992). « Foresight : Its Logical Laws, Its Subjective Sources ». In : *Breakthroughs in Statistics : Foundations and Basic Theory*. Sous la dir. de Samuel KOTZ et Norman L. JOHNSON. New York, NY : Springer New York, p. 134-174. DOI : [10.1007/978-1-4612-0919-5_10](https://doi.org/10.1007/978-1-4612-0919-5_10).
75. GABRIEL, A. H. et al. (déc. 1995). « Global Oscillations at Low Frequency from the SOHO Mission (GOLF) ». In : *Solar Physics* 162.1-2, p. 61-99. DOI : [10.1007/BF00733427](https://doi.org/10.1007/BF00733427).
76. GABRIEL, A. H. et al. (oct. 1997). « Performance and Early Results from the GOLF Instrument Flown on the SOHO Mission ». In : *Solar Physics* 175.2, p. 207-226. DOI : [10.1023/A:1004911408285](https://doi.org/10.1023/A:1004911408285).
77. GILLILAND, Ronald L. et al. (fév. 2010). « Kepler Asteroseismology Program : Introduction and First Results ». In : *Publications of the ASP* 122.888, p. 131. DOI : [10.1086/650399](https://doi.org/10.1086/650399).
78. GOUGH, D. O. (1990). « Comments on Helioseismic Inference ». In : *Progress of Seismology of the Sun and Stars*. Sous la dir. d'Yoji OSAKI et Hiromoto SHIBAHASHI. T. 367, p. 283. DOI : [10.1007/3-540-53091-6](https://doi.org/10.1007/3-540-53091-6).
79. — (jan. 1993). « Linear adiabatic stellar pulsation. » In : *Astrophysical Fluid Dynamics - Les Houches 1987*, p. 399-560.
80. — (jan. 2002). « Helioseismology : some current issues concerning model calibration ». In : *Stellar Structure and Habitable Planet Finding*. Sous la dir. de B. BATTRICK et al. T. 485. ESA Special Publication, p. 65-73.
81. — (mars 2007). « An elementary introduction to the JWKB approximation ». In : *Astronomische Nachrichten* 328.3, p. 273. DOI : [10.1002/asna.200610730](https://doi.org/10.1002/asna.200610730).
82. GOUGH, D. O. et A. G. KOSOVICHEV (sept. 1993). « The Influence of Low-Degree P-Mode Frequencies on the Determination of the Structure of the Solar Interior ». In : *Monthly Notices of the Royal Astronomical Society* 264, p. 522. DOI : [10.1093/mnras/264.2.522](https://doi.org/10.1093/mnras/264.2.522).
83. GOUGH, D. O. et M. J. THOMPSON (1991). « The inversion problem. » In : *Solar Interior and Atmosphere*, p. 519-561.
84. GOUGH, D. O. et N. O. WEISS (sept. 1976). « The calibration of stellar convection theories. » In : *Monthly Notices of the Royal Astronomical Society* 176, p. 589-607. DOI : [10.1093/mnras/176.3.589](https://doi.org/10.1093/mnras/176.3.589).
85. GOUGH, Douglas (oct. 2013). « What Have We Learned from Helioseismology, What Have We Really Learned, and What Do We Aspire to Learn ? » In : *Solar Physics* 287.1-2, p. 9-41. DOI : [10.1007/s11207-012-0099-1](https://doi.org/10.1007/s11207-012-0099-1).
86. GRABOSKE, H. C., D. J. HARWOOD et F. J. ROGERS (oct. 1969). « Thermodynamic Properties of Nonideal Gases. I. Free-Energy Minimization Method ». In : *Physical Review* 186.1, p. 210-225. DOI : [10.1103/PhysRev.186.210](https://doi.org/10.1103/PhysRev.186.210).
87. HARVEY, J. W. et al. (mai 1996). « The Global Oscillation Network Group (GONG) Project ». In : *Science* 272.5266, p. 1284-1286. DOI : [10.1126/science.272.5266.1284](https://doi.org/10.1126/science.272.5266.1284).
88. HO, Tin Kam (1995). « Random decision forests ». In : *Proceedings of 3rd International Conference on Document Analysis and Recognition*. T. 1, 278-282 vol.1. DOI : [10.1109/ICDAR.1995.598994](https://doi.org/10.1109/ICDAR.1995.598994).
89. HOUDAYER, Pierre S., Daniel R. REESE et Marie-Jo GOUPIL (juill. 2022). « Properties of the ionisation glitch. II. Seismic signature of the structural perturbation ». In : *Astronomy and Astrophysics* 663, A60, A60. DOI : [10.1051/0004-6361/202243298](https://doi.org/10.1051/0004-6361/202243298).
90. HOUDAYER, Pierre S. et al. (nov. 2021). « Properties of the ionisation glitch. I. Modelling the ionisation region ». In : *Astronomy and Astrophysics* 655, A85, A85. DOI : [10.1051/0004-6361/202141711](https://doi.org/10.1051/0004-6361/202141711).
91. HOUDEK, G. et D. O. GOUGH (mars 2007). « An asteroseismic signature of helium ionization ». In : *Monthly Notices of the Royal Astronomical Society* 375.3, p. 861-880. DOI : [10.1111/j.1365-2966.2006.11325.x](https://doi.org/10.1111/j.1365-2966.2006.11325.x).

92. HOUDEK, G. et D. O. GOUGH (déc. 2011). « On the seismic age and heavy-element abundance of the Sun ». In : *Monthly Notices of the Royal Astronomical Society* 418.2, p. 1217-1230. DOI : [10.1111/j.1365-2966.2011.19572.x](https://doi.org/10.1111/j.1365-2966.2011.19572.x).
93. HOWE, Rachel (déc. 2009). « Solar Interior Rotation and its Variation ». In : *Living Reviews in Solar Physics* 6.1, 1, p. 1. DOI : [10.12942/lrsp-2009-1](https://doi.org/10.12942/lrsp-2009-1).
94. HUMMER, D. G. et Dimitri MIHALAS (août 1988). « The Equation of State for Stellar Envelopes. I. an Occupation Probability Formalism for the Truncation of Internal Partition Functions ». In : *The Astrophysical Journal* 331, p. 794. DOI : [10.1086/166600](https://doi.org/10.1086/166600).
95. HUNGER, K. (jan. 1955). « Die Atmosphäre des AO-Sternes Alpha Lyrae. Mit 12 Textabbildungen ». In : *Zeitschrift fuer Astrophysik* 36, p. 42.
96. JEFFREY, W. (avr. 1988). « Inversion of Helioseismology Data ». In : *The Astrophysical Journal* 327, p. 987. DOI : [10.1086/166255](https://doi.org/10.1086/166255).
97. JEFFREYS, Harold (1925). « On Certain Approximate Solutions of Lineae Differential Equations of the Second Order* ». In : *Proceedings of the London Mathematical Society* s2-23.1, p. 428-436. DOI : <https://doi.org/10.1112/plms/s2-23.1.428>.
98. — (1946). « An invariant form for the prior probability in estimation problems ». In : *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 186.1007, p. 453-461. DOI : [10.1098/rspa.1946.0056](https://doi.org/10.1098/rspa.1946.0056).
99. KELVIN, L. (mars 1862). « On the Age of the Sun's Heat ». In : *Macmillan's Magazine* 5, p. 388-393.
100. KIPPENHAHN, Rudolf, Alfred WEIGERT et Achim WEISS (2012). *Stellar Structure and Evolution*. DOI : [10.1007/978-3-642-30304-3](https://doi.org/10.1007/978-3-642-30304-3).
101. KIRKWOOD, John G. (1946). « The Statistical Mechanical Theory of Transport Processes I. General Theory ». In : *The Journal of Chemical Physics* 14.3, p. 180-201. DOI : [10.1063/1.1724117](https://doi.org/10.1063/1.1724117).
102. KJELDSSEN, H. et T. R. BEDDING (jan. 1995). « Amplitudes of stellar oscillations : the implications for asteroseismology. » In : *Astronomy and Astrophysics* 293, p. 87-106.
103. KOSOVICHEV, A. G. (déc. 1993). « Seismic measurements of the helium abundance and the depth of stellar convection zones ». In : *Monthly Notices of the Royal Astronomical Society* 265, p. 1053. DOI : [10.1093/mnras/265.4.1053](https://doi.org/10.1093/mnras/265.4.1053).
104. KRAMERS, H. A. (oct. 1926). « Wellenmechanik und halbzahlige Quantisierung ». In : *Zeitschrift fur Physik* 39.10-11, p. 828-840. DOI : [10.1007/BF01451751](https://doi.org/10.1007/BF01451751).
105. LANDAU, Lev Davidovich et E. M. LIFSHITZ (1969). *Statistical physics. Pt.1*.
106. LANGER, Rudolph E. (avr. 1937). « On the Connection Formulas and the Solutions of the Wave Equation ». In : *Physical Review* 51.8, p. 669-676. DOI : [10.1103/PhysRev.51.669](https://doi.org/10.1103/PhysRev.51.669).
107. LARKIN, A I (juin 1960). « Thermodynamic functions at low temperature plasma ». In : *Zhur. Eksptl. i Teoret. Fiz.* 38.
108. LEBRETON, Y. et M. J. GOUPIL (sept. 2014). « Asteroseismology for "à la carte" stellar age-dating and weighing. Age and mass of the CoRoT exoplanet host HD 52265 ». In : *Astronomy and Astrophysics* 569, A21, A21. DOI : [10.1051/0004-6361/201423797](https://doi.org/10.1051/0004-6361/201423797).
109. LEDOUX, Paul et Théodore WALRAVEN (jan. 1958). « Variable Stars. » In : *Handbuch der Physik* 51, p. 353-604. DOI : [10.1007/978-3-642-45908-5_6](https://doi.org/10.1007/978-3-642-45908-5_6).
110. LEVERMORE, C. David (juin 1996). « Moment closure hierarchies for kinetic theories ». In : *Journal of Statistical Physics* 83.5-6, p. 1021-1065. DOI : [10.1007/BF02179552](https://doi.org/10.1007/BF02179552).
111. LIFSHITZ, E. M. et L. P. PITAEVSKII (1981). *Physical kinetics*.
112. LIOUVILLE, J. (1838). « Note sur la Théorie de la Variation des constantes arbitraires. » fre. In : *Journal de Mathématiques Pures et Appliquées*, p. 342-349.

113. LUND, Mikkel N. et al. (fév. 2017). « Standing on the Shoulders of Dwarfs : the Kepler Asteroseismic LEGACY Sample. I. Oscillation Mode Parameters ». In : *The Astrophysical Journal* 835.2, 172, p. 172. DOI : [10.3847/1538-4357/835/2/172](https://doi.org/10.3847/1538-4357/835/2/172).
114. MAEDER, André (2009). *Physics, Formation and Evolution of Rotating Stars*. DOI : [10.1007/978-3-540-76949-1](https://doi.org/10.1007/978-3-540-76949-1).
115. MATHUR, S. et al. (avr. 2012). « A Uniform Asteroseismic Analysis of 22 Solar-type Stars Observed by Kepler ». In : *The Astrophysical Journal* 749.2, 152, p. 152. DOI : [10.1088/0004-637X/749/2/152](https://doi.org/10.1088/0004-637X/749/2/152).
116. MAXWELL, James Clerk (1867). « IV. On the dynamical theory of gases ». In : *Philosophical Transactions of the Royal Society of London* 157, p. 49-88. DOI : [10.1098/rstl.1867.0004](https://doi.org/10.1098/rstl.1867.0004).
117. MAZUMDAR, A. (oct. 2005). « Asteroseismic diagrams for solar-type stars ». In : *Astronomy and Astrophysics* 441.3, p. 1079-1086. DOI : [10.1051/0004-6361:20042348](https://doi.org/10.1051/0004-6361:20042348).
118. MAZUMDAR, A. et H. M. ANTIA (mars 2001). « Seismic detection of stellar tachoclines ». In : *Astronomy and Astrophysics* 368, p. L8-L12. DOI : [10.1051/0004-6361:20010145](https://doi.org/10.1051/0004-6361:20010145).
119. MAZUMDAR, A. et al. (fév. 2014). « Measurement of Acoustic Glitches in Solar-type Stars from Oscillation Frequencies Observed by Kepler ». In : *The Astrophysical Journal* 782.1, 18, p. 18. DOI : [10.1088/0004-637X/782/1/18](https://doi.org/10.1088/0004-637X/782/1/18).
120. METCALFE, T. S. et al. (oct. 2014). « Properties of 42 Solar-type Kepler Targets from the Asteroseismic Modeling Portal ». In : *ApJ Supplements* 214.2, 27, p. 27. DOI : [10.1088/0067-0049/214/2/27](https://doi.org/10.1088/0067-0049/214/2/27).
121. MIHALAS, Dimitri, Werner DAPPEN et D. G. HUMMER (août 1988). « The Equation of State for Stellar Envelopes. II. Algorithm and Selected Results ». In : *The Astrophysical Journal* 331, p. 815. DOI : [10.1086/166601](https://doi.org/10.1086/166601).
122. MINERBO, G. N. (jan. 1978). « Maximum entropy Eddington factors. » In : *Journal of Quantitative Spectroscopy & Radiative Transfer* 20.6, p. 541-545. DOI : [10.1016/0022-4073\(78\)90024-9](https://doi.org/10.1016/0022-4073(78)90024-9).
123. MONTEIRO, M. J. P. F. G., J. CHRISTENSEN-DALSGAARD et M. J. THOMPSON (mars 1994). « Seismic study of overshoot at the base of the solar convective envelope ». In : *Astronomy and Astrophysics* 283.1, p. 247-262.
124. MONTEIRO, M. J. P. F. G. et M. J. THOMPSON (jan. 1998). « On the seismic signature of the HeII ionization zone in stellar envelopes ». In : *New Eyes to See Inside the Sun and Stars*. Sous la dir. de Franz-Ludwig DEUBNER, Joergen CHRISTENSEN-DALSGAARD et Don KURTZ. T. 185, p. 317.
125. MONTEIRO, Mário J. P. F. G., Jørgen CHRISTENSEN-DALSGAARD et Michael J. THOMPSON (juill. 2000). « Seismic study of stellar convective regions : the base of the convective envelope in low-mass stars ». In : *Monthly Notices of the Royal Astronomical Society* 316.1, p. 165-172. DOI : [10.1046/j.1365-8711.2000.03471.x](https://doi.org/10.1046/j.1365-8711.2000.03471.x).
126. MONTEIRO, Mário J. P. F. G. et Michael J. THOMPSON (août 2005). « Seismic analysis of the second ionization region of helium in the Sun - I. Sensitivity study and methodology ». In : *Monthly Notices of the Royal Astronomical Society* 361.4, p. 1187-1196. DOI : [10.1111/j.1365-2966.2005.09246.x](https://doi.org/10.1111/j.1365-2966.2005.09246.x).
127. NOELS, A. et al. (juill. 2010). « Overshooting and semiconvection : structural changes and asteroseismic signatures ». In : *Astrophysics and Space Science* 328.1-2, p. 227-236. DOI : [10.1007/s10509-009-0203-7](https://doi.org/10.1007/s10509-009-0203-7).
128. NSAMBA, B. et al. (juill. 2018). « Asteroseismic modelling of solar-type stars : internal systematics from input physics and surface correction methods ». In : *Monthly Notices of the Royal Astronomical Society* 477.4, p. 5052-5063. DOI : [10.1093/mnras/sty948](https://doi.org/10.1093/mnras/sty948).
129. PIAU, L., J. BALLOT et S. TURCK-CHIÈZE (fév. 2005). « Seismic constraints on open clusters ». In : *Astronomy and Astrophysics* 430, p. 571-583. DOI : [10.1051/0004-6361:20041405](https://doi.org/10.1051/0004-6361:20041405).

130. PIJPERS, F. P. et M. J. THOMPSON (sept. 1992). « Faster formulations of the optimally localized averages method for helioseismic inversions ». In : *Astronomy and Astrophysics* 262.2, p. L33-L36.
131. — (jan. 1994). « The SOLA method for helioseismic inversion ». In : *Astronomy and Astrophysics* 281.1, p. 231-240.
132. PINÇON, C. et al. (sept. 2017). « Can plume-induced internal gravity waves regulate the core rotation of subgiant stars ? » In : *Astronomy & Astrophysics* 605, A31, A31. DOI : [10.1051/0004-6361/201730998](https://doi.org/10.1051/0004-6361/201730998).
133. PLANCK, M. (1967). « 1 - On an Improvement of Wien's Equation for the Spectrum » Verh. Dtsch. Phys. Ges. Berlin 2, 202 (1900). » In : *The Old Quantum Theory*. Sous la dir. de D. Ter HAAR. The Commonwealth and International Library : Selected Readings in Physics. Pergamon, p. 79-81. DOI : <https://doi.org/10.1016/B978-0-08-012102-4.50012-7>.
134. POMRANING, G. C. (1973). *The equations of radiation hydrodynamics*.
135. POMRANING, G C (déc. 1982). « Radiation hydrodynamics ». In : DOI : [10.2172/656708](https://doi.org/10.2172/656708).
136. PRAT, V. et F. LIGNIÈRES (juin 2014). « Shear mixing in stellar radiative zones. I. Effect of thermal diffusion and chemical stratification ». In : *Astronomy and Astrophysics* 566, A110, A110. DOI : [10.1051/0004-6361/201423655](https://doi.org/10.1051/0004-6361/201423655).
137. RASMUSSEN, Carl Edward et Christopher K. I. WILLIAMS (nov. 2005). *Gaussian Processes for Machine Learning*. The MIT Press. DOI : [10.7551/mitpress/3206.001.0001](https://doi.org/10.7551/mitpress/3206.001.0001).
138. RAUER, H. et al. (nov. 2014). « The PLATO 2.0 mission ». In : *Experimental Astronomy* 38.1-2, p. 249-330. DOI : [10.1007/s10686-014-9383-4](https://doi.org/10.1007/s10686-014-9383-4).
139. REESE, D. R. et al. (mars 2012). « Estimating stellar mean density through seismic inversions ». In : *Astronomy and Astrophysics* 539, A63, A63. DOI : [10.1051/0004-6361/201118156](https://doi.org/10.1051/0004-6361/201118156).
140. REMPEL, M. (juin 2004). « Overshoot at the Base of the Solar Convection Zone : A Semianalytical Approach ». In : *The Astrophysical Journal* 607.2, p. 1046-1064. DOI : [10.1086/383605](https://doi.org/10.1086/383605).
141. RENZINI, A. (déc. 1987). « Some embarrassments in current treatments of convective overshooting ». In : *Astronomy and Astrophysics* 188.1, p. 49-54.
142. RICKER, George R. et al. (jan. 2015). « Transiting Exoplanet Survey Satellite (TESS) ». In : *Journal of Astronomical Telescopes, Instruments, and Systems* 1, 014003, p. 014003. DOI : [10.1117/1.JATIS.1.1.014003](https://doi.org/10.1117/1.JATIS.1.1.014003).
143. RIEUTORD, M. (2004). *Formulaire pour la mécanique des fluides dans la sphère (ou Manuel du Plombier des Astres)*.
144. ROGERS, Forrest J. et Carlos A. IGLESIAS (jan. 1994). « Astrophysical Opacity ». In : *Science* 263.5143, p. 50-55. DOI : [10.1126/science.263.5143.50](https://doi.org/10.1126/science.263.5143.50).
145. ROGERS, Forrest J., Fritz J. SWENSON et Carlos A. IGLESIAS (jan. 1996). « OPAL Equation-of-State Tables for Astrophysical Applications ». In : *Astrophysical Journal* 456, p. 902. DOI : [10.1086/176705](https://doi.org/10.1086/176705).
146. ROGERS, Tamara M., Gary A. GLATZMAIER et C. A. JONES (déc. 2006). « Numerical Simulations of Penetration and Overshoot in the Sun ». In : *The Astrophysical Journal* 653.1, p. 765-773. DOI : [10.1086/508482](https://doi.org/10.1086/508482).
147. ROSSELAND, S. (mai 1924). « Note on the absorption of radiation within a star ». In : *MNRAS* 84, p. 525-528. DOI : [10.1093/mnras/84.7.525](https://doi.org/10.1093/mnras/84.7.525).
148. SACKUR, O. (jan. 1911). « Die Anwendung der kinetischen Theorie der Gase auf chemische Probleme ». In : *Annalen der Physik* 341.15, p. 958-980. DOI : [10.1002/andp.19113411505](https://doi.org/10.1002/andp.19113411505).
149. SAHA, Megh Nad (oct. 1920). *LIII. Ionization in the solar chromosphere*. DOI : [10.1080/14786441008636148](https://doi.org/10.1080/14786441008636148).

150. SAHA, Meghnad N et Alfred FOWLER (1921). « On a physical theory of stellar spectra ». In : *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 99.697, p. 135-153. DOI : [10.1098/rspa.1921.0029](https://doi.org/10.1098/rspa.1921.0029).
151. SAMADI, R., K. BELKACEM et T. SONOI (fév. 2015). « Stellar oscillations - II - The non-adiabatic case ». In : *EAS Publications Series*. T. 73-74. EAS Publications Series, p. 111-191. DOI : [10.1051/eas/1573003](https://doi.org/10.1051/eas/1573003).
152. SCHALLER, G. et al. (déc. 1992). « New Grids of Stellar Models from 0.8-SOLAR-MASS to 120-SOLAR-MASSSES at Z=0.020 and Z=0.001 ». In : *Astronomy and Astrophysics Supplement* 96, p. 269.
153. SCHOU, J. et al. (sept. 1998). « Helioseismic Studies of Differential Rotation in the Solar Envelope by the Solar Oscillations Investigation Using the Michelson Doppler Imager ». In : *The Astrophysical Journal* 505.1, p. 390-417. DOI : [10.1086/306146](https://doi.org/10.1086/306146).
154. SCHWARZSCHILD, Martin (1958). *Structure and evolution of the stars*.
155. SERENELLI, Aldo M., W. C. HAXTON et Carlos PEÑA-GARAY (déc. 2011). « Solar Models with Accretion. I. Application to the Solar Abundance Problem ». In : *The Astrophysical Journal* 743.1, 24, p. 24. DOI : [10.1088/0004-637X/743/1/24](https://doi.org/10.1088/0004-637X/743/1/24).
156. SERENELLI, Aldo M. et al. (nov. 2009). « New Solar Composition : The Problem with Solar Models Revisited ». In : *ApJ Letters* 705.2, p. L123-L127. DOI : [10.1088/0004-637X/705/2/L123](https://doi.org/10.1088/0004-637X/705/2/L123).
157. SHANNON, Claude Elwood (1948). « A Mathematical Theory of Communication ». In : *The Bell System Technical Journal* 27, p. 379-423.
158. SILVA AGUIRRE, V. et al. (juin 2013). « Stellar Ages and Convective Cores in Field Main-sequence Stars : First Asteroseismic Application to Two Kepler Targets ». In : *The Astrophysical Journal* 769.2, 141, p. 141. DOI : [10.1088/0004-637X/769/2/141](https://doi.org/10.1088/0004-637X/769/2/141).
159. SONOI, T. et al. (nov. 2015). « Surface-effect corrections for solar-like oscillations using 3D hydrodynamical simulations. I. Adiabatic oscillations ». In : *Astronomy and Astrophysics* 583, A112, A112. DOI : [10.1051/0004-6361/201526838](https://doi.org/10.1051/0004-6361/201526838).
160. STASSUN, Keivan G. et al. (oct. 2019). « The Revised TESS Input Catalog and Candidate Target List ». In : *The Astronomical Journal* 158.4, 138, p. 138. DOI : [10.3847/1538-3881/ab3467](https://doi.org/10.3847/1538-3881/ab3467).
161. TASSOUL, M. (août 1980). « Asymptotic approximations for stellar nonradial pulsations. » In : *ApJ Supplements* 43, p. 469-490. DOI : [10.1086/190678](https://doi.org/10.1086/190678).
162. TETRODE, H. (jan. 1912). « Berichtigung zu meiner Arbeit : Die chemische Konstante der Gase und das elementare Wirkungsquantum ». In : *Annalen der Physik* 344.11, p. 255-256. DOI : [10.1002/andp.19123441112](https://doi.org/10.1002/andp.19123441112).
163. THOMAS, H. C. (jan. 1967). « Sternentwicklung VIII. Der Helium-Flash bei einem Stern von 1.3 Sonnenmassen ». In : *Zeitschrift fur Astrophysik* 67, p. 420.
164. THOMAS, L. H. (jan. 1930). « The radiation field in a fluid in motion ». In : *The Quarterly Journal of Mathematics* os-1.1, p. 239-251. DOI : [10.1093/qmath/os-1.1.239](https://doi.org/10.1093/qmath/os-1.1.239).
165. THOMPSON, Michael J. et al. (jan. 2003). « The Internal Rotation of the Sun ». In : *Annual Review of Astronomy and Astrophysics* 41, p. 599-643. DOI : [10.1146/annurev.astro.41.011802.094848](https://doi.org/10.1146/annurev.astro.41.011802.094848).
166. THOUL, Anne A., John N. BAHCALL et Abraham LOEB (fév. 1994). « Element Diffusion in the Solar Interior ». In : *Astrophysical Journal* 421, p. 828. DOI : [10.1086/173695](https://doi.org/10.1086/173695).
167. TRAVING, G. (jan. 1955). « Die Atmosphäre des BO-Sternes τ Scorpii. Mit 9 Textabbildungen. » In : *Zeitschrift fuer Astrophysik* 36, p. 1.
168. ULRICH, R. K. (juill. 1986). « Determination of Stellar Ages from Asteroseismology ». In : *ApJ Letters* 306, p. L37. DOI : [10.1086/184700](https://doi.org/10.1086/184700).

169. UNNO, Wasaburo et al. (1989). *Nonradial oscillations of stars*.
170. VANDAKUROV, Yu. V. (fév. 1968). « The Frequency Distribution of Stellar Oscillations. » In : *Soviet Astronomy* 11, p. 630.
171. VEDENOV, AA et AI LARKIN (1959). « Equation of state of plasmas ». In : *Sov. Phys. JETP* 9, p. 806-821.
172. VERMA, Kuldeep et Víctor SILVA AGUIRRE (oct. 2019). « Helium settling in F stars : constraining turbulent mixing using observed helium glitch signature ». In : *Monthly Notices of the Royal Astronomical Society* 489.2, p. 1850-1858. DOI : [10.1093/mnras/stz2272](https://doi.org/10.1093/mnras/stz2272).
173. VERMA, Kuldeep et al. (oct. 2014a). « A Theoretical Study of Acoustic Glitches in Low-mass Main-sequence Stars ». In : *The Astrophysical Journal* 794.2, 114, p. 114. DOI : [10.1088/0004-637X/794/2/114](https://doi.org/10.1088/0004-637X/794/2/114).
174. VERMA, Kuldeep et al. (août 2014b). « Asteroseismic Estimate of Helium Abundance of a Solar Analog Binary System ». In : *The Astrophysical Journal* 790.2, 138, p. 138. DOI : [10.1088/0004-637X/790/2/138](https://doi.org/10.1088/0004-637X/790/2/138).
175. VERMA, Kuldeep et al. (mars 2017). « Seismic Measurement of the Locations of the Base of Convection Zone and Helium Ionization Zone for Stars in the Kepler Seismic LEGACY Sample ». In : *The Astrophysical Journal* 837.1, 47, p. 47. DOI : [10.3847/1538-4357/aa5da7](https://doi.org/10.3847/1538-4357/aa5da7).
176. VERMA, Kuldeep et al. (mars 2019). « Helium abundance in a sample of cool stars : measurements from asteroseismology ». In : *Monthly Notices of the Royal Astronomical Society* 483.4, p. 4678-4694. DOI : [10.1093/mnras/sty3374](https://doi.org/10.1093/mnras/sty3374).
177. VERMA, Kuldeep et al. (juill. 2022). « Advanced asteroseismic modelling : breaking the degeneracy between stellar mass and initial helium abundance ». In : *Monthly Notices of the Royal Astronomical Society*. DOI : [10.1093/mnras/stac1860](https://doi.org/10.1093/mnras/stac1860).
178. VILLANI, Cédric (2002). « Chapter 2 – A Review of Mathematical Topics in Collisional Kinetic Theory ». In.
179. VON MISES, Richard (1964). « Chapter VIII - More on distributions ». In : *Mathematical Theory of Probability and Statistics*. Sous la dir. de Richard VON MISES. Academic Press, p. 368-430. DOI : <https://doi.org/10.1016/B978-1-4832-3213-3.50011-7>.
180. VON ZEIPPEL, H. (juin 1924). « The radiative equilibrium of a rotating system of gaseous masses ». In : *Monthly Notices of the Royal Astronomical Society* 84, p. 665-683. DOI : [10.1093/mnras/84.9.665](https://doi.org/10.1093/mnras/84.9.665).
181. VRARD, Mathieu et Margarida S. CUNHA (nov. 2019). « Influence of structural discontinuities present in the core of red-giant stars on the observed mixed-mode pattern and characterization of their properties ». In : *arXiv e-prints*, arXiv :1911.05175, arXiv :1911.05175.
182. WEIDEMANN, V. (jan. 1955). « Metallhäufigkeiten, Druckschichtung und Stoßdämpfung in der Sonnenatmosphäre. Mit 12 Textabbildungen ». In : *Zeitschrift fuer Astrophysik* 36, p. 101.
183. WENTZEL, Gregor (juin 1926). « Eine Verallgemeinerung der Quantenbedingungen für die Zwecke der Wellenmechanik ». In : *Zeitschrift fur Physik* 38.6-7, p. 518-529. DOI : [10.1007/BF01397171](https://doi.org/10.1007/BF01397171).
184. WOODBURY, M.A. et Princeton University. Department of STATISTICS (1950). *Inverting Modified Matrices*. Memorandum Report / Statistical Research Group, Princeton. Department of Statistics, Princeton University.
185. YVON, J. (1935). *La théorie statistique des fluides et l'équation d'état*. Actualités Scientifiques et Industrielles : Théories Mécaniques. Hermann & cie.
186. ZAHN, J. P. (déc. 1991). « Convective penetration in stellar interiors. » In : *Astronomy and Astrophysics* 252, p. 179-188.
187. — (nov. 1992). « Circulation and turbulence in rotating stars. » In : *Astronomy and Astrophysics* 265, p. 115-132.
188. ZAHN, J. P., S. TALON et J. MATIAS (juin 1997). « Angular momentum transport by internal waves in the solar interior. » In : *Astronomy & Astrophysics* 322, p. 320-328.