Revisiting statistical physics tools for the complex time-frequency characterization of physiological signals

Alexandre Guillet

université de BORDEAUX

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR

DE L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES PHYSIQUES ET DE L'INGÉNIEUR

SPÉCIALITÉ LASERS, MATIÈRE ET NANOSCIENCES

Alexandre GUILLET

Revisiting statistical physics tools for the complex time-frequency characterization of physiological signals

---

Sous la direction de Françoise ARGOUL et Alain ARNÉODO

Soutenue le 6 mai 2022

Membres du jury :

M. Bernard CAZELLES	Professeur, Sorbonne Université	**Président**
Mrs. Selin AVIYENTE	Professor, Michigan State University	**Rapporteur**
M. Bruno TORRÉSANI	Professeur, Université d'Aix-Marseille	**Rapporteur**
Mr. Ronny BARTSCH	Associate Professor, Bar-Ilan University	**Examinateur**
Mr. Lucas FAES	Associate Professor, University of Palermo	**Examinateur**
Mme. Françoise ARGOUL	Directrice de Recherche au CNRS	**Directrice de thèse**

Outils de physique statistique revisités pour la caractérisation complexe et temps-fréquence des signaux physiologiques

Résumé :
Un outil analytique pour l'observation de la dynamique des organismes vivants est développé dans cette thèse, et appliqué à l'étude des interactions entre les systèmes cardiaque, respiratoire et nerveux basée sur des enregistrements humains pendant le sommeil. Ces signaux physiologiques sont particuliers et complexes : chaque enregistrement est unique, non stationnaire, et ne peux être reproduit à l'identique. Pour répondre à cette difficulté, nous proposons une analyse temps-fréquence de ces signaux physiologiques, basée sur des ondelettes analytiques aussi appelées « atomes temps-fréquences ». La première partie A de cette thèse introduit ce formalisme, souligne l'adéquation de l'ondelette log-normale dans ce contexte, l'importance du facteur de qualité Q associé et sa correspondance avec le principe d'incertitude d'Heisenberg. En étendant cette approche à l'estimation de la cohérence temps-fréquence entre signaux physiologiques, nous découplons les résolutions temporelle et spectrale, afin d'introduire des degrés de liberté statistiques dans l'analyse. Pour aider à l'observation d'interactions physiologiques entre différentes sources, nous définissons une famille générique d'estimateurs de taux complexes, capturant les modulations tant de la fréquence instantanée que de l'amplitude, dans une bande spectrale d'intérêt. La seconde partie B illustre cette méthodologie sur des signaux physiologiques issus de bases de données polysomnographiques. Les riches spécificités des rythmes fluctuants contenus dans les électrocardiogrammes (ECG), signaux respiratoires et électroencéphalogrammes (EEG) sont discutées en détail pour valider l'estimation générique de la cohérence temps-fréquence entre paires d'enregistrements ou de leurs transformations en taux. Nous comparons l'efficacité des estimateurs, à bande large ou étroite pour extraire les modulations du taux cardiaque, à une méthode d'extraction de référence. Trois échelles d'interaction distinctes sont observées à partir des modulations cardiaques et respiratoires cohérentes : une première rapide à la fréquence fondamentale de respiration (arythmie sinusale respiratoire, RSA), une deuxième lente et moins rythmique et une troisième encore plus lente, toutes œuvrant à la régulation des taux cardio-respiratoires. Dans le formalisme des ondelettes, nous proposons une représentation originale de l'évolution du ratio entre fréquences cardiaque et respiratoire, qui se trouve être efficace pour suivre les traces de l'apnée du sommeil, des phases de sommeil paradoxal (REM) et de sommeil profond. Cette approche du couplage cardio-respiratoire offre également une nouvelle perspective pour détecter ses épisodes de synchronisation et ses fluctuations temporelles. L'activité nerveuse pendant le sommeil est introduite à travers l'analyse temps-fréquence des signaux EEG, et leurs séquences caractéristiques de motifs de densité de puissance et de cohérence dans de multiples bandes. Finalement, nous rassemblons les signaux de taux cardio-respiratoire et de magnitude nerveuse et nous les comparons par paires en termes de cohérence temps-fréquence. À partir d'une analyse extensive de 2650 enregistrements issus d'une large base de données polysomnographiques, nous relevons une augmentation de l'amplitude et de la cohérence des modulations pendant l'apnée, émergeant systématiquement comme un pique spectral comparé à une ligne de base sans apnée. Nous montrons aussi que l'apnée du sommeil est concomitante d'une importante réduction de l'interaction cardio-respiratoire rapide (RSA) et de l'interaction neuro-respiratoire très lente. Cette étude ouvre la voie à de nouveaux développements, esquissés à la fin de cette thèse, tels que la détection des fuseaux de sommeil à partir de la phase des oscillations dans la bande EEG sigma, et la cohérence de leur amplitude avec les oscillations lentes delta pendant le sommeil profond.

Mots-clés : biophysique, temps-fréquence, multi-échelle, ondelette, cohérence, statistique, polysomnographie, traitement du signal

Revisiting statistical physics tools for the complex time-frequency characterization of physiological signals

Abstract:
An analytical tool for the observation of the dynamics of living organisms is developed in this thesis, and applied to the study of interactions between the cardiac, respiratory and neural systems based on human recordings during sleep. These physiological signals are peculiar and complex: each recording is unique, non-stationary and cannot be reproduced. To solve this issue, a time-frequency analysis of physiological signals, based on analytic wavelets also called "time-frequency atoms", is proposed. The first part A of this thesis introduces this formalism, highlighting the adequacy of the log-normal wavelet in that context, the importance of the associated quality factor Q and its correspondence with the Heisenberg uncertainty principle. Extending this approach to the estimation of the time-frequency coherence of physiological signals, we separate temporal from spectral resolutions in order to introduce statistical degrees of freedom in this analysis. To assist the observation of physiological interactions between different sources, we define a generic family of complex rate estimators capturing both the modulations of the instantaneous frequency and amplitude, in a spectral band of interest. The second part B illustrates this methodology on physiological signals from polysomnography databases. The rich specificities of fluctuating rhythms contained in electrocardiograms (ECG), breathing signals and electroencephalograms (EEG) are thoroughly discussed to validate the generic estimation of the time-frequency coherence between pairs of recordings or their rate transformations. We compare the effectiveness of wide-band and narrow-band estimators in retrieving heart rate modulations, relative to a reference extraction method. From coherent cardiac and respiratory modulations, we observe three distinct scales of interaction: a fast one at the fundamental breathing frequency (RSA), a slow and less rhythmic one and an even slower one, regulating cardio-respiratory rates. In the wavelet framework, we propose an original time-dependent representation of the frequency ratios of cardiac and breathing rhythms that turns out very effective for tracking signatures of sleep apnea, rapid eye movement (REM) phases, and deep sleep stages. This approach of the cardio-respiratory coupling also offers a new perspective to detect synchronization episodes of cardio-respiratory coupling and their temporal fluctuations. The neural activity during sleep is introduced from the time-frequency analysis of EEG signals, and their characteristic sequences of multi-band power density and coherence patterns. Finally, we gather cardio-respiratory rates and neural magnitude signals and compare them in pairs in terms of time-frequency coherence. From an extensive analysis of 2650 recordings from a large polysomnography database, we point out the increased amplitude and coherence of slow modulations during apnea, systematically emerging as a spectral peak as compared to a flat baseline without apnea. We also show that apnea is concomitant with a strong coherence reduction of the fast RSA cardio-respiratory interaction and the very slow neuro-respiratory interaction. This study paves the way to new developments, outlined in the end of this thesis, such as the detection of sleep spindles from the phase of the oscillations in the sigma EEG band, and the coherence of their amplitude with slow delta oscillations during deep sleep.

Keywords: biophysics, time-frequency, multi-scale, wavelet, coherence, statistics, polysomnography, signal processing

---

Laboratoire Ondes et Matière d'Aquitaine
UMR 5798 Université de Bordeaux, 33000 Bordeaux, France.

Remerciements

Je souhaite exprimer dans ces quelques lignes ma gratitude pour les personnes qui ont contribué au développement de cette thèse. Merci à Alain, mon premier interlocuteur dans cette aventure, avec qui j'aurais souhaité partager plus longtemps l'enthousiasme des petites découvertes. Je lui rends hommage. Merci à Françoise pour sa confiance, ses recadrages ciblés contre la dispersion et sa bienveillance. Je les remercie tous les deux pour m'avoir appris l'importance de l'observation attentive des données, le bon dosage du flair et cette introduction de choix à la recherche et à l'interdisciplinarité.

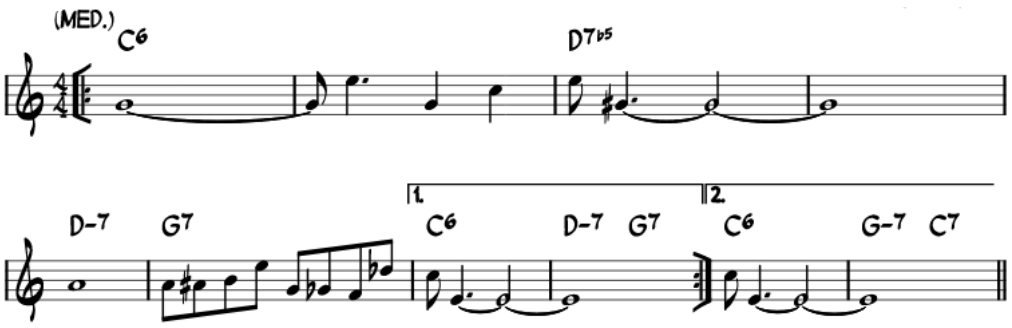
Merci à Cécilia pour son infaillible soutien au quotidien —particulièrement vers la fin de la thèse— dont je serai longtemps redevable.

Merci à Stefano Polizzi, Léo Delmarre, Etienne Harté et Romaric Vincent pour ces moments d'échange, de cohésion dans la difficulté, et d'émulation. Merci à Evgenyia Gerasimova et à Jacques Taillard, de m'avoir donné, avec leurs mesures et leurs questionnements, un point de départ pour l'une et un point d'arrivée pour l'autre. Merci aux volontaires anonymes pour leurs précieux enregistrements physiologiques, et aux chercheurs qui les ont rendus accessibles librement.

Merci à Jonathan Bitton, qui a infléchi le cours de cette thèse en m'ouvrant la piste de la cohérence, lors d'une réunion de la « famille ondelettes ». Merci à Pierre Argoul et à Bruno Torrésani, d'avoir stimulé la clarification du présent usage de l'ondelette log-normale et de la distribution des ratios de fréquence. Merci encore à Françoise, Bruno et Cécilia pour leurs nombreuses remarques, d'une grande aide lors de la relecture et la finalisation du manuscrit de thèse.

Merci au projet Sci-Hub, sans lequel la recherche bibliographique serait bien ennuyeuse.

Merci à Thomas Salez pour nous avoir fait connaître le LOMA en M2, et à tout le laboratoire et ses lomiens pour leur accueil et leur soutien. Mes salutations aux illustres successeurs et prédécesseurs occupants du bureau 120, et un remerciement particulier à Philippe Maire pour m'avoir aidé à garder le contact avec une muse discrète, mais bien réelle de cette thèse :



B. Strayhorn & D. Ellington

Introduction

Le Vivant et ses mouvements sont des phénomènes naturels, faisant partie de la φύσις (phýsis, c'est-à-dire l'Écosystème naturel) ; pourtant, la nature physique des systèmes vivants reste encore très élusive. Ils partagent, avec les systèmes hors équilibre et les systèmes dynamiques non-linéaires, la présence de comportements pulsatiles et quasi-périodiques. Les rythmes physiologiques ayant lieu dans le cœur, la respiration, la locomotion, les signaux nerveux, les horloges cellulaires ou moléculaires, les sécrétions hormonales, les cycles de vie, etc, couvrent plus de 10 décades à travers les échelles de temps. Les organismes vivants ont une existence complexe à de nombreuses échelles d'espace et de temps interconnectées, et souvent peu séparées. En dehors de ces caractéristiques multi-échelles, une autre propriété essentielle qui distingue la matière vivante de la matière inerte est son activité instable et imprédictible.

Toutes ces caractéristiques sont autant de défis adressés aux biophysiciens pour trouver un cadre cohérent de description des systèmes vivants. De nombreuses mesures répétées sont nécessaires pour compenser leur forte variabilité, d'un système à l'autre, mais aussi au sein d'un même système, pour lequel le comportement évolutif ne peut pas être reproduit sur demande dans beaucoup de situations expérimentales. L'enregistrement continu sur le corps humain de l'activité physiologique correspond à ce cas de figure, et se trouve au point focal de cette thèse. Il s'agit d'une expérience à essai unique : chaque trajectoire mesurée est unique et changeante, avec des oscillations régulières inextricablement mêlées à des fluctuations irrégulières aux caractéristiques intermittentes. Spécifique au système vivant durant l'observation, ce mélange de comportements imprédictibles et organisés est le résultat de l'interaction entre les sous-systèmes sous-jacents à de multiples échelles distinctes, des biomolécules, cellules, tissus, organes, à l'organisme complet et son environnement. Leur description réductionniste et intégrante à travers les échelles, en un modèle biophysique unique, représente une tâche formidable, qui restera peut-être au stade d'idée abstraite.

Reconnaître et examiner des rythmes est primordial en physiologie du cœur, de la respiration, du métabolisme, du cerveau, de l'audition, de la voix et du langage, de la locomotion et la proprioception... L'intérêt biomédical à déchiffrer leurs signaux complexes est une motivation forte pour soutenir l'effort de modélisation en proposant une analyse approfondie des observations. Le corps humain a été décrit comme un ensemble de sous-systèmes physiologiques distincts par les cliniciens, tels que les systèmes respiratoire, cardiaque et nerveux, finement inter-régulés. Leurs interactions forment un réseau, dont la dynamique est particulièrement complexe. Pour obtenir un meilleur aperçu de ces interactions, il est essentiel d'observer l'activité physiologique à travers différents enregistrements simultanés, par exemple une polysomnographie, mesurée par différents capteurs sur le corps au repos. L'étude de ces enregistrements concentre les efforts d'une large communauté de physiologistes, neuroscientifiques, cliniciens, biologistes, physiciens et mathématiciens, œuvrant à l'amélioration de la santé des personnes affectées par l'arythmie cardiaque, l'apnée du sommeil, l'insomnie, certains troubles psychiques, ou simplement le vieillissement, ainsi qu'à l'approfondissement de la compréhension des phénomènes associés au sommeil et à l'éveil, l'attention, les perceptions, les émotions, et peut-être même la conscience.

Le contexte de la mesure biophysique peut être assez éloigné de celui de l'expérience classique de physique. Les méthodes d'enregistrement disponibles sont basées sur différentes techniques physiques (électromagnétique, optique, thermique, barométrique), parfois redondantes comme dans le cas du rythme cardiaque. Les méthodes non intrusives et passives ont un avantage pratique comparées aux mesures plus directes et contrôlées précisément ; les premières sont donc préférées pour des applications cliniques ou commerciales. Mais ce sont aussi celles qui requièrent le plus de traitement pour en extraire les informations biophysiques d'intérêt, telles que des taux ou modulations d'intensité. Une profusion d'algorithmes sophistiqués, conçus pour

chacune des méthodes d'enregistrement, accompagnent la difficulté du processus de traitement du signal ou de l'image. La forte spécificité de ces techniques tend à effacer la généricité des relations mathématiques entre les observables biophysiques enregistrées et extraites. Ce manque d'une base commune se fait d'autant plus sentir lorsqu'on essaie de déchiffrer le réseau des interactions physiologiques les reliant. Ces considérations soulignent l'importance d'améliorer les outils mathématiques existants et les techniques numériques d'observation, qui sont pour certains explorés et développés dans cette thèse.

Une physique statistique de ces systèmes ne saurait être théorisée sans considérer à la fois les caractéristiques temporelles et fréquentielles de leurs comportements fluctuants et oscillants. Pour faire face à la complexité des enregistrements physiologiques, une nouvelle représentation temps-fréquence du signal apparaît comme un ingrédient essentiel, pour déployer et diluer en une carte à deux dimensions les riches informations initialement concentrées en une seule (temps ou fréquence). Les caractéristiques multi-échelles des signaux sont dépeintes à chaque instant dans la direction spectrale, et leurs caractéristiques évolutives à chaque fréquence dans la direction temporelle. Une représentation temps-fréquence emploie une fonction test pour scruter le contenu oscillatoire du signal analysé en différentes localisations : soit au moyen de translations en temps et en fréquences (comme la fenêtre glissante de la transformée de Fourier à court terme), ou par une translation et une dilatation du temps (comme dans la transformée en ondelettes, utilisée tout au long de cette thèse). Ces deux méthodes mènent à des interprétations apparemment distinctes, telles que les « fréquences » et les « échelles », diversifiées par les innombrables possibilités de fonctions test (incluant le signal lui-même) et leur mise en œuvre pratique. D'un côté, cela constitue un gage de versatilité, qui s'est traduit en de nombreuses applications scientifiques. De l'autre, cela implique une profusion de détails techniques à gérer, cachant parfois des liens avec des intuitions ou concepts plus fondamentaux.

L'interprétation probabiliste des fluctuations dans les enregistrements naturels est le second ingrédient. Les processus stochastiques prennent une place grandissante dans la modélisation physique des trajectoires fluctuantes, pour les phénomènes micro et mésoscopiques, tandis que les méthodes d'inférence statistique sont le plus souvent adaptées à la complexité des systèmes macroscopiques. Ces approches partagent les mêmes idées issues de la théorie de l'information. La définition d'ensembles statistiques, à la fois au sein de l'enregistrement et entre enregistrements physiologiques, est une tâche délicate, qui vise à extraire des observations robustes de mesures uniques aux caractéristiques non stationnaires et multi-échelles. Ces conditions limitent souvent les estimations au premier et second ordres statistiques, comme l'analyse des corrélations, et l'évaluation de leur signifiance est primordiale. Bien qu'élémentaire, l'analyse des corrélations est efficace pour évaluer la connectivité et les interactions entre les quantités examinées, qui peuvent être de natures physiques très différentes.

La combinaison de l'approche statistique des corrélations et d'une perspective temps-fréquence a permis d'aboutir à une méthode composite appelée la cohérence temps-fréquence. Les données de deux enregistrements simultanés sont décomposées à différents temps et échelles avant d'être analysées statistiquement. Cet outil d'investigation de dynamiques complexes, relativement récent, s'est montré efficace et a de nombreuses applications dans des contextes physiologique, écologique, géophysique, économiques entre autres. À nouveau, ce type d'analyse a été mis en œuvre dans des modalités variées, avec divers degrés de contrôle de la résolution temps-fréquence et de la signifiance statistique. De plus, il dépend souvent de signaux systématiquement pré-traités, ayant subi diverses procédures de retrait de la tendance, blanchissement et filtres, hors du cadre de l'analyse de la cohérence. Finalement, les quantités dynamiques d'intérêt sont rarement celles directement mesurées. Par exemple, les signaux de taux cardiaque et respiratoire ou les ondes cérébrales sont systématiquement extraits, par un moyen algorithmique, des enregistrements physiologiques contenant l'information originelle.

La question centrale au fondement de cette thèse peut être exprimée comme suit : y a-t-il un

langage commun utilisable pour les différentes tâches, allant du traitement des enregistrements physiologiques bruts en de nouvelles observables biophysiques à l'étude de leurs interactions ? Quels sont les paramètres pertinents pour parvenir à ce type d'analyse de manière générique ? Dans un effort de synthèse, nous établissons un cadre générique basé sur une représentation temps-fréquence de l'enregistrement utilisant des ondelettes, avec un intérêt particulier pour l'interprétation physique. Cet outil est appliqué à l'étude des enregistrements polysomnographiques de l'activité cardiaque, respiratoire et nerveuse, et de leur riche phénoménologie.

La construction de ce cadre théorique est présentée dans la partie A de cette thèse, composée des chapitres I, II et III, suivie par une application aux enregistrements physiologiques dans la partie B, composée des chapitres IV et V. Leur bibliographie distincte est introduite progressivement dans les chapitres concernés.

Dans le chapitre I, nous commençons par l'examen des représentations temporelles et spectrales des oscillations, pour un signal réel et son spectre. En élaborant sur les avantages de la version analytique ou « en hélice » complexe du signal pour interpréter ses oscillations dans le temps, nous montrons que l'idée de distinguer des contributions à de multiples échelles mène naturellement à une transformée en ondelettes dont l'ondelette est analytique. L'analyse en ondelettes revient à décomposer un signal en « atomes » temps-fréquence, appelés ondelettes, choisis de manière adéquate pour accéder à l'information souhaitée. En particulier, nous récupérons l'intuition spectrale de la transformée de Fourier en examinant la forme du spectre de l'ondelette. Nous revisitons la famille générale des ondelettes analytiques de Morse, de manière à rendre explicite et ajustable le seul paramètre pertinent : le facteur de qualité. À cet égard, nous trouvons idéal la forme log-normale, cas limite dans la famille de Morse, avec une symétrie, localisation et régularité exceptionnelles. Décrite par R.A. Altes, et proposée par A. Grossmann au début de la théorie des ondelettes comme étant « particulièrement bonne », cette forme demeure paradoxalement quasi-absente de la littérature qui en découle. Les propriétés de l'ondelette log-normale sont détaillées, ainsi que la résolution dans le temps et les fréquences de la représentation associée, caractérisée par un compromis entre incertitudes contrôlé explicitement par l'unique paramètre, le facteur de qualité.

Nous continuons dans le chapitre II avec une simplification de la notation pour la représentation en ondelettes de l'enregistrement, l'illustration de l'amplitude et de la phase des oscillations dans le continuum temps-fréquence, et des considérations géométriques et numériques. Nous mettons en évidence le rôle du facteur de qualité comme un curseur entre les approches temporelle et spectrale, en détaillant les quantités linéaires et quadratiques qui peuvent être récupérées à partir de cette représentation temps-fréquence. En passant des oscillations aux fluctuations, nous discutons ensuite l'approche probabiliste, et nous interprétons les quantités quadratiques comme des estimations statistiques à partir d'une réalisation unique d'une trajectoire stochastique. En particulier, l'analyse des corrélations est détaillée dans le temps et les fréquences, pour définir une cohérence temps-fréquence. La difficulté de définir un estimateur statistique temps-fréquence est exposé : abandonner de la résolution, en « relâchant » le principe d'incertitude d'Heisenberg, est la condition pour gagner des degrés de liberté statistiques. Plusieurs atomes d'information de la décomposition temps-fréquence doivent être recomposés pour constituer une estimation statistique signifiante. Dans le cadre des ondelettes log-normales, le problème se résorbe, ainsi que la variété de techniques existantes, dès lors que l'on réalise l'existence d'un estimateur canonique introduisant un second facteur de qualité. En distinguant le contrôle des résolutions en temps et en fréquence, nous laissons la place pour un nombre de degrés de liberté statistiques donné exactement par le rapport des facteurs de qualité. Pour mettre en pratique ces deux paramètres essentiels, nous proposons une stratégie d'approximation permettant une mise en œuvre numérique efficace.

Le chapitre II s'achève sur l'évaluation de la signifiance de l'estimateur de cohérence, à partir de la distribution de sa valeur résiduelle, ou cohérence fausse, associée à deux bruits

gaussiens indépendants. Nous argumentons et soutenons à partir de simulations le fait que cette distribution est étroitement approximée par une distribution bêta avec un unique paramètre. Précisément identifié à ce paramètre, le nombre de degrés de liberté permet de construire simplement des niveaux de signifiance pour la valeur observée de la cohérence entre deux enregistrements naturels. Ces niveaux permettent de construire une visualisation synthétique de la cohérence signifiante, à partir d'un codage couleur en teinte et en saturation pour les valeurs complexes de cette image temps-fréquence. L'outil qui en résulte est illustré avec des enregistrements polysomnographiques : électroencéphalogramme (EEG), électrocardiogramme (ECG), flux d'air (AF) et substitut.

Nous abordons dans le chapitre III la question de l'estimation des taux fluctuants pour des oscillations rythmiques modulées. Les caractéristiques temporelle et spectrale de ces rythmes sont d'abord détaillées à travers l'exemple de la voix. Sa structure horizontale en temps et ses ordres harmoniques résolus en fréquence peuvent être précisément anticipés à partir du nombre d'oscillations dans l'ondelette et sa bande passante, c'est-à-dire à partir de son facteur de qualité. Après avoir introduit et discuté les crêtes et les méthodes de réallocation, nous définissons une famille générique d'estimateurs de taux complexe, capturant à la fois les modulations de l'amplitude et de la fréquence instantanée, dans un domaine de fréquence choisi. En accord avec le cadre précédemment établi, les estimateurs basés sur des expressions quadratiques se trouvent être plus stables, en rapport avec leur interprétation statistique. Ils sont illustrés dans les limites d'une bande spectrale étroite et large pour un enregistrement vocal, achevant ainsi la partie A.

Basée sur des sources existantes de données cliniques, en particulier 2.3 ans (en cumulé) de polysomnographies obtenues à partir de 2650 enregistrements individuels, la partie B applique le langage développé dans la partie A à l'étude des signaux physiologiques.

Consacré aux rythmes cardio-respiratoires, le chapitre IV commence avec l'activité cardiaque. Nous analysons sa variabilité temporelle telle qu'enregistrée par les très différents électrocardiogramme (ECG) et photopléthysmogramme (PPG). Nous montrons comment utiliser les variables de localisation et les facteurs de qualité pour extraire la variabilité du taux cardiaque à partir des estimateurs de taux basés sur les ondelettes, du plus simple et générique (bande large) aux plus sophistiqués (bande étroite). De plus, nous remplaçons l'ECG par un signal dérivé d'amplitude, toujours dans le cadre des ondelettes, permettant ainsi d'améliorer l'estimation en exploitant une spécificité de l'ECG. Au moyen d'une analyse de la cohérence, ces estimateurs des modulations de la fréquence cardiaque sont comparés à un algorithme d'extraction du taux cardiaque conventionnel basé sur l'ECG.

Nous nous tournons ensuite vers le rythme respiratoire, enregistré comme un flux d'air avec un capteur thermoélectrique. Il module le taux cardiaque en un réflexe appelé arythmie sinusale respiratoire (RSA), fournissant le premier exemple d'interaction entre organes visible dans ces fluctuations. Sur un enregistrement individuel, nous montrons comment la force et la variabilité de cette interaction peut être décrite par les corrélations temporelles ou par la cohérence spectrale, puis détaillée comme la cohérence temps-fréquence entre le taux cardiaque estimé et l'enregistrement respiratoire. En plus de l'interaction RSA, une composante cohérente plus lente apparait, confirmée par une statistique globale sur la plus large base de données polysomnographiques. En raison d'un filtre instrumental, la bande basse fréquence de l'enregistrement respiratoire est limitée et la cohérence qui en résulte est dégradée. Les modulations de fréquence et d'amplitude portées par les oscillations respiratoires n'ont pas ces limitations à basse fréquence et sont donc utilisées pour l'analyse de la cohérence. La statistique globale de la cohérence cardio-respiratoire entre les signaux de taux est calculée et une troisième composante très lente est identifiée. La distribution de la phase entre modulations cardio-respiratoires est aussi obtenue à partir de l'analyse de la cohérence.

À la fin du chapitre IV, nous mettons au point une approche quantitative de la synchroni-

sation des rythmes cardiaque et respiratoire, prenant en compte la valeur de leur taux, au-delà de l'analyse de la cohérence entre leurs modulations. Elle consiste en la construction d'une distribution de leurs rapports de fréquences, à chaque instant, concentrée en un simple nombre rationnel pendant la synchronisation, sinon fluctuant librement, voire oscillant durant l'apnée du sommeil, comme illustré par la polysomnographie de deux sujets. Cette méthode constitue un nouvel outil pour étudier le quotient entre les taux cardiaque et respiratoire.

Le chapitre V introduit l'activité nerveuse, enregistrée par EEG, et les phénomènes du sommeil. D'habitude décomposée en un ensemble empirique de bandes spectrales, nous illustrons en un continuum l'intensité EEG, multi-échelle et évolutive. Les motifs provenant à la fois de la densité de puissance et de la cohérence inter-EEG procurent un aperçu direct des stades du sommeil et de l'éveil et peuvent simplifier la description de l'architecture globale du sommeil. La structure fine temps-fréquence de l'EEG est aussi capturée à l'aide d'un choix minutieux des facteurs de qualité, permettant l'extraction des modulations de la puissance ou magnitude dans chaque bande.

En rassemblant les signaux de taux cardiaque et respiratoire extraits ainsi que les signaux nerveux, nous étudions la modification de leur couplage lors de l'apnée du sommeil. L'analyse de leur cohérence temps-fréquence signifiante est illustrée et calculée individuellement, puis leur statistique collective est conditionnée à l'observation clinique d'apnée du sommeil obstructive ou centrale, d'hypopnée ou d'absence d'apnée. En plus d'un couplage RSA plus faible dans la paire cardio-respiratoire, nous rapportons l'omniprésence d'une lente modulation rythmique associée à différents types d'apnée, augmentant la cohérence de chaque paire de signaux à la fréquence correspondante. La phase des lentes modulations nerveuses est détaillée pour chaque groupe apnéique sur le continuum des bandes EEG, mettant en évidence des spécificités à chaque bande, telles qu'un chirp linéaire dans la bande δ. C'est une illustration claire de la manière étroite dont s'entrecroisent les régulations nerveuse et cardio-respiratoire. Leur altération durant les stades du sommeil et l'apnée est finalement esquissée dans une étude du couplage amplitude-phase des fuseaux du sommeil et des ondes lentes. Ce couplage est décrit comme une cohérence intra-EEG et comparé au quotient des taux cardiaque et respiratoire.

Dans la conclusion, nous détaillons des perspectives pour ce présent travail visant au développement d'une physique statistique des signaux physiologiques.

Dans une partie C en bonus, nous explorons la question musicale de la rationalité (ou l'irrationalité) de la distribution des ratios mise au point au chapitre IV, en introduisant un indice de synchronisation multi-fréquence que nous appelons la *sonance*, basée sur un ensemble statistique défini sur les nombres rationnels.

Conclusion

Un outil analytique pour aider à l'observation de la dynamique des organismes vivants a été développé dans cette thèse, et appliqué à l'étude des interactions entre les systèmes respiratoire, cardiaque et nerveux basée sur des enregistrements polysomnographiques. Le développement de cette thèse est résumé ci-dessous et des perspectives sont ensuite esquissées.

Résumé

Partie A

Dans la première partie, nous avons exposé une manière de déchiffrer des enregistrements naturels complexes par la caractérisation de leur contenu à la fois oscillant et fluctuant. Cela a été réalisé en représentant les enregistrements dans le temps et les fréquences, avant d'effectuer des estimations statistiques dans ces deux dimensions grâce à une combinaison de décomposition et recomposition d'ondelettes.

Nous avons commencé avec l'approche oscillatoire des enregistrements, en particulier avec l'intuition temporelle associée aux signaux oscillants : une onde évolue dans le temps avec une certaine amplitude et une phase parcourant un cycle en une certaine période. La version analytique du signal, une représentation complexe « en hélice » très adaptée d'une oscillation avec son rayon et son angle, est d'un usage limité lorsque des oscillations à de multiples échelles de temps sont superposées. À partir du prolongement analytique du signal dans le plan complexe, la séparation de composantes à des échelles distinctes mène naturellement à la transformée continue en ondelettes basée sur une ondelette analytique, translatée et dilatée dans le temps. Cette décomposition temps-échelle a la spécificité d'être indépendante d'une échelle de référence, contrairement au paradigme alternatif de la transformée de Gabor ou de Fourier à court terme. Cette propriété est particulièrement intéressante pour analyser de multiples événements oscillants simultanément, *chacun depuis sa propre échelle*, possiblement séparée des autres par plusieurs décades.

Pourtant, la description de Fourier des périodicités dans l'enregistrement, de par son spectre, est indispensable ; la variable d'échelle de l'ondelette n'apparaît que comme le reflet du domaine des fréquences. Pour retrouver cette intuition spectrale complémentaire, nous avons revisité le choix de l'ondelette analytique. Une fois son spectre positif standardisé en une forme sans dimension, la précision de la correspondance entre échelle et fréquence est réduite à un paramètre essentiel, son facteur de qualité Q. Proportionnel au nombre d'oscillations de l'ondelette et à l'inverse de la bande passante, il fixe un compromis sur les résolutions, qui partage l'atome d'incertitude temps-fréquence. En d'autres termes, le facteur de qualité est un curseur entre les interprétations temporelle et fréquentielle. Pour cette dernière, un contrôle plus fin de la forme de l'ondelette est superflu tant que le facteur de qualité est assez élevé. La fonction log-normale, un cas limite central dans la famille générale des ondelettes de Morse, se distingue par sa régularité, avec une décroissance plus rapide que polynomiale dans les deux domaines, et par sa symétrie, cristallisant ainsi l'idée d'une échelle logarithmique des fréquences.

Signal et spectre sont récupérés à partir de la transformée en ondelettes de deux manières équivalentes : d'une part comme des marginales en temps et en fréquence, d'autre part comme des asymptotes opposées du facteur de qualité. L'approche probabiliste des fluctuations est ensuite introduite, en considérant l'enregistrement comme la réalisation d'un processus stochastique. L'application d'une des opérations précédentes à une représentation quadratique revient à une estimation statistique basée sur les ondelettes : pour une seule trajectoire, la variance dépendante du temps et la densité spectrale, ou pour un produit de trajectoires, le coefficient de corrélation décliné comme une cohérence soit temporelle, soit spectrale. Leur localisation à la fois dans le temps et les fréquences requiert cependant un second paramètre pour estimer

l'espérance statistique, lequel était identifié à un nombre effectif de « tapers » orthogonaux (ici des ondelettes) interprétés comme des degrés de liberté statistiques. L'approche alternative, mais tout aussi efficace, consiste à réduire la résolution temps-fréquence par une opération de lissage.

Dans le cadre des ondelettes log-normales, ces points de vue probabiliste et temps-fréquence se rencontrent en un estimateur statistique temps-fréquence que nous qualifions de canonique en raison de son rôle très spécial. L'introduction de plusieurs degrés de liberté dans l'estimation statistique coïncide avec la distinction de deux facteurs de qualité différents dans l'analyse temps-fréquence : $n_d = \frac{Q_+}{Q_-}$, le plus élevé contrôlant la résolution temporelle, indépendamment du moins élevé pour la résolution en log-fréquence. Cette correspondance est un pont entre deux points de vue sur l'information, équilibrant signifiance et résolution. La première étape, de décomposition des oscillations en atomes temps-fréquence, est suivie de la seconde étape, de recomposition de n_d atomes d'incertitude pour plus de stabilité contre les fluctuations.

Nous avons dû faire deux compromis, un de chaque côté de ce pont, pour mettre en pratique ces deux paramètres essentiels. Le premier est la mise en œuvre numérique de l'estimateur canonique grâce à un lissage temporel gaussien, que nous montrons être l'approximation de Laplace correcte au premier ordre. Le second est la prédiction simplifiée des niveaux de signifiance statistique attendus pour l'estimateur canonique de cohérence. Nous motivons et vérifions, à partir de la simulation de bruits gaussiens stationnaires et indépendants, que la distribution de la fausse cohérence au carré est presque la loi bêta. La densité de puissance et la cohérence sont respectivement une variance locale et un coefficient de corrélation, détaillés pour toute translation et dilatation du temps. À partir des niveaux de signifiance de la cohérence, nous construisons un codage couleur synthétique par la teinte et la saturation pour l'angle et le module des régions temps-fréquence de cohérence signifiante.

Le problème de l'estimation des modulations d'un rythme fluctuant est finalement introduit. La forme d'onde non-circulaire répétée est décomposée par la transformée en ondelettes en structures horizontales et verticales, interprétées comme des harmoniques jusqu'à un certain ordre $\sim \frac{Q}{\sqrt{\uparrow}}$ et comme des battements au-delà. Les cycles de la phase (et l'amplitude variable) des lignes harmoniques, en particulier ceux du premier ordre (fondamental), sont reliés au taux recherché par une dérivée temporelle, qui se trouve être équivalente à une dilatation en fréquence. Une certaine flexibilité existe dans la sélection en fréquence des cycles de phase, en relation aux concepts de crêtes horizontales stables et instables et de réallocation en fréquence, de sorte que nous postulons l'existence d'une bande d'intérêt en fréquence. Les modulations dans cette bande peuvent être extraites par une famille générique d'estimateurs de taux. En particulier, une version quadratique est conforme à l'interprétation statistique : l'introduction d'un second facteur de qualité (en relation avec la bande passante) assure la stabilité du taux estimé. Pour un enregistrement vocal, les différents résultats, dans les limites à bande étroite ou large, de ce troisième type d'estimateur basé sur les ondelettes, donnent une estimation de la hauteur et du vibrato. L'incertitude sur l'estimation de la hauteur est exprimée en termes de cohérence.

Partie B

La seconde partie est une application à l'analyse des activités cardiaque, respiratoire et nerveuse, capturées par les signaux ECG ou PPG, AF et EEG. Les données physiologiques, polysomnographiques pour la plupart, sont issues de bases de données existantes. Nous nous efforçons de traduire leur traitement en observables biophysiques d'intérêt en utilisant le langage développé dans la première partie.

Nous commençons avec la variabilité du taux cardiaque, extraite du rythme cardiaque comme sa modulation en fréquence. Malgré les très différentes techniques d'enregistrement, le

taux cardiaque peut être estimé précisément avec des estimateurs temps-fréquence génériques du taux fluctuant, dans une bande en fréquence pertinente. Nous montrons que la version large bande de l'estimateur de taux, qui ne nécessite qu'un unique paramètre, est suffisante pour extraire les modulations de la fréquence cardiaque, même si elle est biaisée, particulièrement pour les oscillations fortement non-circulaires de l'ECG. L'estimation de taux cardiaque la plus précise est obtenue en utilisant l'estimateur de taux à bande étroite sur un nouveau signal d'amplitude, dérivé de la transformée en ondelettes du signal ECG à 14 Hz (aECG14). Les estimateurs sont appliqués aux signaux PPG, ECG et aECG14 de tous les sujets dans une base de données. La précision des modulations de la fréquence cardiaque est évaluée en termes de cohérence avec un taux cardiaque de référence estimé à partir des intervalles R-R de l'ECG. Cette cohérence est signifiante dans tous les cas et pour toutes les fréquences de modulation, et elle augmente avec la circularité du signal cardiaque d'entrée et avec la spécificité de l'estimateur.

Signe de l'interaction cardio-respiratoire, la variabilité du taux cardiaque est corrélée à la respiration, introduite avec le signal AF (flux d'air). En comparant ce dernier au taux cardiaque, d'abord dans le domaine du temps ou des fréquences, ensuite dans le plan temps-fréquence, nous observons deux composantes distinctes de la cohérence : l'une rapide à la fréquence fondamentale du rythme respiratoire, connue comme l'arythmie sinusale respiratoire (RSA), et l'autre plus lente et moins rythmique (avec une bande plus large). Ces interactions sont caractérisées par leur phase et module de cohérence, à la fois individuellement et collectivement. La signifiance statistique est évaluée pour l'estimation temps-fréquence individuelle, et l'estimation collective est calculée depuis un second niveau de statistique sur tout temps et pour tous les sujets dans une large base de données. En particulier, l'étendue des fluctuations de la différence de phase entre le taux cardiaque et les oscillations respiratoires est résumée en une densité angulaire de cohérence. En reproduisant cette méthode avec un autre type de signal respiratoire et en comparant les résultats, nous avons mis en évidence des difficultés d'interprétation liées aux limitations instrumentales, visibles dans la réduction de la cohérence et l'incertitude de la phase des oscillations lentes causées par un filtre passe-haut.

L'étude des basses fréquences est facilitée par l'usage de signaux de taux respiratoire, plutôt que l'enregistrement brut, obtenus en extrayant l'activité lente contenue dans ses modulations de fréquence (FM) et d'amplitude (AM). Suffisante à l'étude de la cohérence du rythme cardiaque, l'estimateur de taux générique à bande large est particulièrement approprié pour les oscillations plutôt circulaires de la respiration, dont la période peut beaucoup fluctuer. L'analyse de la cohérence souligne le fait que les FM et AM respiratoires lentes évoluent presque en phase (corrélées positivement) avec les modulations du taux (FM) cardiaque, mais en opposition de phase (anti-corrélation) pour les FM respiratoires plus lentes que 0.01 Hz. Cela illustre la présence d'un mode distinct de régulation cardio-respiratoire, très lent, antagoniste au mode lent.

L'analyse de la cohérence temps-fréquence des FM cardiaques et respiratoires se limite à dépeindre la similarité de leur coévolution. En particulier, elle ne peut pas tenir compte de leur synchronisation à un rapport de fréquence entier, ou caractériser l'évolution du rapport des taux cardiaque et respiratoire. Dans le cadre des ondelettes, nous construisons une distribution de ce rapport de fréquence dépendant du temps, détaillée à une grande résolution dans le temps et à une concentration spectrale précise grâce à une réallocation en fréquence et un faible facteur de qualité. En l'appliquant à une polysomnographie, nous avons illustré individuellement sa capacité à suivre les phénomènes du sommeil, tels que l'apnée sous la forme d'une oscillation régulière du ratio, le stade REM avec ses fortes fluctuations et le sommeil profond avec son ratio stable.

L'observation d'épisodes nets de synchronisation, aussi appelé couplage cardio-respiratoire, et l'analogie musicale avec deux voix chantant un accord consonant ont motivé la formulation d'un indice de synchronisation multi-fréquence. Appelé la sonance, il mesure la rationalité ou

l'irrationalité dans la distribution des ratios. Son développement, actuellement limité à une application musicale, est approfondi dans une partie bonus distincte.

Troisième membre important de l'orchestre physiologique, l'activité du système nerveux est introduite telle qu'enregistrée par l'EEG dans une polysomnographie. Une représentation à facteur de qualité constant est spécialement bien adaptée pour observer des structures multi-échelles dans plusieurs bandes et les stades temporels d'éveil et de sommeil. En fonction du phénomène investigué, la sélection d'un unique ou d'une paire de facteurs de qualité est discutée. $Q_- = 5$ est un minimum pour la distinction spectrale des ondes cérébrales, fixant ainsi une résolution temporelle maximum qui peut être diminuée, soit légèrement ($Q_+ \approx Q_-$) pour définir un champ de magnitude nerveuse stable, soit radicalement pour obtenir une représentation synthétique des motifs de la densité liés à l'éveil et au sommeil ou pour obtenir une estimation très signifiante de la cohérence temps-fréquence. Pour $Q_+ = 500$, nous illustrons des motifs de cohérence inter-EEG, entre enregistrements contra-latéraux, avec un signe et un module qui alternent (d'incohérent à très cohérent), variant indépendamment de la densité de puissance. Les modulations des taux sont reliées aux modulations de la puissance et de la magnitude dans chaque bande EEG ; ces dernières sont préférées comme observables nerveuses.

Pour évaluer les interactions physiologiques, nous rassemblons les taux cardio-respiratoires et les signaux de magnitude nerveuse dans des bandes distinctes, et nous les comparons par paires au moyen de la cohérence temps-fréquence. Celle-ci est appliquée au cas d'un sujet affecté sévèrement par de l'apnée du sommeil obstructive ; nous identifions la signature de l'apnée comme une composante très cohérente et à bande étroite présente dans toutes les paires entre 0.01 et 0.04 Hz. Une analyse approfondie de la cohérence des modulations neuro-cardio-respiratoires est réalisée individuellement pour chacun des 2650 sujets d'une large base de données polysomnographiques. Ensuite, des spectres de cohérence typiques sont calculés par des moyennes temporelles conditionnelles, sur 2.3 ans en cumulé d'enregistrement du sommeil, en les divisant en une sélection d'intervalles et en sous-groupes sur la base des annotations cliniques d'événements apnéiques des différents types : hypopnée, apnée obstructive ou centrale, ou absence d'apnée. Les modulations apnéiques sont caractérisées par une amplitude et une cohérence augmentée, apparaissant comme un pique spectral en comparaison à la ligne de base plate sans apnée ou au profil intermédiaire de l'hypopnée. L'apnée est aussi concomitante à une importante réduction de la cohérence, non seulement au niveau de l'interaction cardio-respiratoire RSA rapide et à bande étroite, mais aussi pour la composante très lente et à bande large (irrégulière) neuro-respiratoire et cardio-respiratoire.

La phase des modulations nerveuses est ensuite comparée à celle des modulations cardio-respiratoires à travers les bandes EEG. Nous parvenons à une relation phase-fréquence spécifique et en grande partie conservée, à une phase globale près, pour tous les profils (avec et sans apnée), pour des modulations lentes ou très lentes et pour toutes les paires de cohérence neuro-cardiaque et neuro-respiratoire. Cette interaction est composée d'un chirp linéaire dans la bande δ, incohérente dans la bande θ, d'un écho du chirp dans la bande α-σ et d'une phase constante dans la bande β-γ.

À l'aide d'une polysomnographie individuelle, nous illustrons finalement les couplages phase-amplitude durant différents stades du sommeil, décrits comme une cohérence intra-EEG, en particulier celui entre les oscillations δ lentes et les fuseaux du sommeil de la bande σ pendant le sommeil profond. Nous présentons aussi une stratégie pour détecter les fuseaux du sommeil à partir de la phase dans la bande σ. Le rapport des fréquences instantanées du cœur et de la respiration complète le tableau nocturne de la séquence des stades du sommeil.

Discussion et perspectives

Il est maintenant clair que la description de la dynamique des systèmes vivants doit pouvoir faire la distinction entre une grande diversité de comportements. Le contenu du bruit physiologique spontané s'étend de fluctuations complètement stochastiques à des rythmes pulsatiles, apériodiques, quasi-périodiques et des oscillations régulières, qui peuvent se côtoyer à des fréquences distinctes et évoluer en une séquence de périodes temporelles de durées variables. Cette riche phénoménologie n'est pas seulement distribuée parmi les enregistrements simultanés, mais aussi au sein de chacun d'entre eux sous la forme de modulations.

La pratique de la physique, dans ce thème interdisciplinaire, a consisté à assurer la visibilité, le contrôle et la solidité des notions élémentaires de facteur de qualité, fréquence, temps, phase et amplitude des oscillations. Pour construire une physique statistique de ces signaux physiologiques, une perspective temps-fréquence est indispensable : ces phénomènes oscillatoires sont correctement décrits lorsque leur facteur de qualité intrinsèque correspond à celui utilisé pour la représentation, permettant ainsi leur localisation. Le découplage du temps et des fréquences revient à dissocier une description cohérente en deux points de vue extrêmes et déformés. L'adoption du point de vue intermédiaire adéquat revient à choisir la forme de l'atome d'incertitude, l'ondelette, dans la décomposition temps-fréquence.

L'outil probabiliste basique des corrélations par paires correspond à la recomposition des atomes de deux signaux distincts en une moyenne locale. Le coefficient circulaire (pour ne pas dire complexe) qui en résulte exprime leur cohérence. Cette information statistique est aussi précise que le nombre d'atomes est important, au détriment de l'information de localisation, qui est réduite d'autant. L'interprétation probabiliste est par conséquent permise par la séparation de longues échelles de temps d'évolution et de courtes échelles spectrales. Des paradigmes différents sont disponibles pour équilibrer ces deux types d'information, chacun amenant une géométrie particulière de la composition dans le plan temps-fréquence. Le formalisme de l'ondelette log-normale proposé dans cette thèse permet de suivre la forme et le nombre d'atomes dans la composition par un rapport de deux facteurs de qualité, sans introduire d'échelle arbitraire de référence.

Toutes les paires de signaux physiologiques forment un graphe complet, dont chaque lien est une image de cohérence temps-fréquence, se résumant à un coefficient de corrélation (circulaire) si l'information était purement statistique et non locale. La donnée de deux paramètres de contrôle (les facteurs de qualité) est suffisante pour dériver de ces signaux un prototype du réseau d'interactions physiologiques, dont la structure est en cours d'examen [Bartsch, Liu, Bashan, et al. 2015; Ivanov, Liu, and Bartsch 2016]. Nos procédés de visualisation au code couleur synthétique de ces relations temps-fréquence vont faciliter ces observations et peuvent être convenablement réduites en composantes principales.

Nous attendons d'une approche physique pleinement développée que soit spécifié la directionnalité des interactions. Autrement dit, la causalité doit être mise à jour dans ce réseau d'interactions, pour mettre en évidence la finesse des mécanismes d'inter-régulation entre les fonctions physiologiques. Réduite en sa forme la plus simple, un pur délai $e^{i\Upsilon f\tau}$, la causalité est codée dans la direction spectrale comme des cycles de la phase. Il devient possible de récupérer efficacement la causalité [Ephremidze, Saied, and Spitkovsky 2018] et son interprétation statistique est déjà bien étayée [Faes, Pernice, et al. 2021; Faes, Stramaglia, and Marinazzo 2017]. La caractérisation évolutive et multi-échelle des interactions causales dans l'orchestre physiologique est bientôt atteinte.

La prochaine étape vers une physique statistique des signaux physiologiques est plus lointaine, mais la question de leur irréversibilité est en vue [Arneodo, Argoul, et al. 1993] et le terrain est peut-être déjà en train d'être préparé [Sekimoto 2010].

Contents

Introduction

Life and its motions are natural phenomena, hence part of the φύσις (phýsis, as the natural Ecosystem); yet, the physical nature of living systems remains largely elusive. They share, with out-of-equilibrium and non-linear dynamical systems, the ubiquitous appearance of pulsatile and quasi-periodic behaviours. The physiological rhythms that occur in the heart, breathing, locomotion, neural signals, cellular or molecular clocks, hormone secretion, life cycle, etc, are spanning over 10 decades across temporal scales. Living organisms have a complex existence at numerous interconnected spatial and temporal scales, often not clearly separated. Apart from these multi-scale characteristics, a strongly unsteady and unpredictable activity, far from being stationary, is another essential property that distinguishes living from inert matter.

All these features are as many challenges addressed to biophysicists for finding a sound framework to describe living systems. Many repeated measurements are required to compensate their strong variability, from one system to another, but also within the same system, whose evolving behaviour cannot be reproduced on demand in many experimental situations. The continuous recording of the physiological activity of the human body is one of them, at the focus of this thesis. It is a single trial experiment: each measured trajectory is unique and ever changing, with regular oscillations inextricably mingled with irregular fluctuations and intermittent characteristics. Specific to the living system during the observation, this mixture of unpredictable and organized behaviours results from the interplays between underlying subsystems at many different scales, from biomolecules, cells, tissues, organs, to the full organism and its environment. Their reductionist and integrative description across scales into a single biophysical model is a formidable task, that may, in the end, remain an abstract idea.

Recognizing and unfolding rhythms is paramount in the physiology of the heart, respiration, metabolism, brain, hearing, voice and language, locomotion and proprioception... The biomedical motivations to decipher their complex signals are a strong incentive to assist the difficult modelling with an improved analysis of observations. The human body has been partitioned by clinicians into distinct physiological subsystems; among them, the finely interregulated respiratory, cardiac and nervous systems. Their interactions form a network, that is the field of highly complex dynamics. To gain better insights into these interactions, a crucial step is the observation of the physiological activity through simultaneous recordings, such as a polysomnography measured by different sensors on the body during the resting state. Their study concentrates the efforts of a large community of physiologists, neuroscientists, clinicians, biologists, physicists and mathematicians, aiming at improving health for persons affected by cardiac arrhythmia, sleep apnea, insomnia, certain psychic disorders, or simply aging, and possibly at getting a deeper understanding of phenomena associated to sleep and wake, attention, perception, emotions, and perhaps even consciousness.

The context of biophysical measures can be quite different from the one of classical physics experiments. Available recording methods are based on different physical techniques (electromagnetic, optical, thermal, barometric), which are sometimes redundant, for instance for the cardiac rhythm. Non-intrusive and passive recording methods have a practical advantage over more direct and precisely controlled measurements; the former are hence preferred for clinical or commercial applications. But they are also the ones that require the most processing to

extract the biophysical information of interest, such as rates or intensity modulations. The difficulty of signal or image processing tasks results in a profusion of sophisticated algorithms, engineered for each of the recording methods. The strong specificity of these techniques tends to oppose the genericity of mathematical relations between recorded and extracted biophysical observables. When trying to decipher the network of physiological interactions between them, this lack of a common ground gets more prominent. These considerations stress the importance in improving existing mathematical tools and numerical techniques of observation, some of which are explored and developed in this thesis.

Constructing a statistical physics' theory of these systems cannot be achieved without considering jointly time and frequency characteristics of their fluctuating and oscillating behaviours. To handle the complexity of physiological recordings, a new time-frequency representation of the signal appears as a crucial ingredient to expand and dilutes into a two-dimensional map its rich information, that was initially concentrated along a single temporal (or spectral) dimension. Multi-scale characteristics of the signals are depicted at each time in the spectral direction, and its evolutionary characteristics are depicted at each frequency in the temporal direction. A time-frequency representation employs a test function to scrutinize the oscillatory content of the analysed signal at different locations: either from translations in time and frequency (as the sliding window of the short-time Fourier transform), or from time translation and dilation (as in the wavelet transform, employed throughout this entire thesis). These two methods lead to apparently distinct interpretations, such as "frequency" and "scale", further diversified by the numerous possibilities for the test function (that include the signal itself) and their implementation. On one hand, this is a token of versatility in practical situations, that has been translated into numerous scientific applications. On the other hand, this is an additional profusion of technicalities to be handled, sometimes hiding relations with more fundamental concepts and intuitions.

The probabilistic interpretation of fluctuations in natural recordings is the second ingredient. Stochastic processes take an increasingly important place in the physical modelling of fluctuating trajectories in micro and mesoscopic phenomena, whereas statistical inference methods are often more suited for the complexity of macroscopic systems. These approaches share common information-theoretic ideas. The definition of statistical ensembles, both within and among physiological recordings, is a delicate task, aiming at extracting robust observations from these individual and single trial measurements with non-stationary and multi-scale characteristics. These conditions often limit estimations to the first and second order statistics, such as the analysis of correlations, and the evaluation of their significance is essential. Although elementary, the correlation analysis is effective to assess connectivity and interactions between quantities of interest, that may be of very different physical nature.

The statistical approach of correlations has been combined with a time-frequency perspective into a composite method called the time-frequency coherence. The data from two simultaneous recordings is decomposed at different times and scales prior to their statistical analysis. This relatively recent tool of investigation for complex dynamics proves efficient and has found many applications in physiological, ecological, geophysical, economical contexts and more. Here again, this analysis has been implemented in many flavours and refinements, with varying degrees of control over the time-frequency resolution and the statistical significance. Furthermore, it often relies on the outcome of systematically preprocessed signals, that have undergone various filtering, whitening, detrending procedures, left out of the framework of the coherence analysis. Eventually, the relevant dynamical quantities are rarely the one directly measured. For instance, the heart and breathing rate signals or neural waves are systematically extracted, by algorithmic means, from the physiological recordings that contain all the original information.

The central issue underlying this thesis can be expressed as follows: is there a common

language able to handle the different tasks, from the processing of the bare physiological recordings into new biophysical observables, to the study of their interactions? What are the relevant parameters to achieve this analysis in a generic way? In an effort of synthesis, we establish a generic framework upon a wavelet-based time-frequency representation of the recordings, with care for a physicist interpretation. This tool is applied to the study of polysomnographic recordings of the cardiac, respiratory and neural activities, and their rich phenomenology.

The construction of this framework is presented in the part A of the thesis, composed of chapters I, II and III, followed by an application to physiological recordings in the part B, composed of chapters IV and V. Their distinct bibliography is introduced progressively in the corresponding chapters.

In chapter I, we start examining the temporal and spectral representation of oscillations, in a real signal and its spectrum. Elaborating upon the convenience of the analytic or complex helical version of the signal to interpret oscillations in time, we show how the idea of distinguishing contributions at multiple scales naturally leads to wavelet transform with an analytic wavelet. The wavelet analysis amount to decomposing a signal as time-scale "atoms", called wavelets, chosen adequately to access relevant information. In particular, we retrieve the spectral intuition of the Fourier transform by examining the shape of the wavelet spectrum. We revisit the general family of Morse analytic wavelets, so as to make the only relevant parameter –the quality factor– explicit and adjustable. We find that the log-normal shape, limit case in the Morse family, is ideal in this respect, with exceptional symmetry, localization and regularity. Described by R.A. Altes, and proposed by A. Grossmann at the very beginning of the wavelet theory as a "particularly good" one, this shape has remained paradoxically quasi-absent from the downstream literature. The properties of the log-normal wavelet are detailed, together with the resolution in time and frequency of the associated representation, characterized by an uncertainty trade-off explicitly controlled by the unique parameter, the quality factor.

We continue in chapter II with a simplification of the notation of this wavelet representation of the recording, the illustration of the amplitude and phase of oscillations in the time-frequency continuum, geometrical and numerical considerations. We highlight the role of the quality factor as a cursor between the time and the frequency approaches, by detailing the linear and quadratic quantities that can be retrieved from this time-frequency representation. Sliding from oscillations to fluctuations, the probabilist approach is then discussed, and the quadratic quantities interpreted as statistical estimations on a single realization of a stochastic trajectory. In particular, the analysis of correlations is detailed in time and in frequency, to define a time-frequency coherence. The problem of defining statistical time-frequency estimators is exposed: giving up some resolution by "relaxing" the Heisenberg uncertainty principle is the condition to gain statistical degrees of freedom. Several atoms of information from the time-frequency decomposition must be recomposed to constitute a significant statistical estimation. In the log-normal wavelet framework, the problem resorbs itself, and the various existing techniques with it, by realizing the existence of a canonical estimator that introduces a second quality factor. By distinguishing the control of the time and frequency resolutions, we open room for statistical degrees of freedom, counted exactly by the ratio of the quality factors. In order to put these two essential parameters in practice, we propose an approximation strategy for an efficient numerical implementation.

Chapter II ends on the evaluation of the significance of the coherence estimator, from the distribution of its residual value, or spurious coherence, associated to independent Gaussian noises. We argue and support from simulations that this distribution is closely approximated by a single-parameter beta distribution. Accurately identified with this parameter, the number of degrees of freedom allows a straightforward construction of significance levels for the observed values of the coherence between two natural recordings. These levels allow to build a synthetic visualization of the significant coherence, from a hue-saturation colour-coding of its complex-

valued time-frequency map. The resulting tool is illustrated on polysomnography recordings: electroencephalogram (EEG), electrocardiogram (ECG), airflow (AF) and surrogate.

We address in chapter III the question of estimating fluctuating rates for modulated rhythmic oscillations. The temporal and spectral characteristics of these rhythms are first detailed through the example of the voice. Its horizontal structure and resolved harmonic orders in time and frequency can be precisely anticipated from the number of oscillations in the wavelet and its bandwidth, i.e. from its quality factor. After introducing and discussing ridges and reassignment methods, we define a generic family of complex rate estimators, capturing both the modulations of the instantaneous frequency and amplitude, in a frequency range of interest. Consistent with the previously developed framework, estimators based on quadratic expressions prove more stable, in relation to their statistical interpretation. They are illustrated in the narrow-band and wide-band limits for a voice recording, ending part A.

Based on existing sources of clinical data, in particular 2.3 years of cumulated polysomnography from 2650 individual recordings, Part B applies to the study of physiological signals the language developed in part A.

Chapter IV, focusing on cardio-respiratory rhythms, starts with the cardiac one. We analyse its temporal variability as recorded by the very different electrocardiograms (ECG) and photoplethysmograms (PPG). We show how to utilize the localization variables and the quality factors to extract the heart rate variability from the wavelet rate estimators, from the most simple and generic (wide-band) to more refined ones (narrow-band). In addition, we show how replacing the input ECG by a derived amplitude signal, still in the wavelet framework, allows to improve the estimation by exploiting a specificity of the ECG. These estimators for the cardiac frequency modulations are compared to a conventional ECG-based heart rate extraction algorithm, by means of a coherence analysis.

The breathing rhythm follows, recorded as an airflow with a thermoelectric sensor. It modulates the heart rate, in a reflex called respiratory sinus "arrhythmia" (RSA), providing a first example of organs interaction visible in these fluctuations. We show on an individual recording how the strength and the variability of this interaction can be described from temporal correlations and spectral coherence, and is eventually detailed as the time-frequency coherence between the estimated heart rate and the respiratory recording. In addition to the RSA interaction, a slower coherent component appears, confirmed from a global statistics on the largest polysomnographic database. Due to instrumental filtering, the low frequency range of the respiratory recording is limited, and the resulting coherence degraded. The frequency and amplitude modulations, carried by the breathing oscillations, have not such limitation at low frequency, hence used for the coherence analysis. The global statistics for the cardio-respiratory coherence between rate signals is computed, and a third very slow coherent component is identified. The distribution of the phase between cardio-respiratory modulations is also obtained from the coherence analysis.

In the end of Chapter IV, we devise a quantitative approach of synchronization between the cardiac and respiratory rhythms, taking into account their actual rate values, beyond the analysis of coherence between their modulations. It consists in constructing their instantaneous distribution of frequency ratios, concentrated on a simple rational number during synchronization, freely fluctuating otherwise or even oscillating during sleep apnea, as illustrated from the polysomnography of two subjects. This method constitutes a new tool to study the pulse-respiration quotient.

Chapter V introduces the neural activity, recorded from EEG, and sleep phenomena. Usually decomposed into an empirical set of spectral bands, we illustrate continuously the multiscale and time-evolving EEG intensity. Both the power density and the inter-EEG coherence patterns provide a direct insight into the wake-sleep stages, and can simplify the description of the global sleep architecture. The fine time-frequency structure of the EEG is also captured

from a careful selection of quality factors, allowing the extraction of power and magnitude modulations in each band.

Gathering the extracted cardiac and respiratory rate signals as well as neural signals, we study the modification of their coupling when sleep apnea occurs. The analysis of their significant time-frequency coherence is illustrated and computed individually, before conditioning their collective statistics to the clinical observation of obstructive and centra sleep apnea, hypopnea or no apnea. In addition to a weaker RSA coupling in the cardio-respiratory pair, we report the ubiquitous appearance of a slow rhythmic modulation associated to the different types of apnea, enhancing coherence in all pairs at the corresponding frequency. The phase of slow neural modulations is detailed for each apnea group on the continuum of EEG band frequency, uncovering band-specific characteristics, such as a linear chirp in the δ band. This is a clear illustration of the tight intertwining of neural and cardio-respiratory regulations. Their alteration during sleep stages and apnea is finally outlined in a study of the amplitude-phase coupling of sleep spindles and slow oscillations. It is described as an intra-EEG coherence and compared to the pulse-respiration quotient.

In the conclusion, we finally detail perspectives for the present work towards the development of a statistical physics of physiological signals.

In a bonus part C, we explore the musical question of quantifying the (ir)rationality of the ratio distribution raised in chapter IV, by introducing a multi-frequency synchronization index that we call *sonance*, based on the definition of a statistical ensemble over rational numbers.

Part A

Oscillations and fluctuations with one foot in time, the other in frequency?

Chapter I

Time-frequency representation: the way of the analytic wavelet

Jumping back and forth between the time and frequency domains (or spatial position and frequency), using a spectral (Fourier or Laplace) transform, is a fundamental skill in Physics to describe and understand a huge diversity of phenomena.

Although the assumptions —of periodicity, localization, stationarity, etc— associated to this tool can face strong difficulties when confronted to the complexity of living systems, most of the intuition and concepts can still bring useful insights. In this chapter, we start by recalling the terms and notions used in the context of the temporal evolution of an observable, before showing how they have steered the physicist community to a certain use of the wavelet transform, in order to capture both its multi-scale and time-dependent characteristics.

1 Anatomy of a real signal: spectrum and analytic version

In general, natural recordings are real signals $x(t)$, possibly multi-dimensional, evolving as a function of time. Therefore, what we call signal is the time-domain representation of the recording. We call its frequency-domain representation the spectrum, $\hat{x}(f)$, related to $x(t)$ by the Fourier transform and its inverse, defined here as:

$$\hat{x}(f) = \int x(t) e^{-i\tau ft} \mathrm{d}t \qquad ; \qquad x(t) = \int \hat{x}(f) e^{i\tau ft} \mathrm{d}f \quad . \tag{I.1}$$

The spectrum can be understood as a projection of the time-localized values of x onto a continuous basis of pure complex waves $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, whose sinusoidal oscillations have the frequency f. The angle $\theta = \tau ft$ is the phase of the oscillation, where ft is a number of cycles and

$$\tau = 2\pi \approx 6.283 \qquad ; \qquad e^{i\tau} = 1$$

is the natural angular unit [Harremoës 2012; Hartl 2010; Palais 2001; White 2015], measuring the tour of a unit-radius circle (π is half of it). In general, the spectrum takes complex values, that can be decomposed as a modulus, given by the absolute value $|\hat{x}(f)|$, and an angle, the complex argument $\arg \hat{x}(f) = \Im\{\log \hat{x}(f)\}$, so that $\hat{x}(f) = |\hat{x}(f)| e^{i \arg \hat{x}(f)}$.

1.1 Time and frequency: similar variables, different treatments

In spite of the formally symmetric situation between the time variable $t \in \mathbb{R}$ and the frequency one $f \in \mathbb{R}$, they often have a very different physical nature. Both the time and the frequency

are counted as multiples of a *unit scale*. Their physical dimensions are inversely related, for instance the second (s) is the standard time unit and the Hertz ($\mathrm{Hz} = \mathrm{s}^{-1}$) the corresponding frequency unit. However, the time variable is also counted starting from an initial time t_0, a reference often arbitrarily set to zero: $t_0 + t = t$, so that the time variable is a *relative* quantity. In contrast, the frequency variable is an *absolute* quantity (the reference $f = 0$ is absolute) since f unambiguously relates to the duration $\tau = f^{-1}$ spent between two oscillations, the period. This is an example of a temporal quantity which is absolute (as the difference between two relative times) and can be used to characterize a system (its periodicity, delay, growth or decay time, etc); we use a distinct notation τ for these characteristic times, or *timescales*, often taken as natural units when they appear in a physical system.

1.2 Real or analytic signal, same spectrum

Although the spectrum is a function of an absolute quantity, the possibility to shift the initial time of the signal translates into an arbitrary angle: $\hat{x}(f) \to \hat{x}(f)e^{i\Upsilon f t_0}$; the modulus is not affected. The real value of the signal corresponds to the (Hermitian) symmetry of its spectrum, $\hat{x}(-f) = \overline{\hat{x}(f)}$, where the bar denotes the complex conjugation. This means that the information in the spectrum is redundant, due to complex values holding twice as much information as real ones. Therefore, the negative support of the frequencies can be safely discarded without loss of information, yielding in the time domain a complex helical version of the signal, called the *analytic signal*:

$$x^+(t) = \int \Theta(f)\hat{x}(f)e^{i\Upsilon f t}\mathrm{d}f = \int_0^\infty \hat{x}(f)e^{i\Upsilon f t}\mathrm{d}f \quad , \tag{I.2}$$

where the (Heaviside) step function Θ filters out negative frequency components and halves the constant one:

$$\Theta(f) = \begin{cases} 0\,, & f < 0 \\ \frac{1}{2}\,, & f = 0 \\ 1\,, & f > 0 \end{cases}$$

This restriction to positive frequencies essentially turns the Fourier transform into a Laplace transform of imaginary parameter (second equality, with special care at $f = 0$). Why bother with this complex version while we could stay simple and real? This complex signal in its polar form is actually most convenient to study modulated waves, and can be returned to the real territory by dropping the imaginary part, i.e. adding the conjugate of x^+, x^-:

$$x(t) = x^-(t) + x^+(t) = 2\Re\{x^+(t)\} \quad . \tag{I.3}$$

The relation between the real and imaginary parts of $x^+(t)$ is called the Hilbert transform, also known as the Kramers-Kronig relations.

For a complex signal the complex helical terms in the decomposition $x^-(t) + x^+(t)$ are no more complex conjugates; x^- is called the *anti-analytic* signal defined from negative frequencies using $\Theta(-f)$ in Eq.(I.2).

1.3 Example of the complex wave model

For instance, the real wave of amplitude A and frequency f, $x(t) = A\cos(\Upsilon f t) = \frac{A}{2}(e^{i\Upsilon f t} + e^{-i\Upsilon f t})$ has for analytic version the complex wave $x^+(t) = \frac{A}{2}e^{i\Upsilon f t}$, from which the phase at each time and the amplitude are conveniently obtained as the angle and twice the modulus. In general, the so-called canonical amplitude and phase $A(t) = 2|x^+(t)|$ and $\phi(t) = \arg x^+(t)$ are useful to bring a

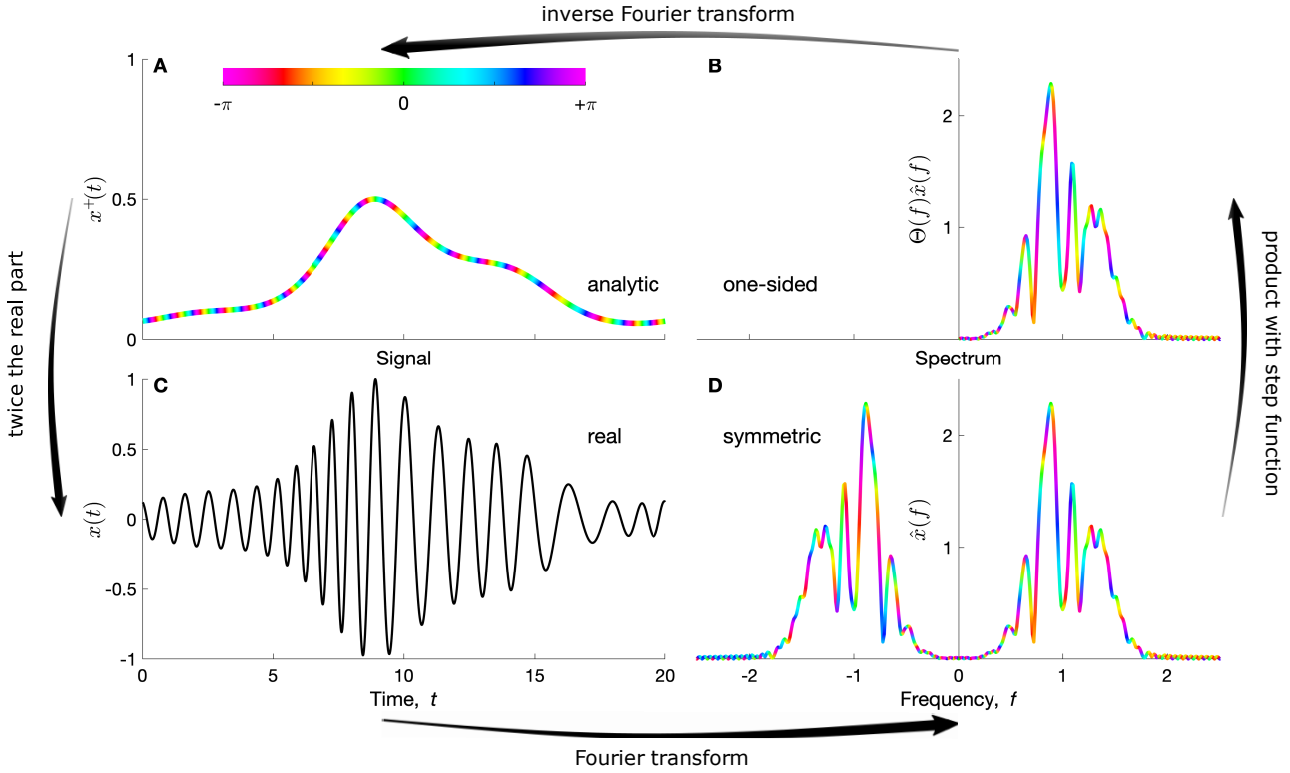


Figure I.1: *Equivalent representations for a 1D recording x. This signal is defined from Eq.*(I.5) *with f_0, r_1, f_1, r_2, f_2 equal to 1, 0.3, 0.05, 0.2 and 0.15 respectively, and normalized by its maximum value. (A) Analytic signal $x^+(t)$. (B) One-sided spectrum $\Theta(f)\hat{x}(t)$. (C) Real signal $x(t)$. (D) Hermitian symmetric spectrum $\hat{x}(f) = \overline{\hat{x}(-f)}$. The complex values are represented in (A,B,D) as follows: the modulus is plotted as a function of time and the angle (complex argument θ in $e^{i\theta}$) is colour-coded with the hues of the chromatic circle (periodic colour bar). Arrows indicate the mapping relating the different representations and form a commutative diagram.*

physical insight into time-varying or transient oscillations. Their correspondence with the true amplitude and phase that has generated the signal $x(t) = A_0(t)\cos(\phi_0(t))$ is only correct under specific assumptions that have been long studied [Bedrosian 1963; Picinbono 1997], namely the separation of scales of the phase and amplitude modulations, the latter being much slower. For this reason, the analytic version of the signal is an important tool for the study of a harmonic wave carrying slow modulations, also called asymptotic signals.

One such oscillatory signal is presented in Fig.I.1, both in the time and frequency domains, and in its real and analytic versions, with a colour-coding of the angles. This signal has been generated from a complex exponent, in particular from their time-derivative (denoted with a dot), as:

$$x^+(t) = e^{m(t)+i\phi(t)} \quad ; \quad r_x(t) = \dot{m}(t) \quad ; \quad f_x(t) = \frac{\dot{\phi}(t)}{\Upsilon} \quad , \tag{I.4}$$

where we call $m(t) = \log|x^+(t)|$ the magnitude (a logarithm, as the one of astronomy), $r_x(t)$ the instantaneous rate of magnitude change (whose modulus is also interpreted as an instantaneous bandwidth [Cohen and Lee 1990]) and $f_x(t)$ the instantaneous frequency of the signal. In Fig.I.1, we have chosen a modulation $f_x(t)$ around the mean frequency f_0 and the same modulation for the rate $r_x(t)$:

$$r_x(t) = f_x(t) - f_0 = r_1\sin(\Upsilon f_1 t) + r_2\cos(\Upsilon f_2 t) \quad \Rightarrow \quad x^+(t) = e^{\int_0^t r_x(t')+i\Upsilon f_x(t')\mathrm{d}t'} \quad . \tag{I.5}$$

The usefulness of the analytic signal finds its limit when its contains distinct oscillatory activities at multiple scales, requiring their separation with time-frequency methods [Delprat et al. 1992; Le and Argoul 2004; Lilly and Olhede 2010].

1.4 Note on causality: swapping frequency and delay

The Kramers-Kronig relations in physics are often applied to real and imaginary parts of a spectrum rather than a signal. For instance in linear response theory, the fluctuation-dissipation theorem allows to characterize a dynamical system close to equilibrium from its thermal fluctuations [Kubo 1966]. This is achieved assuming that the response function (or susceptibility) of the physical system $\chi(\tau)$ is a function of positive delays only, i.e. it is *causal*: $\chi(\tau) = \Theta(\tau)\chi(\tau)$. Since the spectrum of this susceptibility $\hat{\chi}(f)$ is analytic, we wish here to emphasize the symmetry with the previously exposed situation, which accounts to switching the labels "spectrum" and "frequency" with the ones "signal" and "time" in Fig.I.1.

Note that we cannot define an analytic signal whose spectrum is also analytic (and non-zero). However, a causal and real signal (or a one-sided and real spectrum) can be extended into an analytical one without loss of information; only the imaginary part is non-causal.

2 Introducing the wavelet transform

Time-frequency analyses were first based on time and frequency-translated windowing of the Fourier transform; this is the Gabor or short-time Fourier transform, defined in the Appendix 1 [Gabor 1946]. A method based on time translation and scaling (frequency dilation), namely the *wavelet transform*, was then introduced in the end of the twentieth century [Carmona, Hwang, and Torrésani 1997; Carmona, Hwang, and Torrésani 1995; Combes, Grossmann, and Tchamitchian 1989; Delprat et al. 1992; Grossmann and Morlet 1984; Kronland-Martinet, Morlet, and Grossmann 1987]. It has since been applied to many scientific domains for analysing and modelling non-stationary and multi-scale objects.

Both methods constitute two paradigms for transforming complex recordings in a two-dimensional representation, that aims at facilitating the reading and the manipulation of its rich content. They share many common characteristics, and have been classified in a broad theory of the time-frequency representations [Carmona, Hwang, and Torrésani 1998; Flandrin 1998b; Papandreou-Suppappola, Hlawatsch, and Boudreaux-Bartels 1998]. The former is often presented as closer to the spectral intuition inherited from the Fourier analysis, and the latter is the most suited to a multi-scale analysis, hence called a time-scale representation. We agree with the second argument, justifying our choice of the wavelet paradigm to study natural recordings, and we show that a certain use of the wavelet transform inherits the oscillatory intuition associated to analytic signals.

2.1 Definition of the wavelet transform

The wavelet transform of a finite energy signal $x(t) \in L^2(\mathbb{R})$ is defined as its inner product with the shifted and scaled copies of an analysing absolute integrable and finite energy wavelet : $\psi(t) \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ [Chui 1992; Torrésani 1995]:

$$\mathcal{W}_\psi[x](a,b;p) = \int x(t)\overline{\psi\left(\frac{t-b}{a}\right)}|a|^{-\frac{1}{p}}\mathrm{d}t \quad , \tag{I.6}$$

b and a are real shift and scale variables. It quantifies the matching of local features of the signal with the shape of wavelet, used as a macro-to-microscope scanning the signal along time and

across timescales [Arneodo, Grasseau, and Holschneider 1988]. Often hidden, p is a parameter that can be interpreted as the exponent of the norm which is conserved across scales for the dilated wavelet $|a|^{-\frac{1}{p}}\psi(t/a)$: the conventions $p = 1$ and $p = 2$ are usually found in the literature and this choice varies with the application and the preference of the authors. As will be made clear in the course of this chapter, p should not matter in the theory of wavelets, but $p = 1$ suits best the ideas of scaling and of analytic signal's amplitude, while $p = 2$ is helpful to match the intuition of Fourier's frequency domain.

In the following, $p = 1$ is preferred and omitted in $\mathcal{W}_\psi[x](a, b)$, so that the occasional use of the convention $p = 2$ will be denoted $\sqrt{a}\mathcal{W}_\psi[x](a, b)$. While it is quite straightforward to use the shift parameter b as a time variable, the interpretation of $\mathcal{W}_\psi[x](a, b)$ as a time-frequency representation of x relies on the key properties of time and frequency localization of the analysing wavelet (see section 3.2).

The computation of the wavelet transform from the frequency domain highlights the interpretation of the wavelet as a signal processing filter:

$$\mathcal{W}_\psi[x](a, b) = \int \hat{x}(f)\overline{\hat{\psi}(af)}e^{i\Upsilon f b}\mathrm{d}f \quad . \tag{I.7}$$

Of very practical interest, this expression takes the form of an inverse Fourier transform and operates a band-pass filtering of the signal, whose typical frequency f_a scales inversely with the scale variable a. Therefore, the correct relation between f_a and a gives a correspondence from wavelet scale to frequency, and the collection of wavelet filters indexed by a determines a new frequency sampling.

Contrary to the time-shift variable b, the physical dimension of the scale a is left free of interpretation in the definition of the wavelet transform: it can either be a dimensionless factor if the wavelet ψ has the same physical nature as the signal x, or it can be used as an explicit and adjustable time unit, turning the wavelet into a function of a dimensionless variable. Although of physical importance, this aspect of the wavelet is often implicit, with a physical unit set to 1 when appropriate. In the following, the use of parametric functions for the wavelet rapidly leads us to adopt the second interpretation. Together with the convention $p = 1$, it also means that the wavelet coefficients $\mathcal{W}_\psi[x](a, b)$ have the same physical dimension as the signal x.

As a localized version of the pure complex wave in the Fourier transform, a complex helical shape is a great choice of wavelet to complete the spectral intuition associated to the scale parameter. We support this idea with a general argument and a useful property favouring the choice of an analytic wavelet, before discussing its localization in time and frequency.

2.2 From analytic continuation to analytic wavelet transform: a natural derivation

For the same reasons as discussed previously with the analytic signals, using a real wavelet to analyse real signals accounts to computing only the real part of a corresponding analytic wavelet transform. We show below how a basic application of complex analysis to the separation of scales in a signal naturally leads to the analytic wavelet transform.

The analytic signal $x^+(t)$ is called so because its expression Eq.(I.2) can extend to an analytic function of the upper-half complex plane $\mathbb{H} = \{z = t + i\tau \mid \tau > 0\,;\, t, \tau \in \mathbb{R}\}$:

$$x^+(z) = \int_0^\infty \hat{x}(f)e^{i\Upsilon f' z}\mathrm{d}f' = \int_{-\infty}^{+\infty} x(t')\frac{1}{i\Upsilon(t' - z)}\mathrm{d}t' \quad . \tag{I.8}$$

Form the time domain, it takes the form of a Cauchy integral with a contour following the real axis, whose meaning for $z = t$ (real) is made precise using the principal value (see also the Sokhotski-Plemelj theorem [Plemelj 1964]). As a remark, the study of causal filters is in

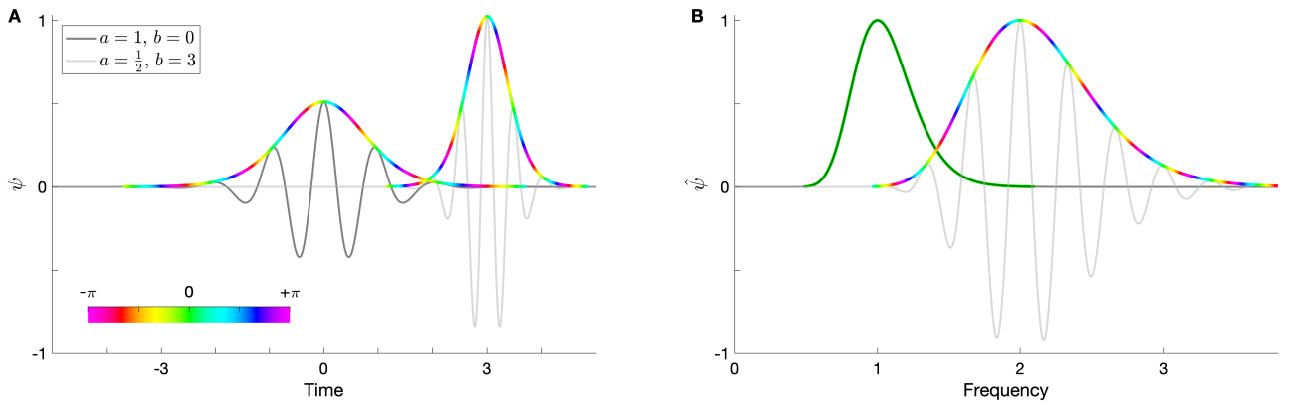


Figure I.2: *Two copies of the same analytic wavelet, one centred, the other dilated and shifted. They are plotted (A)* $\psi\left(\frac{t-b}{a}\right)\frac{1}{|a|}$ *in the time domain, and (B)* $\hat{\psi}(af)e^{-i\Upsilon f b}$ *in the frequency domain. In both panels, the thick lines represent their modulus with colour-coded angle, while the thin grey lines represent their real part only, in darker grey for the centred wavelet (*$a = 1$, $b = 0$*) and lighter for the dilated and shifted wavelet (*$a = \frac{1}{2}$, $b = 3$*). The wavelet is the most used one in this thesis,* ψ_Q *with* $Q = 5$*, as defined later in Eq.*(I.33).

the exact symmetric situation: instead of the signal, the spectrum (of the filter) is analytically continued: $x(t), t$ and $\hat{x}(f), f$ are swapped in Eq.(I.8), see the note 1.4.

The imaginary coordinate $\tau = \Im\{z\}$ introduced in this analytical continuation of the signal has everything of a scale parameter, while the real part $t = \Re\{z\}$ is a shift parameter in a proto-wavelet ψ_0, put into evidence with the change to dimensionless variables $u = \frac{t'-t}{\Upsilon\tau}$ and $v = \Upsilon f'\tau$:

$$\psi_0(u) = \frac{1}{1 - i\Upsilon u} \qquad ; \qquad \hat{\psi}_0(v) = \Theta(v)e^{-v}$$
$$x^+(t + i\tau) = \mathcal{W}_{\psi_0}[x](\Upsilon\tau, t) \quad . \tag{I.9}$$

From this viewpoint, the identified scale variable $a = \Upsilon\tau$ has the same physical dimension as the shift variable $b = t$, and ψ_0 is a pure mathematical shape which makes coincide analytic continuation and wavelet transform of the signal. In the frequency domain, this particular wavelet sets an exponential cut-off at the frequency $f = (\Upsilon\tau)^{-1}$, so that $x^+(t+i\tau)$ is a smoothed version of the analytic signal, cumulating the influence of lower frequencies. The precise decaying shape of this low-pass filter has no influence on the high resolution limit $x^+(t)$.

In contrast to this cumulated representation of scales, the idea of the wavelet transform is to separate each scale's contributions to the signal. A rudimentary way to do it is to derive with respect to the cut-off scale τ, equivalent to an imaginary time derivative from the Cauchy-Riemann property of analytic functions:

$$\partial_\tau x^+(t + i\tau) = i\partial_t x^+(t + i\tau) \quad . \tag{I.10}$$

Rewritten as a derivative with respect to the cut-off frequency f, the proto-wavelet in Eq.(I.9) turns into a Cauchy wavelet:

$$f\partial_f x^+(t + i/\Upsilon f) = \frac{1}{i\Upsilon f}\partial_t x^+(t + i/\Upsilon f) = \mathcal{W}_{\psi_1}[x](f^{-1}, t) \tag{I.11}$$

$$\psi_n(u) = \frac{1}{(1 - i\Upsilon u)^{n+1}} \qquad ; \qquad \hat{\psi}_n(v) = \Theta(v)v^n e^{-v} \quad , \tag{I.12}$$

where the time unit brought by the frequency derivative is compensated by the frequency factor, that keeps the wavelet dimensionless and the wavelet coefficient with the dimension of

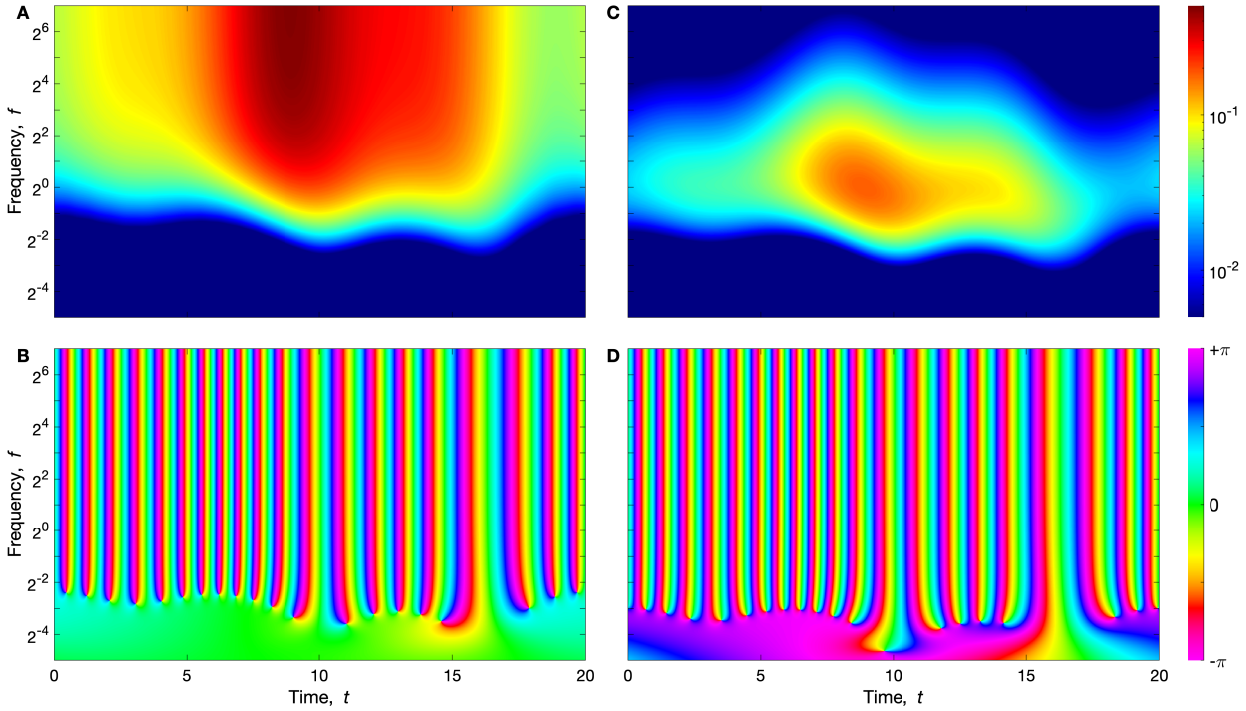


Figure I.3: *Comparison of analytic continuation and analytic wavelet transform for the signal x of Fig.I.1. (A,B) Time-frequency representation of the analytic continuation $x^+(t+i\tau)$, Eq.*(I.9)*, whose modulus (A) and angle (B) are colour-coded for each cut-off frequency $f = (\Upsilon\tau)^{-1}$ (on a log-scale), and time t. (C,D) Same representation for the analytic wavelet transform $\mathcal{W}_{\psi_1}[x](f^{-1},t)$, with the first Cauchy wavelet ψ_1 defined in Eq.*(I.11)*.*

x. Notice how the logarithm of the cut-off frequency appears in $f\partial_f = \partial_{\log f}$, consistent with the multiplicative behaviour of the scale variable. The differences between these representations of a signal x are illustrated in Fig.I.3. A wavelet transform that relates to an analytic function of the complex plane is necessarily based on a Cauchy wavelet, slightly generalized in [Holighaus et al. 2019].

This yields the simplest example of analytic wavelet (and its transform), ψ_1, member of the Cauchy family Eq.(I.12), of which the proto-wavelet can be considered the order zero. Notice here how the analyticity of the signal, enforced by the Heaviside step function in Eqs.(I.2, I.8), is absorbed into the construction of an analytic wavelet in Eq.(I.12).

By construction of this wavelet transform as a derivative with respect to scale in Eq.(I.11), the analytic signal is recovered from the cumulation of the scale contributions:

$$\int_0^\infty \mathcal{W}_{\psi_1}[x](a,t)\mathrm{d}\log a = -\Big(x^+(t+i\infty) - x^+(t+i0)\Big) = x^+(t) \quad , \tag{I.13}$$

with $a = f^{-1} = \Upsilon\tau$.

2.3 Signal reconstruction and synthesis wavelets

The real part Eq.(I.3) of the expression Eq.(I.13) is a remarkably simple inversion formula for the wavelet transform of a real signal, which applies quite generally to any analytic wavelets ψ (of positive frequency support), providing the existence of a finite correction factor C_ψ:

$$x^+(t) = C_\psi^{-1}\int_0^\infty \mathcal{W}_\psi[x](a,t)\mathrm{d}\log a \qquad ; \qquad C_\psi = \int_0^\infty \overline{\hat{\psi}(v)}\mathrm{d}\log v \quad . \tag{I.14}$$

In practice, the signal is thus reconstructed via its analytic version by summing the wavelet transform coefficients over geometrically sampled positive scales (evenly spaced on a log-scale).

This reconstruction formula is amply sufficient for the needs of this thesis. Nevertheless, we mention here for the completeness of the presentation a more flexible and general reconstruction formula since it involves the choice of a second synthesis wavelet φ [Daubechies 1992; Holschneider and Tchamitchian 1990]:

$$x(t) = C_{\varphi\psi}^{-1} \iint \mathcal{W}_\psi[x](a,b)\varphi\left(\frac{t-b}{a}\right)\frac{1}{|a|}\frac{\mathrm{d}a\mathrm{d}b}{|a|} \quad ; \quad C_{\varphi\psi} = \int \hat{\varphi}(v)\overline{\hat{\psi}(v)}\frac{\mathrm{d}v}{|v|} \quad , \qquad \text{(I.15)}$$

for any pair of analysis and synthesis wavelets, ψ and φ (not necessarily analytic), satisfying the admissibility condition of a finite non-zero factor $C_{\varphi\psi}$. A common choice is $\varphi = \psi$, for which $C_{\psi\psi}$ is referred to as the admissibility coefficient of the wavelet. The complex exponential wave is also common, and Eq.(I.14) corresponds to yet another possibility: $\varphi = \delta$ the Dirac delta.

In Eq.(I.15), the integral over positive and negative scales, $\int \frac{\mathrm{d}a}{|a|}$ does not only apply to real (or analytic) signal: the analytic wavelet transform can be used for a general complex signal. In this case, however, we cannot overlook the computation of the wavelet transform at negative scales (i.e. negative frequencies). A change of the convention $p = 1$ only changes the exponent on the scale integrator that becomes $\int \frac{\mathrm{d}a}{|a|^{2-\frac{1}{p}}}$.

These reconstruction formulas are in practice restricted to a subspace of the values (a, b), leading to a partial reconstruction of the signal in the time-scale domain of interest. Finally, applying another wavelet transform to these reconstruction formulas provides us with ways of changing of wavelet (continuous basis or frame). This more general formula is the one proved in Appendix 3; in particular, it leads to the reproducing kernel formula when the new wavelet is identical to the former (allowing a Hilbert space approach of the wavelet transform), and the reconstruction formulas are recovered.

Summary of sections 1 and 2

The oscillatory interpretation of a real signal $x(t)$ benefits from its analytic versions $x^+(t)$. Without loss of information, the phase and amplitude of these oscillations are straightforwardly obtained from the angle (complex argument) and radius (modulus) of the complex helical signal $x^+(t)$.
The complex analysis associated to the concept of analytic signal introduces a timescale variable. The distinction of the contributions to the signal at each scale naturally leads to a wavelet transform defined with an analytic wavelet. This natural derivation identifies the translation variable b with time, while the scale variable a is a timescale, conferring to the analytic wavelet the role of a mathematical function with a dimensionless variable. The use of an analytic wavelet brings several advantages and simplification to the wavelet analysis.
There is a complete correspondence between positive and negative values of the scale variable and positive and negative frequency components of the signal, which are well separated. This simplifies the computation of the wavelet transform and the reconstruction of the signal when it is real. The natural scale decomposition of the signal from the analytic wavelet confers to the transform the same oscillatory interpretation of the angle and modulus as phase and amplitude. As a complex helical function, the wavelet gets closer to the pure complex wave of the Fourier transform.

3 Wavelet shape: time-frequency localization and uncertainty atom

The time-frequency interpretation of the wavelet transform relies on the localization properties of the chosen wavelet. For a general family of analytic wavelets, we present in this section how these properties relate to their shape. We select a one-parameter wavelet that has ideal properties in many aspects, and will be used in the following parts.

To focus on their mathematical shapes, we pursue the approach introduced in the section 2.2 that consists in using the variables:

$$u = \frac{t-b}{a} \quad ; \quad v = af \quad , \tag{I.16}$$

dimensionless time and frequency respectively, by considering a as a reference timescale (or unit).

3.1 Introducing the Morse family of analytic wavelets

Most (if not all) analytic wavelets defined as a real band-pass filter can be expressed in terms of (or approximated by) a very general parametric family called the Morse wavelets [Daubechies 1988; Lilly and Olhede 2012]. Their Fourier spectra is the following real function of a positive dimensionless frequency $v > 0$:

$$\hat{\psi}^{(k)}_{\beta,\gamma}(v) = v^\beta e^{-v^\gamma} L_k^{(\alpha-1)}(2v^\gamma) \ , \quad \alpha = \frac{2\beta+1}{\gamma} \ , \quad L_k^{(\alpha')}(x) = \sum_{\ell=0}^{k} \binom{k+\alpha'}{k-\ell} \frac{x^\ell}{\ell!} \quad . \tag{I.17}$$

The parameter k fixes a polynomial order in the (generalized) Laguerre orthogonal family $L_k^{(\alpha')}$, where the binomial coefficient extends to non-integer parameters via the Gamma function Γ.

The exponents β, γ, usually positive, control the shape of this band-pass filter: while β sets a polynomial filtering at low frequencies (high-pass), γ sets a stretched exponential filtering at high frequencies (low-pass); in-between stands a characteristic frequency. Decays of different nature at high and low frequencies can be switched by choosing negatives exponents, for which the wavelet has an inversion property:

$$\hat{\psi}^{(k)}_{-\beta,-\gamma}(v) = \hat{\psi}^{(k)}_{\beta-1,\gamma}(v^{-1})v^{-1} \quad . \tag{I.18}$$

In the case of exponents of different signs, $\beta\gamma < 0$, the Morse wavelet diverges and cannot be considered as a band-pass filter any more.

The case $\gamma = 1$ of this shape was introduced in the context of quantum mechanics as the eigenstates in the Morse potential [Morse 1929]; the order $k > 0$ generalizes the Cauchy wavelets: $\psi_n = \psi^{(0)}_{n,1}$. It has been extended to any γ for the partial reconstruction of a signal (with $\phi = \psi$), restricted to a time-scale disc region of the upper-half plane $\mathbb{H}$ [Daubechies 1988]. The corresponding (squared) eigenvalues with respect to this partial reconstruction operator was shown to quantify the time-frequency (energy) concentration of the Morse wavelets [Bayram and Baraniuk 1996; Olhede and Walden 2002]. Their property of orthogonality in both time and frequency domains writes:

$$\int_0^\infty \hat{\psi}^{(m)}_{\beta,\gamma}(v)\hat{\psi}^{(n)}_{\beta,\gamma}(v)\mathrm{d}v = \frac{\Gamma(n+\alpha)}{n!2^\alpha\gamma}\delta_{m,n} = \int \psi^{(m)}_{\beta,\gamma}(u)\overline{\psi^{(n)}_{\beta,\gamma}(u)}\mathrm{d}u \quad , \tag{I.19}$$

with $\delta_{m,n}$ the Kronecker delta, preceded by the energy (norm 2) of the wavelet.

The study of this shape (characteristic frequency, bandwidth, symmetry) can be restricted to the case $k = 0$ [Lilly and Olhede 2009; Lilly and Olhede 2012], positive and non-oscillating in the frequency domain, simply because $\psi^{(k)}_{\beta,\gamma}$ is a linear combination of $\psi^{(0)}_{\beta+\ell\gamma,\gamma}$, $\ell \leq k$. Furthermore, the time-frequency concentration decreases with the order k [Olhede and Walden 2002], making $k = 0$ a first choice in many practical situations. For this reason, the Morse wavelet term will simply refer to $\psi_{\beta,\gamma} = \psi^{(0)}_{\beta,\gamma}$ in the following, while the notation $\psi^{(k)}_{\beta,\gamma}$ will be reserved to particular applications where the full orthogonal family is needed.

In general, the Morse wavelet lacks of a simple expression in the time domain (except for simple integer values of γ). Nevertheless, its time derivative to an arbitrary (possibly fractional) order n has the interesting property of being another Morse wavelet:

$$\partial^n_u \psi_{\beta,\gamma}(u) = (i\Upsilon)^n \psi_{\beta+n,\gamma}(u) \quad . \tag{I.20}$$

3.2 Time-frequency localization

The wavelet transform of a signal can be given a time-frequency interpretation when the wavelet is well localized both in the time and frequency domains. Therefore, it is essential to understand and carefully select the shape of the wavelet, determining the time-frequency localization.

It is common, when working both in time and frequency, to approach the localizing kernel (here the wavelet, but also its counterpart the Gabor window) as a "wave function", whose time-domain and frequency-domain moduli raised to a certain power, $|\psi(u)|^p$, $|\hat{\psi}(v)|^q$, are considered as (unnormalized) density functions. For instance, the spectrum of the Morse wavelet corresponds to the generalized gamma distribution in a probabilistic context, for which all moments can be computed. Treating similarly time and frequency, we may define a time-frequency position (u_p, v_q) and width $(\Delta u_p, \Delta v_q)$ from the mean and standard deviation of these densities. Along with alternative definitions, their expressions can be found later in Table I.1. They provide the essential information regarding time-frequency localization and uncertainty. Higher order moments or cumulants contain more precise information about the shape, that can be considered secondary.

3.2.1 Morse position: easy in time, not so much in frequency

For Morse wavelets, the exponent p does not influence the time position, since a real wavelet spectrum $\hat{\psi} \in \mathbb{R}$ results in a Hermitian time symmetry $\overline{\psi(u)} = \psi(-u)$, implying:

$$u_p = 0 \quad . \tag{I.21}$$

However, the central frequencies v_q of analytic wavelets, see Eq.(I.23), have both different values and interpretations. For instance, v_∞ selects the maximal value of $\hat{\psi}$ so that $f_\infty = v_\infty/a$ is called the peak frequency of the wavelet at scale a. Similarly, when the wavelet spectrum is real and non-negative, such as $\psi_{\beta,\gamma}$, the centre instantaneous frequency (in the sense of Eq.(I.4)) coincides with the norm 1:

$$\Im\left\{\frac{\partial_u \psi(0)}{\Upsilon\psi(0)}\right\} = \Im\left\{\frac{\int_0^\infty i\Upsilon v\hat{\psi}(v)\mathrm{d}v}{\Upsilon\int_0^\infty \hat{\psi}(v)\mathrm{d}v}\right\} = \frac{\int_0^\infty v|\hat{\psi}(v)|\mathrm{d}v}{\int_0^\infty |\hat{\psi}(v)|\mathrm{d}v} = v_1 \quad , \tag{I.22}$$

where the first equality is the definition of the instantaneous frequency at time 0, expressed from the spectrum, and its non-negativity is used in the second one.

In addition, alternative approaches are available for defining a characteristic scale of an analytic wavelet. While the first moment corresponds to an (integral) arithmetic average of the frequency, the harmonic average (first moment of negative order) has also proved useful in

geophysics [Claerbout 1985] (p. 68 for $q = 2$) since it averages a characteristic time (inverse frequency). Let us denote it v'_q; it takes the specific form of a ratio between the wavelet norm and admissibility coefficient, here compared to v_p:

$$v_q = \frac{\int_0^\infty v|\hat{\psi}(v)|^q \mathrm{d}v}{\int_0^\infty |\hat{\psi}(v)|^q \mathrm{d}v} \quad ; \quad v'_q = \frac{\int_0^\infty |\hat{\psi}(v)|^q \mathrm{d}v}{\int_0^\infty |\hat{\psi}(v)|^q \mathrm{d}\log v} \quad . \tag{I.23}$$

Mitigating the choice between an average time or frequency scale, the geometric average (from the first logarithmic moment) gives an intermediary value. The relevance of logarithmic scales to the wavelet could be taken into account even more radically by interpreting the wavelet spectrum as a density with respects to $\mathrm{d}\log v$ as in Eqs.(I.14, I.15), yielding yet another class of expressions.

The following observation saves us from these numerous possible choices: from its maximizing property, the limit $q \to +\infty$ always agrees on the peak frequency for all these alternatives, defining our preference for the wavelet characteristic scale. For Morse wavelets, it relates to the parameters β and γ as follows:

$$v_\infty = v'_\infty = \left(\frac{\beta}{\gamma}\right)^{\frac{1}{\gamma}} \quad . \tag{I.24}$$

reaching the maximum value $\hat{\psi}_{\beta,\gamma}(v_\infty) = \left(\frac{\beta}{e\gamma}\right)^{\frac{\beta}{\gamma}}$ of this parametric band-pass filter. This approach is consistent with the definition of the characteristic time since it can be proved that the modulus of Morse wavelets is maximized at time zero.

For an easier comparison of the Morse wavelet shapes, it is convenient to standardize it to the same peak frequency and corresponding maximum value [Lilly and Olhede 2009]:

$$\hat{\varphi}_{\beta,\gamma}(v) = \frac{\hat{\psi}_{\beta,\gamma}(v_\infty v)}{\hat{\psi}_{\beta,\gamma}(v_\infty)} = \left(v\, e^{\frac{1-v^\gamma}{\gamma}}\right)^\beta \quad , \tag{I.25}$$

where the dilation that centres the wavelet at the new peak frequency $\hat{\varphi}_{\beta,\gamma}(1) = 1$ can be considered as a change of the scale variable. By extension, we call *standardized* a wavelet that has a unit dimensionless peak frequency: $v_\infty = 1$. Causing no loss of generality, this requirement is helpful to interpret the scale variable directly as an inverse frequency $f_\infty = 1/a$, and to select the Morse parameters β, γ (see Fig.I.4).

3.2.2 Widths for Morse wavelets? Introducing the quality factor

At first sight, the question of defining the widths of the Morse wavelet, which determines the time and frequency resolutions of the corresponding transform, seems cluttered by the same difficulties as the one of the characteristic frequency, namely a profusion of possibilities depending on the norm we consider, the linear, inverse or logarithmic frequency scale, etc. The lack of simple wavelet expressions in the time domain appears to restrain us to the energy approach $p = 2 = q$ of the second cumulants, which remains computationally tractable for general parameters β, γ. The resulting widths are cumbersome expressions involving the Gamma function.

Pursuing the idea of characterizing the Morse wavelet from its properties at its peak, the use of its curvature has been proposed in [Lilly and Olhede 2009], expressed for a wavelet $\hat{\psi}$ from its peak frequency v_∞:

$$Q = \sqrt{-\frac{v_\infty^2 \partial_v^2 \hat{\psi}(v_\infty)}{\hat{\psi}(v_\infty)}} \quad . \tag{I.26}$$

Also known to specify a width in a different parametrization of the general Gamma distribution, this quantity simplifies greatly the interpretation of the Morse parameters:

$$Q = \sqrt{-\partial_v^2 \hat{\varphi}_{\beta,\gamma}(1)} = \sqrt{\beta\gamma} \quad , \tag{I.27}$$

advantageously expressed from the standardized Morse wavelet Eq.(I.25). We interpret this positive number as a *quality factor*, usually defined as the inverse of a relative bandwidth, $\frac{v_2}{2\Delta v_2}$, that represents the prominence of the peak. Conversely, the wavelet $\hat{\varphi}_{\beta,\gamma}(af)$ of peak frequency a^{-1} can be given a bandwidth $(aQ)^{-1}$. The particular definition Eq.(I.26) has several advantages. It is unchanged by the logarithmic approach of scales and frequencies: $\partial^2_{\log v}\hat{\varphi}_{\beta,\gamma}(1) = \partial_v\hat{\varphi}_{\beta,\gamma}(1) + \partial_v^2\hat{\varphi}_{\beta,\gamma}(1) = \partial_v^2\hat{\varphi}_{\beta,\gamma}(1)$, since $\partial_v\hat{\varphi}_{\beta,\gamma}(1) = 0$ from the unit peak frequency of standardized wavelets, so that this quality factor also expresses to log-frequency bandwidth:

$$\Delta \log v = \frac{1}{Q} \quad . \tag{I.28}$$

In addition, Eq.(I.26) provides a time-domain interpretation from the following "variance" Δu^2 of the wavelet demodulated by its peak frequency [Lilly and Olhede 2009]:

$$(\Upsilon v_\infty \Delta u)^2 = \frac{\int (\Upsilon v_\infty u)^2 \psi(u) e^{-i\Upsilon u v_\infty} \mathrm{d}u}{\int \psi(u) e^{-i\Upsilon u v_\infty} \mathrm{d}u} = Q^2 \quad . \tag{I.29}$$

The use of the demodulated wavelet $\psi(u)e^{-i\Upsilon u v_\infty}$ is a complex approximation of the positive time-domain envelop $|\psi(u)|$ ($p = 1$), and is expected to slightly underestimate Δu_1 without requiring an explicit temporal expression for its computation; it is developed further in [Lilly and Olhede 2010]. As a result, the duration of the dilated wavelet $\varphi_{\beta,\gamma}(\frac{t-b}{a})$ is proportional to the quality factor and scales with a: $\Delta t = \frac{aQ}{\Upsilon}$.

3.2.3 Uncertainty and wavelet widths

Although the previous approach does not yield distinct expressions for the time and frequency widths, it manages to provide a simple control on the localization trade-off between time and frequency

$$\Delta u \Delta v = \Delta u \Delta \log v = \frac{1}{\Upsilon} \quad , \tag{I.30}$$

which is compatible with the Heisenberg uncertainty relation:

$$\Delta u_2 \Delta v_2 \geq \frac{1}{2\Upsilon} \quad . \tag{I.31}$$

We thus have a simple expression for the "uncertainty" on the Morse wavelet position. Although its simplicity is charming, we should keep in mind that the uncertainties defined above are about local properties around the peak, relying on the regularity of the Morse wavelet spectrum (single smooth bump). Therefore, it might be very different from the standard deviation used in the common uncertainty relation, especially in the case of very asymmetric profiles.

A more precise and general approach of time and frequency uncertainty is based on the full wavelet shapes $|\psi(u)|^p, |\hat{\psi}(v)|^q$. For conjugated exponents $\frac{1}{p} + \frac{1}{q} = 1$, the framework of L^p spaces provides an entropic uncertainty relation [Beckner 1975; Hirschman 1957] (see Appendix 4). Giving a lower bound to the concentration in time from the known concentration in frequency, and vice versa, it is expressed in terms of Rényi entropies:

$$e^{H_{\frac{p}{2}}[|\psi|^2] + H_{\frac{q}{2}}[|\hat{\psi}|^2]} \geq p^{\frac{1}{p-2}} q^{\frac{1}{q-2}} \quad ; \quad H_\alpha[\rho] = \frac{1}{1-\alpha} \log \frac{\int \rho(u)^\alpha \mathrm{d}u}{(\int \rho(u)\mathrm{d}u)^\alpha} \quad . \tag{I.32}$$

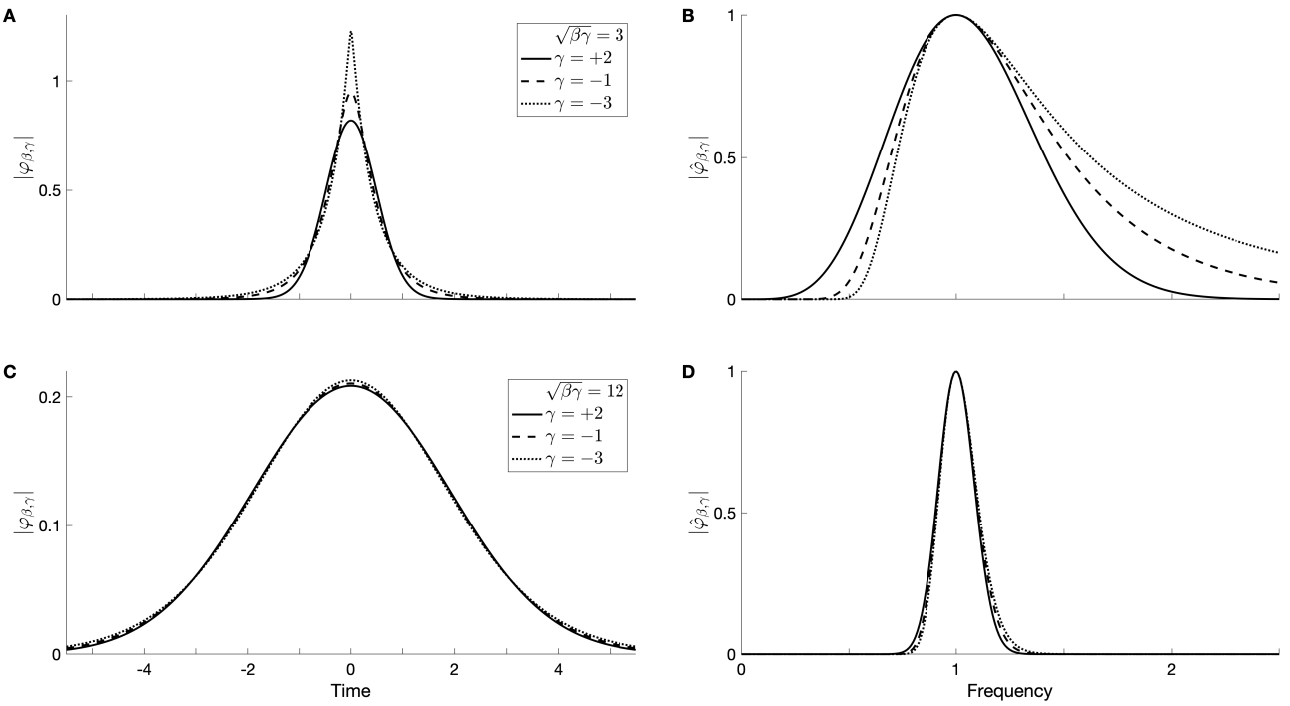


Figure I.4: *Comparison of standardized Morse wavelets: (A, C) $|\varphi_{\beta,\gamma}|$ in time, and (B, D) $|\hat{\varphi}_{\beta,\gamma}|$ in frequency, with different parameters β, γ but the same quality factors: $Q = \sqrt{\beta\gamma} = 3$ in (A, B), and $Q = 12$ in (C, D). The modulus is represented in black continuous lines for $\gamma = +2$, dashed lines for $\gamma = -1$ and dotted lines for $\gamma = -3$.*

The central case of the energy ($p = 2 = q$) implies the Heisenberg inequality Eq.(I.31) through its relation to the standard deviation: $e^{H_{\frac{p}{2}}[|\psi|^2]} \leq \sqrt{\uparrow e}\,\Delta u_2$ (idem for Δv_2). Therefore, exponential entropies are also good candidates to measure the extent of the wavelet dispersion, whatever its shape.

3.2.4 Time-frequency interpretation of the wavelet transform

Once the Morse wavelet is standardized, Eq.(I.25), and the quality factor $Q = \sqrt{\beta\gamma}$ is fixed (requiring $\beta\gamma > 0$), we have a good control on its time-frequency localization, i.e. on the time-frequency interpretation of the associated wavelet transform.

For instance, the exponential wave $e^{i\uparrow f_0 t}$ is the (analytic) signal that is fully localized at the frequency f_0 and delocalized in time t. Its wavelet transform is given by $\mathcal{W}_\psi[e^{i\uparrow f_0 t}](a,b) = \overline{\hat{\psi}(af_0)}e^{i\uparrow f_0 b}$, whose modulus is maximal at $a = f_0^{-1}$ for a standardized wavelet. The frequency width of $\hat{\psi}$ provides a natural frequency resolution Eq.(I.29). The phase of the wave at time t is also recovered from the angle at $b = t$.

Conversely, a strong impulse signal localized at time t_0 (delocalized in frequency) can be modelled by the Dirac function $\delta(t - t_0)$, whose wavelet transform is $\mathcal{W}_\psi[\delta_{t_0}](a,b) = \overline{\psi\left(\frac{t_0 - b}{a}\right)}\frac{1}{|a|}$. The time of the pulse is easily recovered for a symmetric wavelet, of maximal modulus at the time $b = t_0$, with the duration of the wavelet as a natural time resolution.

Therefore, for an analytic wavelet ψ that is symmetric, Eq.(I.21) (i.e. of real spectrum), and standardized, Eq.(I.25), the wavelet transform $\mathcal{W}_\psi[x](f^{-1}, t)$ can be used directly as a time-frequency representation of x. Its resolution is determined by the choice of a quality factor Q, fixing a constant log-frequency resolution $\Delta \log f = Q^{-1}$ and a time resolution that adapts to the frequency, $f\Delta t = \frac{Q}{\uparrow}$, i.e. a constant number of oscillations whatever their period. Low quality factors are most suited to distinguish time-localized singular (or few oscillating) events

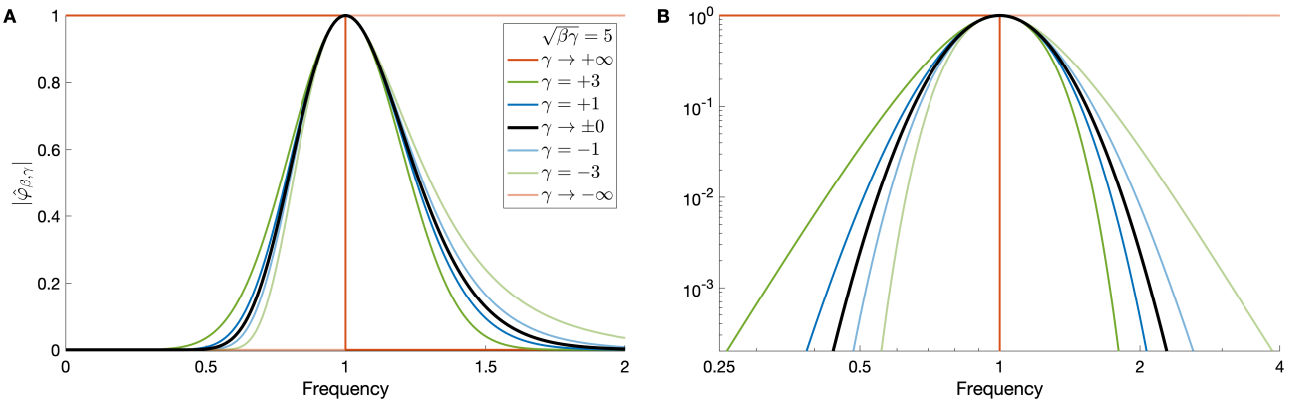


Figure I.5: *Comparison of the asymmetry in the frequency domain of standardized Morse wavelets $\hat{\varphi}_{\beta,\gamma}$ for different parameters but the same fixed quality factors $Q = \sqrt{\beta\gamma} = 5$: (A) in linear axes, and (B) in logarithmic axes. Morse wavelets of parameter $\gamma = \pm 1, \pm 3, \pm\infty$ are respectively in blue, green and red lines, lighter when γ is negative, and Grossmann's log-normal wavelet corresponds to $\gamma \to 0$, in black.*

in the analysed signal, whereas high quality factors can discriminate simultaneous oscillating components of well localized frequencies.

In the examples above, the perfect frequency localization of the pure wave can only be fully resolved from an infinite quality factor, whereas the time localization of the pulse is fully resolved for a vanishing quality factor. The resolution in the other dimension is lost for both limit cases, so that we only obtain a time or a frequency representation: the analytic signal or the Fourier spectrum in certain conditions (see Appendix 6). The intermediary situation of a finite quality factor is the one of interest in multi-scale and time-dependent situations, bridging the gap between the time and frequency representations. For Morse wavelets parametrized by the pair (β, γ), we are still left with a second degree of freedom once $Q = \sqrt{\beta\gamma}$ is fixed. Less relevant than the quality factor for a time-frequency interpretation, we show in the following how to discard it.

3.3 Grossmann's log-normal wavelet, like the nose on Morse's face

Here we introduce the analytic wavelet that is used throughout all practical applications in this thesis. It has a single parameter Q, that determines the quality factor of the wavelet, whose spectrum is log-normal:

$$\hat{\psi}_Q(v) = e^{-\frac{1}{2}(Q \log v)^2} \quad . \tag{I.33}$$

Described by [Altes 1976], and introduced by A. Grossmann as a "particularly good" wavelet shape in one of the most cited early paper in the theory of wavelets [Grossmann and Morlet 1984], it remained paradoxically anecdotal (not to say forgotten) since then, eclipsed by the non-analytic but famous Morlet wavelet, defined the same year in another important paper [Goupillaud, Grossmann, and Morlet 1984].

We show below its central role in the Morse family of analytic wavelets (although it does not seem to count among its members). We also discuss various criteria that lead us to select this shape among the continuum of Morse parameters, along with its remarkable properties and related quantities.

3.3.1 Symmetry of the wavelet across scales

A first criterion of choice consists in selecting the most symmetric Morse wavelet, in some sense, for a fixed quality factor $Q = \sqrt{\beta\gamma}$. The symmetry in the time domain being satisfied for the full Morse family, we look for a similarly convenient symmetry around the peak frequency. This question has been treated in [Lilly and Olhede 2012], where the parameter $\gamma = 3$, called the Airy wavelet, was found the most Gaussian and concentrated in Heisenberg sense Eq.(I.31). A key to this observation lies in the following Taylor expansion of the standardized Morse wavelet $\hat{\varphi}_{\beta,\gamma}(v)$ Eq.(I.25) near its peak at $v = 1$:

$$\log \hat{\varphi}_{\beta,\gamma}(v) = -\beta\gamma\frac{(v-1)^2}{2!} + \beta\gamma(3-\gamma)\frac{(v-1)^3}{3!} + \mathcal{O}\left(\frac{(v-1)^4}{4!}\right) \quad , \tag{I.34}$$

in which the term of order 3 vanishes for $\gamma = 3$. This expansion relies on an additive view of either the signal frequency f or the timescale variable a that composes the variable $v = af$ of $\hat{\varphi}_{\beta,\gamma}(v)$. We could argue that the local symmetry of the wavelet near its peak should be considered, in a time-frequency perspective, with respect to the wavelet transform frequency variable $f_a = a^{-1}$. The Taylor expansion that corresponds to an additive perturbation near $f_a = a^{-1}$ shares the same leading second order term, but the third order would vanish for $\gamma = -3$. Only, pairs of negative parameters (β, γ) were not considered in [Lilly and Olhede 2012].

In our opinion, neither of these symmetry arguments appears valuable for the wavelet transform, that intrinsically treat scales and frequencies in a multiplicative way, so that we should consider perturbing $\log f$ or $\log a$. An application that would require an additive approach of the frequency domain may rather use the Gabor transform (see Appendix 1). Surprisingly, the log-frequency expansion was also considered in [Lilly and Olhede 2012], as a corner case:

$$\log \hat{\varphi}_{\beta,\gamma}(v) = -\beta\gamma\frac{(\log v)^2}{2!} - \beta\,\gamma^2\frac{(\log v)^3}{3!} - \ldots - \frac{\beta}{\gamma}\frac{(\gamma \log v)^n}{n!}\ldots \tag{I.35}$$

All terms but the leading one vanish in the following scaling limit: $\gamma \to \pm 0$ and $\beta \to \pm\infty$ so that their product is fixed to a finite positive number. This explains the appearance of the log-normal wavelet in the Morse family, as the limit case:

$$\psi_Q = \lim_{\substack{\gamma\to 0\\ \beta\to\infty}} \varphi_{\gamma,\beta} \;, \quad 0 < \beta\gamma = Q^2 < +\infty \quad , \tag{I.36}$$

that is fully symmetric with respect to multiplicative scales a. An illustration of its central place in the Morse family is given in Fig.I.5 for variable γ and fixed quality factor.

The converse of this symmetric case are the fully asymmetric ones, $\gamma \to \pm\infty$ and $\beta \to \pm 0$ with $Q^2 = \beta\gamma > 0$, yielding sharp analytic low-pass and high-pass filters, equal to 1 for $v \in (0, 1]$ when $\gamma \to +\infty$ or for $v \in [1, \infty)$ when $\gamma \to -\infty$, and vanishing anywhere else. We can notice in Fig.I.5 that the interpretation of Q as a quality factor totally collapses in these asymmetric limit cases.

3.3.2 Power law measurement and wavelet independence to trends

The order of any wavelet ψ is defined as the largest integer number n_ψ that satisfies either of the time or frequency equivalent properties:

$$\int u^m \psi(u)\mathrm{d}u = 0 = \partial_v^m \hat{\psi}(0), \quad \forall m\,,\ 0 \le m < n_\psi \quad . \tag{I.37}$$

Such a wavelet is orthogonal to polynomial trends in the analysed signal up to the order $n_\psi - 1$. This property, that quantifies an order of independence to slow (vanishing frequency)

components in the analysed signal, is important when measuring power laws from the wavelet transform. For instance the regularity exponent $h(t)$ (such as the Hölder exponent h, or the Hurst exponent H in a stochastic context) of the signal x near time t depends on a vanishing scale variable $a \to 0$ as follows

$$x(at) \sim a^{h(t)} x(t) \tag{I.38}$$

$$\mathcal{W}_\psi[x](a,t) \sim a^{\min\{h(t), n_\psi\}}, \quad a \longrightarrow 0 \quad , \tag{I.39}$$

which means that the wavelet order n_ψ has to be large enough to measure all possible $h(t)$ [Venugopal et al. 2006]. Also note the importance of the norm convention $p = 1$ here (see the definition Eq.(I.6) of the wavelet transform). The higher the exponent $h(t)$, the more regular the signal around t. The white noise has a Hurst exponent $H = -\frac{1}{2}$, the pink noise $H = 0$ and the brown noise $H = \frac{1}{2}$.

Morse wavelets have a finite order $n_\psi = \beta > 0$ for all β, $\gamma > 0$, from their polynomial behaviour at zero frequency, but an infinite order for all β, $\gamma \leq 0$, from their stretched exponential (or log-normal) behaviour. Therefore,

$$n_{\psi_Q} = +\infty \tag{I.40}$$

for the log-normal wavelets ($\gamma = 0$). This kind of application usually employ wavelets with few oscillations (i.e. a small quality factor); we conclude that the quality factor can be chosen arbitrarily without raising any concern about the unknown order of regularity of the singular events in the signal.

In the time domain, the order of the wavelet is related to the asymptotic decay of its envelope at large times, which is $|\psi(u)| \sim |u|^{-(\beta+1)}$ for Morse wavelet of parameter β, $\gamma > 0$ [Lilly and Olhede 2009]. This means that the log-normal wavelet ψ_Q decays faster than polynomially in time, in frequency and in log-frequency, on both sides, making it especially well-localized and regular.

3.3.3 A close look at the widths of the log-normal wavelet

Numerous strategies are possible to define wavelet widths in time and frequency. All those that can be computed exactly for the log-normal wavelet, essentially from Gaussian integrals, are compared in Table I.1 with their meaning and dimensionless variables Eq.(I.16).

How to choose among all these possibilities? We can notice that they are all proportional, at least asymptotically for a large enough Q, to the following quantities:

$$\delta u = f \delta t = \frac{Q}{\sqrt{p\uparrow}} \qquad ; \qquad \delta \log v = \delta \log f = \frac{\sqrt{\uparrow}}{\sqrt{q} Q} \quad , \tag{I.41}$$

where both p and q refers to norm applied to the wavelet, and by extension, to the exponent at which the considered wavelet transform modulus is raised (1 in the linear case, 2 in the quadratic case). We call them the practical time and frequency resolutions for the log-normal wavelet. They are reduced versions of the Morse widths Eqs.(I.28, I.29), by a factor $\sqrt{\uparrow} \approx 2.5$ so that their product, the time-frequency uncertainty atom, is 1.

Apart from a constant coefficient, these widths can also differ by a factor:

$$\lambda = e^{\frac{1}{Q^2}} \quad , \tag{I.42}$$

related to the different conventions to define a central frequency. This factor is very close to 1 whenever the quality factor is not too small ($Q > 4$), hence omitted in the practical resolutions. We give it the symbol λ because it is a recurring dilation parameter the following.

symbol	*time*	*frequency*	*norm*	*value*	*meaning*
					weight
	$\psi(0)$	$\int \hat{\psi}(v)\mathrm{d}v$	1	$\frac{\sqrt{\Upsilon}}{Q}\lambda^{\frac{1}{2}}$	max time-modulus
$E_{\psi\psi}$	$\int \lvert\psi(u)\rvert^2\mathrm{d}u$	$\int \lvert\hat{\psi}(v)\rvert^2\mathrm{d}v$	2	$\frac{\sqrt{\pi}}{Q}\lambda^{\frac{1}{4}}$	energy
		$\int \lvert\hat{\psi}(v)\rvert^q\mathrm{d}v$	q	$\frac{\sqrt{\Upsilon}}{\sqrt{q}Q}\lambda^{\frac{1}{2q}}$	freq q-norm
C_ψ		$\int \overline{\hat{\psi}(v)}\mathrm{d}\log v$	1	$\frac{\sqrt{\Upsilon}}{Q}$	reconstruction coef
$C_{\psi\psi}$		$\int \lvert\hat{\psi}(v)\rvert^2\mathrm{d}\log v$	2	$\frac{\sqrt{\pi}}{Q}$	admissibility coef
		$\int \lvert\hat{\psi}(v)\rvert^q\mathrm{d}\log v$	q	$\frac{\sqrt{\Upsilon}}{\sqrt{q}Q}$	log-freq q-norm
					position
u_p	$\frac{\int u\lvert\psi(u)\rvert^p\mathrm{d}u}{\int \lvert\psi(u)\rvert^p\mathrm{d}u}$		p	0	any mean time
v_q		$\frac{\int v\lvert\hat{\psi}(v)\rvert^q\mathrm{d}v}{\int \lvert\hat{\psi}(v)\rvert^q\mathrm{d}v}$	q	$\lambda^{\frac{3}{2q}}$	arithmetic freq
		$\exp\frac{\int \log v\,\lvert\hat{\psi}(v)\rvert^q\mathrm{d}v}{\int \lvert\hat{\psi}(v)\rvert^q\mathrm{d}v}$	q	$\lambda^{\frac{1}{q}}$	geometric freq
v'_q		$\frac{\int \lvert\hat{\psi}(v)\rvert^q\mathrm{d}v}{\int v^{-1}\lvert\hat{\psi}(v)\rvert^q\mathrm{d}v}$	q	$\lambda^{\frac{1}{2q}}$	harmonic freq
$v_\infty = v'_\infty$		$\operatorname{argmax}\lvert\hat{\psi}(v)\rvert$	∞	1	peak freq
		$\frac{\int \log v\lvert\hat{\psi}(v)\rvert^q\mathrm{d}\log v}{\int \lvert\hat{\psi}(v)\rvert^q\mathrm{d}\log v}$	q	0	any mean log-freq
					width
Δu_2	$\left(\frac{\int u^2\lvert\psi(u)\rvert^2\mathrm{d}u}{\int \lvert\psi(u)\rvert^2\mathrm{d}u}\right)^{\frac{1}{2}}$	$\left(\frac{\int \lvert\partial_v\hat{\psi}(v)\rvert^2\mathrm{d}v}{\int \lvert\hat{\psi}(v)\rvert^2\mathrm{d}v}\right)^{\frac{1}{2}}$	2	$\frac{\sqrt{1+2Q^2}}{2\Upsilon}$ $\sim \frac{Q}{\sqrt{2\Upsilon}}$	time std
Δv_q		$\left(\frac{\int v^2\lvert\hat{\psi}(v)\rvert^q\mathrm{d}v}{\int \lvert\hat{\psi}(v)\rvert^q\mathrm{d}v} - v_p^2\right)^{\frac{1}{2}}$	q	$\sqrt{\lambda^{\frac{4}{q}} - \lambda^{\frac{3}{q}}}$ $\sim \frac{1}{\sqrt{qQ}}$	freq std
$\lvert\hat{\psi}(v_\pm)\rvert = \frac{1}{2}$		$v_+ - v_-$	q	$2\sinh\frac{\sqrt{\log 2}}{\sqrt{q}Q}$ $\sim \frac{2\sqrt{\log 2}}{\sqrt{q}Q}$	freq FWHM
$\lvert\hat{\psi}(v_\pm)\rvert = \frac{1}{2}$		$\log\frac{v_+}{v_-}$	q	$\frac{2\sqrt{\log 2}}{\sqrt{q}Q}$	log-freq FWHM
$\Delta u = f\Delta t$	$\left(\frac{\int u^2\psi(u)e^{-i\Upsilon u}\mathrm{d}u}{\int \psi(u)e^{-i\Upsilon u}\mathrm{d}u}\right)^{\frac{1}{2}}$	$\left(-\frac{\partial_v^2\hat{\psi}(1)}{\Upsilon^2\,\hat{\psi}(1)}\right)^{\frac{1}{2}}$	1	$\frac{Q}{\Upsilon}$	time demod std
$\Delta v = \frac{\Delta f}{f}$ $\Delta\log v = \Delta\log f$		$\left(-\frac{\partial_v^2\hat{\psi}(1)}{\hat{\psi}(1)}\right)^{-\frac{1}{2}}$	1	$\frac{1}{Q}$	freq peak curvature log-freq std
N_ψ	$\frac{\Im\{\int \partial_u\psi(u)\,\overline{\psi(u)}\mathrm{d}u\}}{\Upsilon\lvert\psi(0)\rvert^2}$	$\frac{\int v\lvert\hat{\psi}(v)\rvert^2\mathrm{d}v}{\lvert\int \hat{\psi}(v)\mathrm{d}v\rvert^2}$	2	$\frac{Q}{\sqrt{2\Upsilon}}$	equivalent nb full oscillations
$e^{H_{\frac{1}{2}}[\lvert\hat{\psi}\rvert^2]} = e^{-H_\infty[\lvert\psi\rvert^2]}$	$\frac{\lvert\psi(0)\rvert^2}{\int \lvert\psi(u)\rvert^2\mathrm{d}u}$	$\frac{\lvert\int \hat{\psi}(v)\mathrm{d}v\rvert^2}{\int \lvert\hat{\psi}(v)\rvert^2\mathrm{d}v}$	2	$\frac{\sqrt{2\Upsilon}}{Q}\lambda^{\frac{3}{4}}$	exp Rényi entropy
$\delta u = f\delta t$			p	$\frac{Q}{\sqrt{p\Upsilon}}$	practical time resol
$\delta\log v = \delta\log f$			q	$\frac{\sqrt{\Upsilon}}{\sqrt{q}Q}$	practical freq resol

Table I.1: *Comparison of strategies to define the time-frequency localization of a standardized analytic wavelet, applied to the log-normal wavelet $\psi = \psi_Q$. Only quantities that have a frequency-domain expression (third column), hence that are computationally tractable (as Gaussian integrals), are presented. Expressions are simplified using $\lambda = e^{\frac{1}{Q^2}}$, and the large quality factor asymptotics (preceded by $\sim$) is provided to facilitate the comparison when Q is large enough. Abbreviations: standard deviation (std), full width at half maximum (FWHM).*

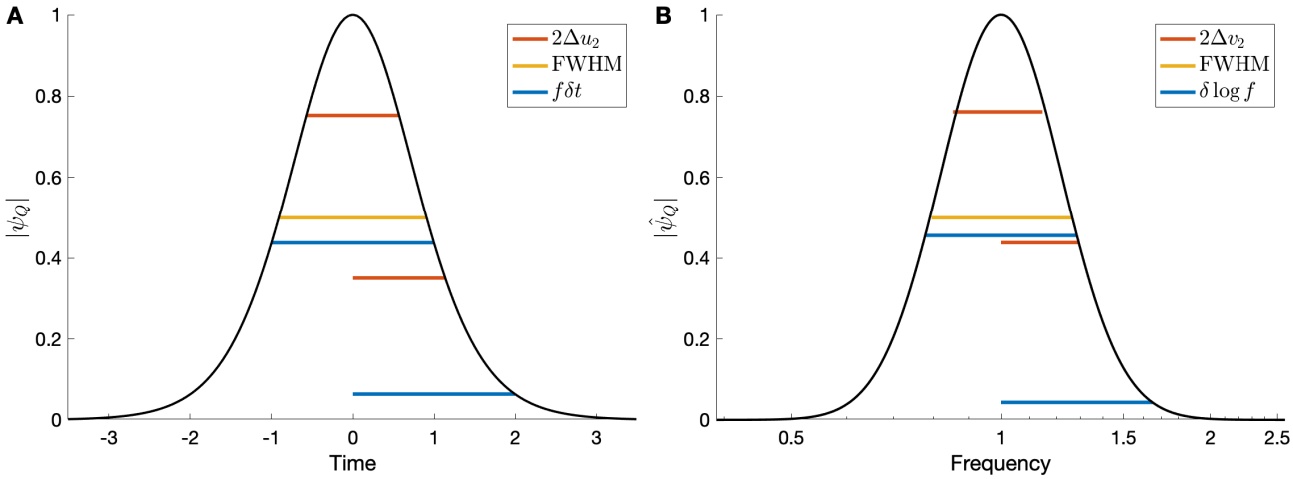


Figure I.6: *Comparison of the widths from the Heisenberg uncertainty relation $\Delta u_2, \Delta v_2$ (red), the full width at half maximum (FWHM, yellow), and the practical resolutions $f\delta t = (\delta \log f)^{-1} = \frac{Q}{\sqrt{\top}}$ (blue) for the log-normal wavelet ψ_Q with quality factor $Q = 5$. (A) Wavelet modulus in the time domain. (B) Wavelet modulus in the frequency domain.*

In Fig.I.6, these practical widths (blue segment) are compared with other common definitions, as wavelet radii and diameters of the log-normal wavelet (its norm $p = 1 = q$). For the quality factor $Q = 5$, intensively used in later applications, practical resolutions turn out to be very close to the full widths at half maximum (FWHM). In comparison, the Heisenberg widths appear impractical to represent the wavelet widths in the norm 1.

3.3.4 Time derivative: quarter of turn and small scaling

As for Morse wavelets in Eq.(I.20), successive time-derivatives of the log-normal wavelet yield a new log-normal wavelet. But contrary to Morse wavelet, its only parameter Q does not change! Instead, it is simply scaled so that the new peak is at λ^n, where n is the order of the derivative:

$$v^n \hat{\psi}_Q(v) = \lambda^{\frac{n}{2}} \hat{\psi}_Q(\lambda^{-n} v) \quad \Leftrightarrow \quad \partial_u^n \psi_Q(u) = \left(i\top\lambda^{\frac{3}{2}}\right)^n \psi_Q(\lambda^n u) \quad . \tag{I.43}$$

The factor i^n changes the phase of the wavelet (by a quarter of turn for $n = 1$), as for a pure wave. Note that the scaling of the wavelet is "small" compared to the wavelet resolution: $\log \lambda = \frac{1}{Q^2} \ll \delta \log f = \frac{\sqrt{\top}}{Q}$, for Q not too low.

This property is very practical because the wavelet transform of signal derivatives:

$$\mathcal{W}_\psi[\partial_t^n x](a,b) = \partial_b^n \mathcal{W}_\psi[x](a,b) = a^{-n} \mathcal{W}_{(-\partial)^n \psi}[x](a,b) \quad , \tag{I.44}$$

is proportional to the rescaled original wavelet transform and needs not be recomputed:

$$\mathcal{W}_{\psi_Q}[\partial_t^n x](a,b) = \left(\frac{i\top\lambda^{\frac{1}{2}}}{a}\right)^n \mathcal{W}_{\psi_Q}[x](\lambda^{-n} a, b) \quad . \tag{I.45}$$

Before the formulation of wavelet theory, this fact was already known in signal processing with constant bandwidth filters. The invariance to time differentiation of the signal of log-normal spectrum, together with its good locality, makes it an exceptional choice of impulse signal to design a sonar [Altes 1976]. So that this wavelet is observed in animal echo-location, up to a certain chirp modulation: n can be complex in $v^n \hat{\psi}_Q(v)$, without changing the log-normal modulus.

3.3.5 Recovering the orthogonal family

The scaling limit in which the log-normal is obtained from the standardized Morse wavelet of order $k = 0$, see Eq.(I.35), can also be applied to the full family of orthogonal Morse wavelets $k \geq 0$. In this limit, Laguerre polynomials turn into Hermite polynomials of the log-frequency:

$$\hat{\varphi}_Q^{(k)}(v) = e^{-\frac{1}{2}(Q\log v)^2} H_k(\tfrac{1}{2Q} - Q\log v) \quad , \quad H_k(x) = \sum_{\ell=0}^{\lfloor k/2 \rfloor} \frac{k!}{(k-2\ell)!\ell!}(2x)^{k-2\ell} \quad . \tag{I.46}$$

This is proven in Appendix 5, using the generating function of both families of orthogonal polynomials.

There are in fact two distinct limits, confounded for $k = 0$ and hidden in the above result, depending on the sign of the Morse parameters, see the inversion property Eq.(I.18): the positive case (β, $\gamma > 0$), and the negative case (β, $\gamma < 0$). We get the latter case replacing Q by $-Q$ in the above expression (hence denoted $\hat{\varphi}_{-Q}^{(k)}$), which can be understood as the negative square root of $Q^2 = \beta\gamma$ in Eq.(I.27).

The shift $\frac{1}{2Q}$, that corresponds to the frequency scaling $\lambda^{-\frac{1}{2}}$, arises to enforce the orthogonality in additive time and frequency:

$$\int_0^\infty \hat{\varphi}_{\pm Q}^{(m)}(v)\hat{\varphi}_{\pm Q}^{(n)}(v)\mathrm{d}v = n!2^n\frac{\sqrt{\pi}}{Q}e^{\frac{1}{4Q^2}}\delta_{m,n} = \int \varphi_{\pm Q}^{(m)}(u)\overline{\hat{\varphi}_{\pm Q}^{(n)}(u)}\mathrm{d}u \quad . \tag{I.47}$$

This simplifies when we are interested instead in the orthogonality with respect to log-scales or log-frequencies:

$$\int_0^\infty e^{-(Q\log v)^2} H_m(\mp Q\log v)H_n(\mp Q\log v)\mathrm{d}\log v = n!2^n\frac{\sqrt{\pi}}{Q}\delta_{m,n} \quad , \tag{I.48}$$

leading to the slightly simpler and alternative definition of "log-Hermite wavelets $\psi_{\pm Q}^{(k)}$:

$$\hat{\psi}_Q^{(k)}(v) = e^{-\frac{1}{2}(Q\log v)^2} H_k(-Q\log v) \quad . \tag{I.49}$$

These 4 lines of orthogonal families (2 for the positive and negative limits, and 2 with respect to linear or logarithmic frequencies) all agree on the order zero: $\psi_Q = \varphi_{\pm Q}^{(0)} = \psi_{\pm Q}^{(0)}$.

3.3.6 Final thoughts on log-normal wavelets: before Morse, after Morlet, near Cauchy

There is a widespread idea in the literature, that the Morse wavelet and its generalized Laguerre orthogonal family are to the wavelet transform what the Gaussian window and its Hermite orthogonal "tapers" are to the Gabor (short-time Fourier) transform [Bayram and Baraniuk 1996; Daubechies and Paul 1988; Flandrin 1988]. We think that the log-normal wavelet of Altes and Grossmann alone should enter this analogy. Introducing a secondary "asymmetry" parameter to the more important quality factor parameter, generalized Laguerre polynomials are a sophistication, that could as well be mapped in the context of the Gabor transform using exponential frequencies, inverse of the mapping of Hermite polynomials in the logarithmic scales in Eq.(I.48).

There are different reasons to the curious absence of popularity of Grossmann's log-normal wavelet, the first of which is the lack of simple expression in the time domain, contrary to the Morlet wavelet: a complex wave modulated by a Gaussian window. Thus, the Morlet wavelet inherited the intuition of the Gabor window in the wavelet domain, even though it is not analytic. This conceptual limitation gave room to the Cauchy wavelets ($\gamma = 1$), and their

beautifully simple expressions both in time and frequency domains, then generalized to all Morse wavelets. Their faster (exponential) decay at high frequency may have been perceived as an advantage over the log-normal shape for certain reconstruction applications, although this is mitigated by a slower polynomial decay in the time domain. Of different analytical properties, the closeness of the Cauchy and log-normal wavelets can be noticed in Fig.I.5, which translates in practice into indistinguishable wavelet transforms as soon as Q is large enough (more than 4 or 5).

Devoted to a logarithmic (i.e. multiplicative) view of scales (and to their Fourier-conjugated domain called the Mellin domain), the log-normal wavelet reflects in the time domain in a still mysterious way. It has been described in the original paper [Grossmann and Morlet 1984] as a special function generalizing Gamma function and Gaussian integrals with nice algebraic properties. In a probabilistic context, it is known as the characteristic function of the log-normal distribution. Bridging multiplicative frequencies and additive times, an approximation involving the Lambert W function is known [Asmussen, Jensen, and Rojas-Nandayapa 2016], derived from a stationary phase or Laplace's method (also used in Appendix 8) for a large enough —but in fact rather small— quality factor. In any case, all its numerical implementations use the definition of the wavelet transform Eq.(I.7) from the frequency domain.

Summary of section 3

In spite of the complex helical shape of the wavelet in the time domain, the time-frequency interpretation of the wavelet transform is uneasy. While the time variable is clearly identified as the translation variable b, the correspondence between the inverse scale variable and the frequency domain heavily relies on the shape of the wavelet spectrum.
The Morse wavelet is a general family of parametric shapes for the spectrum of analytic wavelets, $\varphi_{\beta,\gamma}$, with two shape parameters. In the frequency domain, this shape is a positive band-pass filter, more or less broad and asymmetric. It can be further expanded into a series of orthogonal variations, indexed by an integer order. The order zero has a unique maximum, whose frequency can be standardized to 1 Hz. This standardization of the wavelet, i.e. its rescaling, ensures an approximate correspondence between the scale variable a of the wavelet transform and an inverse frequency f^{-1}.
The precision of this approximate scale-to-frequency relation increases with the alternative parameter $Q = \sqrt{\beta\gamma}$. Interpreted as a quality factor, the parameter Q is sufficient to specify the time and frequency widths of the Morse wavelet. It fixes a trade-off between time and frequency localization, i.e. a relation between the time-domain and frequency-domain uncertainties. The remaining degree of freedom controls symmetry and regularity properties of the wavelet shape, which are especially convenient in a particular limit: $\gamma \to 0$ and $\beta \to \infty$, while Q is fixed. This leads us to select the resulting one-parameter wavelet ψ_Q, that has a log-normal spectrum, symmetric in the log-frequency domain with respect to $\log 1 = 0$.
Therefore, we fix the scale-to-frequency relation to $a = f^{-1}$, that differs by a small factor $\lambda^\alpha = e^{\frac{\alpha}{Q^2}}$ from alternative definitions of the wavelet frequency, together with practical time and frequency widths. These widths are the time-frequency resolution of the wavelet transform: $f\delta t, \delta \log f$ fixed to $\frac{Q}{\sqrt{\pi}}$ and its inverse respectively. This results into a precise and explicit control of the time-frequency interpretation for the selected wavelet transform, from the specification of the quality factor Q.

Chapter II

Oscillations and fluctuations in complex natural recordings

In chapter I, we have carefully selected a single-parameter analytic wavelet to transform complex recordings and represent them into the time-frequency plane. These two dimensions serve to dilute the complexity, i.e. non-stationary and multi-scale characteristics of natural signals such as a voice recording or an electroencephalography (EEG), into a suitable representation.

It is now time to simplify the notations and summarize important properties of this wavelet transform in the first section and to revisit quadratic expressions in the next sections, with a particular emphasis on the role of the quality factor. This covers essential time-frequency properties [Carmona, Hwang, and Torrésani 1998; Flandrin 1998b; Torrésani 1995], applied to the log-normal wavelet framework. In particular, their probabilistic interpretation as statistical estimators is developed in the third section and the fourth is dedicated to the analysis of correlation and coherence. This development culminates in the last section with the general problem of controlling both the significance of statistical estimators and their time-frequency resolution; problem that beautifully simplifies into the introduction of a second quality factor. This exact result constitutes a bridge between the undulatory and stochastic descriptions of complex natural signals, either oscillating or fluctuating.

The collection of properties and relations in this second chapter are the basic blocks of the wavelet framework, that will serve as a dictionary for the following chapters.

1 Time-frequency recording: properties and illustrations

First, we denote the time-frequency representation of a recording x, possibly multidimensional, based on the log-normal analytic wavelet ψ_Q introduced in Eq.(I.33), as follows:

$$X(t,f;Q) = \mathcal{W}_{\psi_Q}[x](f^{-1},t) \ , \quad \hat{\psi}_Q(v) = \Theta(v) e^{-\frac{1}{2}(Q \log v)^2} \quad , \tag{II.1}$$

where $(t,f) \in \mathbb{R}^2$, $Q \in \mathbb{R}^+$ and we use the convention $p = 1$ in the definition Eq.(I.6) of the wavelet transform. We will loosely refer to it as the wavelet transform of x, importantly but sometimes implicitly, with quality factor Q. If needed, this time-frequency notation may also be extended to any other standardized analytic wavelet ψ as $X_\psi(t,f)$, or by giving its explicit parameters such as $X(t,f;\pm Q,k)$ for the log-Hermite orthogonal family $\varphi^{(k)}_{\pm Q}$ defined in Eq.(I.46).

1.1 Representations of the wavelet transform

We illustrate the time-frequency representation $X(t,f;Q)$ for a pedagogical signal in Fig.II.1 and for natural recording in Fig.II.2.

1.1.1 Chirp and pink noise

The signal represented in Fig.II.1(A) is the sum $x(t) = s(t) + \xi_0(t)$ of a deterministic oscillation $s(t)$ with increasing frequency (a chirp) and decreasing amplitude, and a pink noise $\xi_0(t)$ of lower amplitude (the "1/f" noise, in reference to the scaling of its power spectral density). Since $X(t, f; Q) = \frac{1}{2}A(t, f)e^{i\phi(t,f)}$ takes complex values, it is represented by the amplitude A (or its square) and the phase ϕ. The factor 2 can be computed from a simple wave $A\cos(\Upsilon f_0 t)$ for which the modulus of the wavelet transform is $\frac{A}{2}\hat{\psi}_Q(f)$, i.e. at most $\frac{A}{2}$ for $f = f_0$.

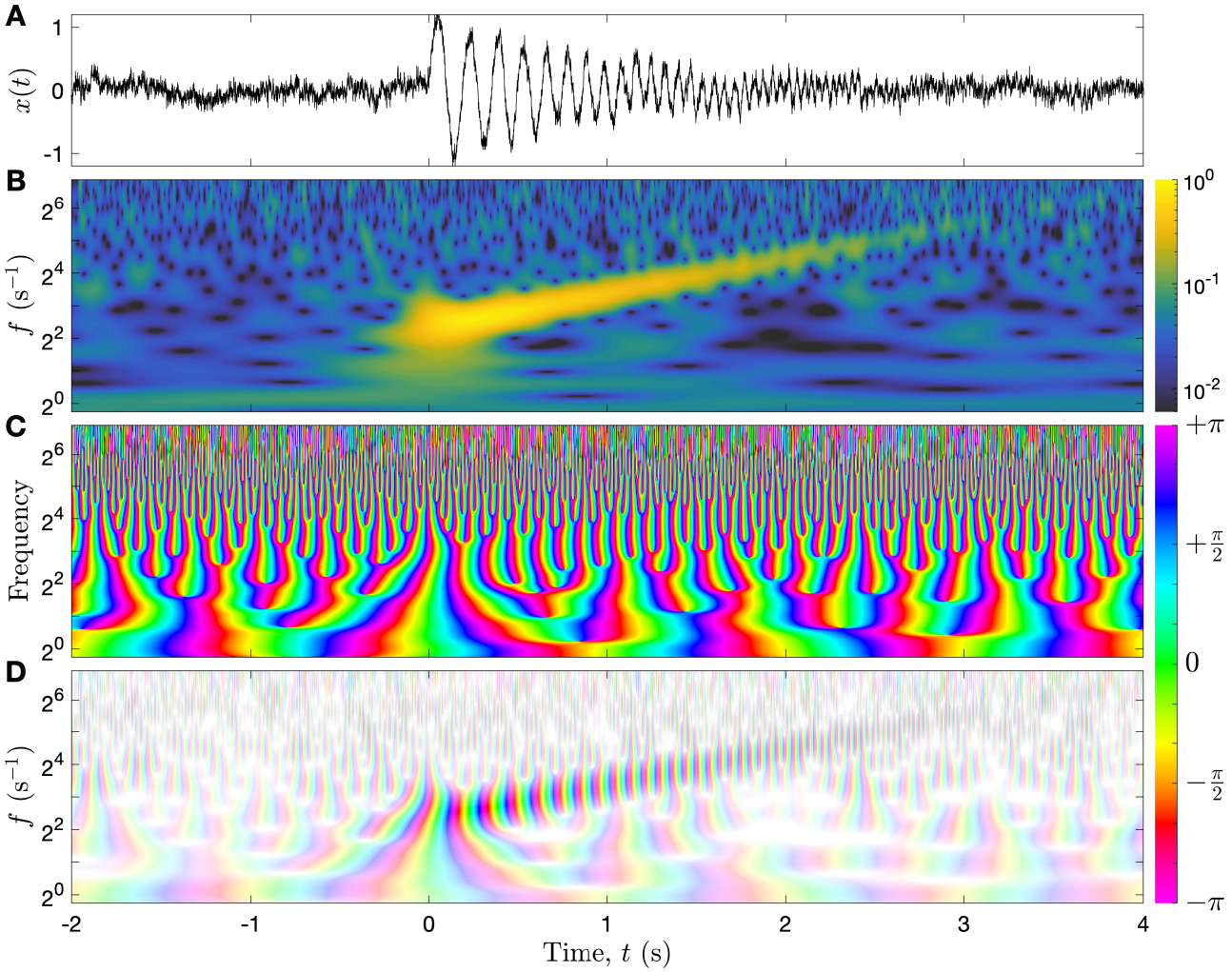


Figure II.1: *Graphical representations of the wavelet transform for a pedagogical signal $x(t)$. The signal $x(t)$ is represented in (A), and its wavelet transform, Eq.*(II.1)*, of quality factor $Q = 5$ is computed. The modulus is colour-coded in a logarithmic scale in (B), and the angle (complex argument) with the chromatic circle in (C). Note that we represented twice the modulus of the wavelet transform in (B) so that it is directly comparable to the amplitude of the signal (A). Both dimensions of the complex values taken by the wavelet transform are combined in (D) with a two-dimensional colour coding: the angle is still associated to a hue in the chromatic circle but the modulus is now coded by the saturation of the colour.*

The image in Fig.II.1 (B) is the modulus of the wavelet transform, that is maximum at or very close to the time-dependent frequencies of the chirp, for $t > 0$. The amplitude of the chirp decreases while its local frequency increases in time. With a perceptually uniform colour-coding [Kovesi 2015] of $A(t, f) = 2|X(t, f; Q)|$ we observe that the maximum amplitude of the chirp in Fig.II.1 (B) matches closely the amplitude of the oscillation in Fig.II.1 (A). Since we code the amplitude on a logarithmic scale, the image of its square, called the scalogram,

is identical. Regions of lower amplitude fluctuate around a constant value at any time and frequency, that corresponds to the amplitude of the pink noise $\xi_0(t)$. This is a specificity of the pink noise: it has a constant power density per decade or per octave and has a strong physiological interest since it has been proposed to describe the scale invariance of many natural stochastic signals, such as EEG [Allegrini et al. 2009; He 2014; He et al. 2010; Lee et al. 2002]. We refer to section 3.3.2 and Fig.II.3 for more details.

The next image in Fig.II.1 (C) represents the angle of $X(t, f; Q)$, i.e. the complex argument $\phi(t, f) = \Im\{\log X(t, f; Q)\}$, which is conveniently represented with the hues in the chromatic circle since it is a circular quantity. At the maximum amplitude of the chirp, this angle corresponds to the phase of the signal. In this work, the angle 0 is represented in green, $\pm\pi = \pm\frac{1}{2}\Upsilon$ is in magenta and the interval from $-\pi$ to $+\pi$ follows the progression of the colours in the visible light spectrum (apart from the magenta, which is not in the physical spectrum since it closes the circle). The angle in time and frequency has a particular behaviour: it always increases continuously and monotonically in time, at a rate that is consistent with the frequency f: $\frac{\partial}{\partial t}\phi(t, f) \approx \Upsilon f$. To satisfy this condition both at low and high frequencies, the lines of constant phase have a tree structure, branching towards higher frequencies. The condition fails near the branching singular points or phase defects (vortices of unit charge), for which the amplitude vanishes. No such singular point is observed in the high amplitude region associated to the chirp, where the angular progression directly represents the phase of its oscillation.

The last image in Fig.II.1 (D) combines both the modulus and the angle of the complex value $X(t, f; Q)$ in a two-dimensional colour map. This type of colour coding, possible because the colour space is at least two-dimensional (three-dimensional for at least 96% of human beings), could be represented in polar coordinates (in $\mathbb{C}$) as a chromatic disc where the angle is the hue and the radius (modulus) is the saturation of the colour (there is no defined hue/angle if the saturation/modulus is vanishing). Here the colour of vanishing modulus is set to white, the low amplitude of the noisy regions indeed appears with very faint and pastel colours, whereas the chirp has a more intense colour. Freely inspired from the *domain colouring* techniques in complex analysis, this colour coding is a synthetic way of representing a map of complex values at the scale of few oscillations (otherwise the colours would hardly be distinguishable).

1.1.2 Skin temperature signal

As a first example of complex natural recording, we chose a non-contact measure of the skin temperature fluctuations. This type of signal is obtained from a dynamic infrared thermogram of the chest of volunteers, studied to improve the early diagnostic of breast cancer in [Gerasimova et al. 2014; Gerasimova-Chechkina et al. 2016]. The spatial information is reduced by averaging each frame of the thermogram (sampled at 50 Hz) inside a region of interest, here the entire left breast of subject 8, see Appendix 1. The temperature signal is shown in Fig.II.2 (A). The origin of its fluctuations is complex to interpret, as they result not only from the actual variations of the global skin temperature, but also from the unsteady geometric projection of the body in the infrared camera.

Physiological rhythms are visible as horizontal lines in the scalogram (or amplitude) Fig.II.2 (B), scattered but with a high amplitude for breathing (at low frequencies), and with a low amplitude but a precise and constant higher frequency for the heart beat. The higher quality factor $Q = 10$ helps to visualize rhythmic lines. A sudden drop in the temperature signal around 470 s is marked by a vertical line. This is an artefact, possibly associated to an important motion of the subject.

All these features are reflected more subtly into the phase patterns, Fig.II.2 (C), which are harder to interpret. The fast progression of the phase at a high frequency f blurs its visualization at large timescales compared to the period f^{-1}, and generates moiré patterns.

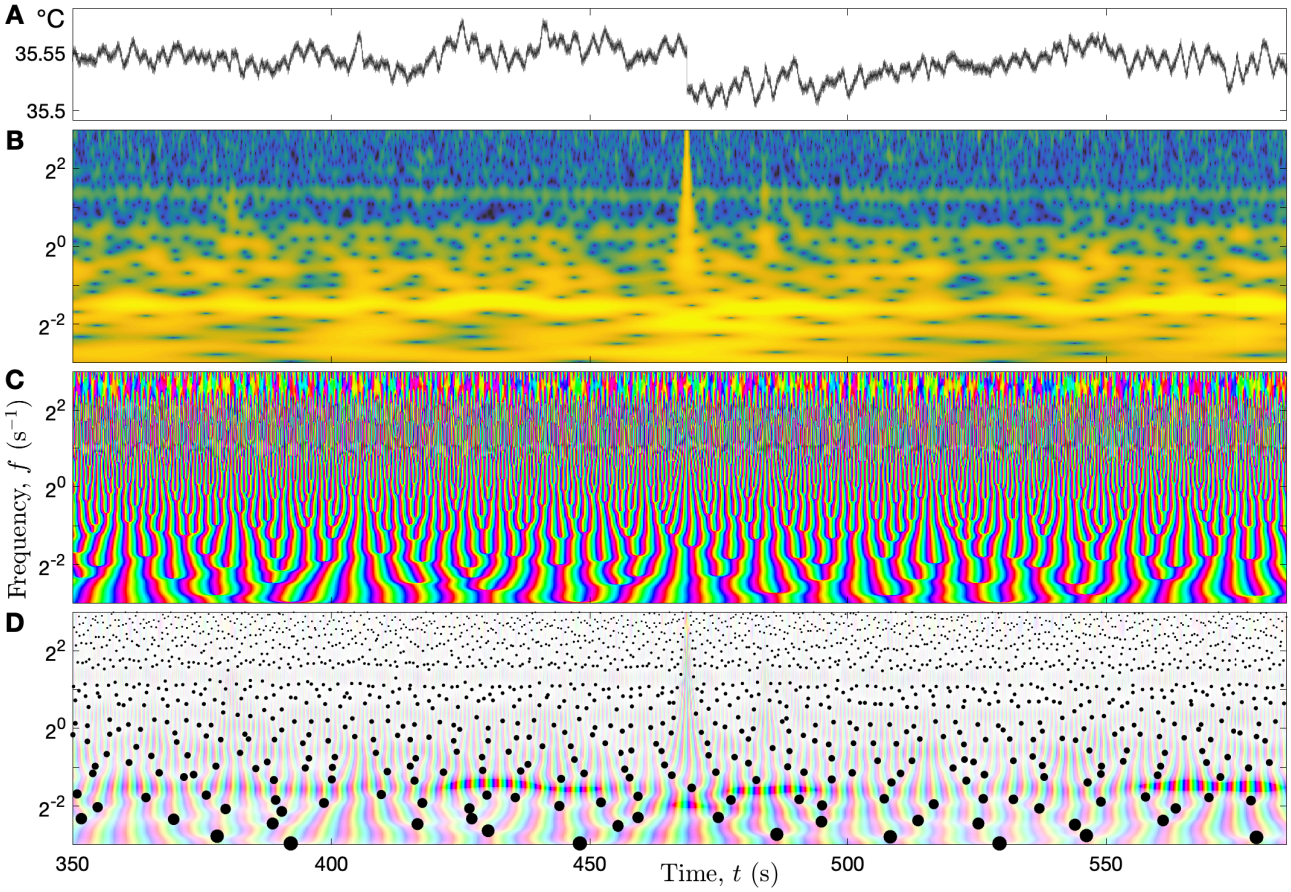


Figure II.2: *Evolution of the skin temperature, as derived from a dynamic infrared thermogram. (A) Mean temperature signal (in °C), and its wavelet transform of quality factor $Q = 10$ is computed: (B) amplitude, (C) angle (complex argument), and (D) both, colour-coded as in Fig.II.1. Singularities (amplitude zeros, phase vortices) are represented in (D) as black spots, with a constant size with respect to the time resolution.*

This representation of the phase is not suitable at the scale of numerous oscillations.

1.1.3 Between wavelets, vortices: a hyperbolic sea

We propose in Fig.II.2 (D) to locate the singularities or phase vortices, in order to summarize the behaviour of the phase: they are repelled from time-frequency features of the signal. A clear horizontal channel, without singularities, is present at a constant frequency, about 2.5 Hz, associated to the double heart beat. Although the corresponding amplitude is very weak compared to the breathing oscillation, no clear channel is observed at lower frequencies, except intermittently. The time-localized artefact also gives rise to a repulsion of the phase defects out of its vertical region.

The global density of these singular points appear to correspond to one per time-frequency atoms of uncertainty. This is a topological constraint of the phase branching, that creates a singular point per Hz and per s. It is also verified quite locally, even in regions with time-frequency "features" (chirp, rhythmic lines) that are depleted but surrounded by accumulated vortices. This depletion is as strong as the signal-to-noise ratio of these features is important, so that it can be used to differentiate the "signal" from the "noise" [Flandrin 2015]. In the case of a Gaussian stochastic process, these "particles" has been shown repel each other, following the random distribution of a determinantal point process in [Bardenet, Flamant, and Chainais 2020; Bardenet and Hardy 2021] for some time-frequency representations (based on the Cauchy

wavelet and the Gaussian Gabor window).

For the log-normal wavelet, we simulated numerically this density, which is not exactly 1, as could be anticipated from the alternative definitions of the frequency variable f, that vary by a small factor $\lambda^\alpha = e^{\frac{\alpha}{Q^2}}$, see Table I.1. As test signals $x(t)$, we used self-similar Gaussian noises of Hurst exponent H (see discussion in section 3.3). Their numbering N from numerical contour integration of the phase of $X(t, f; Q)$, enclosing a large time-frequency region Ω of area $\|\Omega\| = \iint_\Omega \mathrm{d}t\mathrm{d}f$, yields $\frac{N}{\|\Omega\|} = \lambda^{\frac{1}{6}-H}$. It depends on both the noise colour H and quality factor Q, with few percent of uncertainty on the exponent for large regions Ω. The fraction $\frac{1}{6}$ is surprising since it has no corresponding exponent in Table I.1, and no relation to a similar critical density (of sampling for wavelet frames) conjectured to be $\frac{E_{\psi\psi}}{C_{\psi\psi}}$ [Seip 1989; Seip 1993] i.e. equal to $\lambda^{\frac{1}{4}}$ for ψ_Q, and recently proved correct for Cauchy wavelets [Abreu and Speckbacher 2020].

A single difference in this "gas" of singularities between the Gabor and the wavelet transforms has been highlighted in [Bardenet, Flamant, and Chainais 2020]: it consists in the geometry associated to the time-frequency plane, Euclidean (flat) in the first case and hyperbolic in the second. The negative curvature of the hyperbolic case has the specificity of introducing a natural unit of area, that can be identified with the time-frequency atom of uncertainty and the inverse density of singular points. We can be convinced of these geometrical facts from two interpretations of the surface element $\mathrm{d}t\mathrm{d}f$. While its Euclidean interpretation is immediate, its hyperbolic interpretation can be reached considering the wavelet resolutions $f\delta t$ and $\delta \log f$ Eq.(I.41): i.e. rewriting it $f\mathrm{d}t\frac{\mathrm{d}f}{f}$ $(f > 0)$. The use of a timescale variable $\tau \propto f^{-1}$ leads to the Poincaré half-plane model of hyperbolic geometry, with the metric $\mathrm{d}s^2 = \frac{\mathrm{d}t^2+\mathrm{d}\tau^2}{\tau^2}$.

The choice of the log-frequency axis in the time-frequency representation flips and stretches the Poincaré model, without changing the hyperbolic nature of the geometry. It is consistent with the resolution $\delta \log f$, but compresses $f\delta t$ at increasing frequencies, so that it gives the illusion of an increasing density of vortices. Like the spherical geometry on a planisphere, this deformation cannot be avoided in any flat illustration. To alleviate this graphical difficulty, we adapted the markers of the singular points, black spots in Fig.II.2 (D), so that their size remains constant with respect to $f\delta t$. As when looking at a boat: near the coast or far in the sea. The hyperbolic geometry of the wavelet perspective on time and frequency is the key for a truly multi-scale representation.

1.1.4 A note on the numerical implementation of the continuum

The hyperbolic geometry is also the key for an efficient numerical implementation: the discrete counterpart of the hyperbolic plane is indeed the binary tree structure, or dyadic grid, central to minimize the redundancy of wavelet transforms defined on an orthogonal wavelet basis. We refer to [Selesnick 2001, 2002; Selesnick, Baraniuk, and Kingsbury 2005] for the closest realization of an analytic wavelet transform in the world of discrete wavelet bases. The quality factor associated to these discrete wavelet transforms is fixed, as can be estimated from the size of the time-frequency atom in the associated dyadic grid: the scale sampling is once per octave $\delta \log f \sim \log 2$, and the time-step is the scale parameter, i.e. $f\delta t \sim 1$, which yields the estimate $Q \in [2.5, 3.6]$.

In applications, we consider greater and adjustable values of the quality factor Q, we care about the time-frequency continuum of physics and the oscillatory interpretation associated to the analytic wavelet. In agreement with Eq. (I.41), we sample in log-frequency every fraction of the resolution $\delta \log f$ (small for smoothness, close to 1 for speed), but linearly in time for the computational convenience of numerical matrices. However, when the frequency range of interest is far greater than few octaves, a time sampling that is sufficient at high frequency is a waste of computational resources (especially memory) at low frequency. It is much more

efficient to adjust the time resolution proportionally to the frequency, as a fraction of δt to conserve smoothness.

For resource intensive applications, this has led us to divide the time-frequency plane in horizontal bands wide of an octave and discretize them on a log-frequency-linear-time grid. Then, we downsample the linear-time grid by a factor 2 each time we change to the octave below, as for a discrete wavelet basis. The wavelet transform is then evaluated numerically as series of complex matrices, one for each band, applying the wavelet filter to the recording x between two fast Fourier transforms (one to get the spectrum $\hat{x}(f)$, the other for the transform Eq.(I.7)).

1.2 Useful properties

We sum up here the elementary properties of the wavelet transform outlined in Eq.(II.1). Let us recall explicitly how it is computed from Fourier transform of the signal $x(t)$ Eq.(I.1), i.e. from the spectrum $\hat{x}(f')$:

$$X(t,f;Q) = \int \hat{x}(f')\Theta(f'/f)e^{i\Upsilon f' t - \frac{1}{2}(Q\log\frac{f'}{f})^2}\mathrm{d}f' \quad . \tag{II.2}$$

Here, the step function Θ allows selecting positive spectral components for $f > 0$. Negative ones, $f < 0$, are useful to describe general complex signals. In the following, all integrals are over $\mathbb{R}$, unless specified otherwise with integration bounds.

1.2.1 Linearity, shift, scaling and physical unit

The wavelet transform is a linear transformation, it behaves just as the signal and the spectrum when they contain superimposed components:

$$ax(t) + by(t) \qquad \longleftrightarrow \qquad aX(t,f;Q) + bY(t,f;Q) \qquad \longleftrightarrow \qquad a\hat{x}(f) + b\hat{y}(f) \quad , \tag{II.3}$$

for any constants a, b.

Now consider the common situation of a translated and dilated recording. For instance, an experiment is recorded as a signal $x(t)$. The recording starts at time t_0 and is sampled at frequency f_s, hence it is a discrete time series x_n of time step $\mathrm{d}t = f_s^{-1}$. The signal $x(t)$ that represents the physical experiment is obtained from the time series as $x(\mathrm{d}t\, n + t_0) = x_n$. For any real scale and shift parameters c, d, the wavelet transform behaves just like the signal:

$$x(ct+d) \qquad \longleftrightarrow \qquad X(ct+d, f/c; Q) \qquad \longleftrightarrow \qquad \hat{x}(f/c)e^{i\Upsilon f d}/c \quad . \tag{II.4}$$

The frequency variable is also scaled, inversely to the time variable in agreement with their inverse physical units. But, there is no supplementary phase and scale factor as for the spectrum.

This correspondence of the wavelet transform with the signal, especially for the scaling, is a consequence of the convention $p = 1$ in the definition Eq.(I.6). This is a physical choice which enforces a direct agreement between the modulus of the wavelet transform and the amplitude of the signal (more specifically of its analytical version, see section 1), and gives them the same physical unit. In contrast, the spectrum (Fourier transform) is a time integral that departs from the original signal's unit. The linear view of the frequency domain commonly attached to it is reassessed in this thesis.

With any other convention p, the wavelet transform at location $(ct+d, f/c)$ would correspond to $x(ct)c^{1-\frac{1}{p}} \leftrightarrow \hat{x}(f/c)c^{-\frac{1}{p}}$.

1.2.2 Complex conjugation and Hermitian symmetry

In the general complex case ($x \in \mathbb{C}$), complex conjugation behaves as the signal and the spectrum:

$$\overline{x(t)} \qquad \longleftrightarrow \qquad \overline{X(t,-f;Q)} \qquad \longleftrightarrow \qquad \overline{\hat{x}(-f)} \quad . \tag{II.5}$$

This property results from the realness of the wavelet spectrum (otherwise the wavelet is changed by the conjugation).

Therefore, the Hermitian symmetry of the spectrum $\hat{x}(-f) = \overline{\hat{x}(f)}$ in the real case ($x \in \mathbb{R}$) is transmitted to the wavelet transform:

$$X(t,-f;Q) = \overline{X(t,f;Q)} \quad . \tag{II.6}$$

This symmetry of the wavelet transform means that its computation for negative frequencies is not necessary in the real case, i.e. for most recordings, so that the symbol $\int \Theta(f'/f)$ can be replaced by $\int_0^\infty$ in the definition Eq.(II.2), and we can restrict frequencies to positive values.

1.2.3 Relations to the analytic signal and its spectrum

The spectrum of the signal can be recovered in two ways, either from a Fourier transform of the time or for a diverging quality factor:

$$\int X(t,f;Q) e^{-i\Upsilon f t} \mathrm{d}t = \hat{x}(f) \tag{II.7}$$

$$\lim_{Q\to\infty} X(t,f;Q) \frac{Q}{\sqrt{\Upsilon}} = \hat{x}(f) e^{i\Upsilon f t} |f| \quad . \tag{II.8}$$

The factor in the second line corresponds to the inverse of $C_{\psi_Q} = \frac{\sqrt{\Upsilon}}{Q}$.

Similarly, the analytic signal can be recovered in two ways, either from an integral over log-frequencies or a vanishing quality factor:

$$\int_0^\infty X(t,f;Q) \mathrm{d}\log f = x^+(t) \frac{\sqrt{\Upsilon}}{Q} \tag{II.9}$$

$$\lim_{Q\to 0} X(t,f;Q) = x^+(t) \quad . \tag{II.10}$$

The second line is valid only for $f > 0$, otherwise $f < 0$ gives the anti-analytic signal $x^-(t)$: this asymptotic limit only depends on the sign of f.

Unless we assume a centred signal, $\hat{x}(0) = 0$, note that these reconstruction formulas are ambiguous at $f = 0$. In practice, $f = 0$ is not part of a geometric frequency sampling and must be treated separately from the wavelet transform. This ambiguity is resolved when we also take into account negative frequencies:

$$x(t) = \int X(t,f;Q) \frac{\mathrm{d}f}{|f|} \frac{Q}{\sqrt{\Upsilon}} = x^-(t) + x^+(t) \quad , \tag{II.11}$$

where the anti-analytic signal $x^-(t)$, just $\overline{x^+(t)}$ in the real case, is relevant for general complex signals.

Considering the wavelet as a band-pass filter, the asymptotic formulas Eqs.(II.8, II.10) correspond to narrow-band and wide-band limits, detailed in Appendix 6.

1.2.4 Change of quality factor

We may need to change the quality factor after the computation of the wavelet transform. This operation changes of wavelet by integrating over both times and frequencies. For analytic wavelets, this is equivalent to reconstructing the signal and then compute the new wavelet transform.

However, there are single-integral short-cuts for log-normal wavelets, different whether we want to increase or decrease the quality factor Q (to $Q_- < Q < Q_+$):

$$X(t,f;Q_-) = \int X(t,f';Q)\overline{\hat{\psi}_{\epsilon_-}(f'/f)}\frac{\mathrm{d}f'}{|f'|}\sqrt{\frac{\epsilon_-^2+Q^2}{\curlyvee}}\ , \qquad \epsilon_-^{-2} = Q_-^{-2} - Q^{-2} \tag{II.12}$$

$$X(t,f;Q_+) = \int X(t',f;Q)\overline{\psi_{\epsilon_+}(f(t'-t))}f\mathrm{d}t'\ , \qquad \epsilon_+^2 = Q_+^2 - Q^2 \quad . \tag{II.13}$$

These are convolution (or cross-correlation) integrals with respect to log-frequency and time respectively, which are computed efficiently from the convolution theorem for each fixed second variable. In practice, this is precisely how the wavelet transform is obtained from the original signal, so that we cannot expect much numerical advantage in changing the quality factor in this way rather than from a new wavelet transform. These relations will bring conceptual simplifications, and they have a pivotal role at this end of this chapter.

1.2.5 Derivatives

The time derivative to any order n of the signal, $\partial_t^n x(t)$, has the spectrum $(i\curlyvee f)^n\hat{x}(f)$ and the wavelet transform:

$$\partial_t^n X(t,f;Q) = \left(i\curlyvee\lambda^{\frac{1}{2}}f\right)^n X(t,\lambda^n f;Q)\ , \quad \lambda = e^{\frac{1}{Q^2}} \quad , \tag{II.14}$$

where the order n can in fact be any real number. This property means that a time derivative not only changes the angle ($i^n = e^{in\frac{\pi}{2}}$) and the slope in scales, but due to the finite quality factor, it also causes a slight dilation in frequency, scaled up by the factor $\lambda > 1$. The frequency $\sqrt{\lambda}f$ is the intermediate one between the initial and final frequencies (f before and λf after a derivative). In practice, it also means that we need not recompute the wavelet transform of the signal time derivative when we already have it for the signal.

For $n = 1$, this property is both connected to the spectrum of the signal derivative, with the factor $i\curlyvee f$, and to the Cauchy-Riemann relation Eq.(I.11), which cannot be exact for the log-normal wavelet, hence the presence of λ. $X(t,\lambda f;Q)$ looks like the first term of a finite difference with respect to $\log f$, that would reproduce the derivative in the imaginary direction of the analytic function of the complex plane. There is no differentiation because the analytic wavelet transform is already a frequency-differentiated version of the analytical continuation (see chapter I, section 2.2).

For the completeness of the discussion, we could also ask what are the derivatives with respect to $\log f$ of $X(t,f;Q)$. They produce wavelet transforms of the same signal x with respect to new wavelets, very similar to the orthogonal family Eq.(I.49). It is indeed the case for the first order, up to a factor $\frac{Q^2}{2}$, but not for higher orders because the Hermite polynomials $H_k(-Q\log v)$ are obtained via $\log v$-derivatives of $\hat{\psi}_{\sqrt{2}Q}(v) = \hat{\psi}_Q(v)^2$ (instead of $\hat{\psi}_Q(v)$).

The time derivative is applied to study oscillatory signals in chapter III, while the log-frequency derivative would be more suited to characterize singular events across scales.

2 Quadratic representations: from energy to temporal and spectral densities

There exists a great diversity of processing techniques for natural recordings, most of which can be recast into operations on time-frequency representations. These representations can be linear, as the wavelet transform Eq.(II.2), or quadratic [Hlawatsch and Boudreaux-Bartels 1992] as introduced in this section. This section serves to introduce notations for quadratic wavelet-based expressions; we collect their properties and relations, emphasizing the role of the quality factor.

2.1 Energy of a pair of signals

We saw in the previous section how the signal and its spectrum are recovered from log-frequency and time integration of the wavelet transform respectively, Eqs.(II.9, II.7). More precisely, the spectrum is recovered from a time integral of the wavelet transform demodulated by its frequency: $X(t,f;Q)e^{-i\tau ft}$. Otherwise, the time integral of $X(t,f;Q)$ vanishes for any non-zero frequency f, since it oscillates circularly around zero. For this reason, wavelet-based integral estimators often involve a conjugated product such as the squared modulus $|X(t,f;Q)|^2$, which does not oscillate circularly, the simplest of which is the energy of the signal:

$$E_x = \int |x(t)|^2 \mathrm{d}t = \int |\hat{x}(f)|^2 \mathrm{d}f \tag{II.15}$$

$$= \iint |X(t,f;Q)|^2 \frac{\mathrm{d}f}{|f|}\mathrm{d}t\frac{Q}{\sqrt{\pi}} \quad , \tag{II.16}$$

where the first line is known as the Plancherel-Parseval equality. The correction factor corresponds to the inverse of the admissibility coefficient $C_{\psi_Q\psi_Q} = \frac{\sqrt{\pi}}{Q}$, Eq.(I.15), ubiquitous in quadratic estimators.

This quantity, which cumulates the intensity of the signal at all times and scales, extends to a pair of signals:

$$E_{xy} = \int x(t)\overline{y(t)}\mathrm{d}t = \int \hat{x}(f)\overline{\hat{y}(f)}\mathrm{d}f \tag{II.17}$$

$$= \iint X(t,f;Q)\overline{Y(t,f;Q)}\frac{\mathrm{d}f}{|f|}\mathrm{d}t\frac{Q}{\sqrt{\pi}} \quad , \tag{II.18}$$

that can be called the cross-energy of the signals, or their inner product. Note that the wavelet transform is already the inner product of the signal x with the wavelet ψ_Q dilated and shifted.

For a multidimensional signal $\vec{x}(t) = [x_i(t)]_{i=1,2,3,\ldots}$, we can define its energy as a full matrix $\mathbf{E}$ where each entry E_{ij} compares a pair of signals x_i, x_j as in Eq.(II.17). Diagonal entries E_{ii} correspond to Eq.(II.15), the energy of each signal x_i. Since $E_{yx} = \overline{E_{xy}}$, $\mathbf{E}$ is a Hermitian symmetric matrix (and positive definite).

Since the cross-energy E_{xy} serves to express both the multidimensional case $\mathbf{E}$ (each of its entries) and the simple energy case ($E_x = E_{xx}$), it is sufficient in the following, without loss of generality, to discuss only the comparison of a pair of signals x and y through related cross-quantities (where we will not specify the "cross-" each time).

2.2 Energy densities: power and power spectral density

In many cases, the intensity of the signals (their conjugated product) at infinite times or scales is still significant, so that the energy diverges. For instance, the pure wave $e^{i\tau ft}$ is periodic of intensity 1 at any time. Like its counterpart the pulse $\delta(t)$, of spectrum equal to 1 at any

frequency, they have an infinite energy. In practice, this energy is limited by the total duration T of the recording and the bandwidth of the measuring apparatus. When it does not have low frequency limitations, the slowest measurable oscillation has frequency of the order T^{-1}. A high frequency limit which constrains all digitalized measurements comes from the Nyquist sampling theorem, which states that in a recording sampled at the frequency f_s, the fastest resolved oscillations have a frequency at most equal to $\frac{f_s}{2}$.

The recorded energy is time-limited and band-limited in practice, corresponding to a bounded time-frequency domain of integration in Eq.(II.18). It is therefore useful to consider instead energy densities in time and frequency.

2.2.1 Power

In particular, the power refers to the average energy density per unit of time during the full recording, which reads in the limit of an infinite duration:

$$P_{xy} = \lim_{T\to\infty} \frac{1}{T} \int_{-\frac{T}{2}}^{+\frac{T}{2}} x(t)\overline{y(t)}\mathrm{d}t = \lim_{T\to\infty} \frac{1}{T} \int \hat{x}_T(f)\overline{\hat{y}_T(f)}\mathrm{d}f \quad , \tag{II.19}$$

where $\hat{x}_T(f)$ refers to the spectrum of the signal $x_T(t)$, the version of $x(t)$ restricted to the interval of duration T (zero outside).

For more clarity, let us rewrite the time average as $\langle\ \rangle_t$, so that the energy from a time and frequency integration in Eq.(II.18) turns into a power:

$$P_{xy} = \left\langle x(t)\overline{y(t)} \right\rangle_t = \int \left\langle X(t,f;Q)\overline{Y(t,f;Q)} \right\rangle_t \frac{\mathrm{d}f}{|f|} \frac{Q}{\sqrt{\pi}} \quad . \tag{II.20}$$

2.2.2 Time-dependent power or temporal energy density

The wavelet transform naturally yields a time-dependent power, or energy temporal density, when we omit the time average:

$$P_{xy}(t;Q) = \int X(t,f;Q)\overline{Y(t,f;Q)} \frac{\mathrm{d}f}{|f|} \frac{Q}{\sqrt{\pi}} \quad , \tag{II.21}$$

of mean value P_{xy} and integrating to E_{xy}. Note that this time-dependent power is free of any timescale, but still has a notion of time localization through the quality factor Q. It is most localized in the limit of a vanishing quality factor $Q \to 0$, that is called the instantaneous power:

$$P_{xy}(t) = x^+(t)\overline{y^+(t)} + x^-(t)\overline{y^-(t)} \quad . \tag{II.22}$$

The analytic and anti-analytic signals appear from the contribution of positive and negative frequencies, which are conjugated in the real case.

2.2.3 Power log-frequency and spectral densities

Similarly, we obtain a *power log-frequency density* when we omit the integral over log-frequencies in Eq.(II.20):

$$S_{xy}(f;Q)|f| = \left\langle X(t,f;Q)\overline{Y(t,f;Q)} \right\rangle_t \frac{Q}{\sqrt{\pi}} \quad , \tag{II.23}$$

where $S_{xy}(f;Q)$ has already the common interpretation of a *power spectral density* $S_{xy}(f) = S_{xy}(f;\infty)$ in the limit of an infinite quality factor $Q \to \infty$:

$$S_{xy}(f) = \int R_{xy}(\tau) e^{-i\Upsilon f \tau} \mathrm{d}\tau \ , \qquad R_{xy}(\tau) = \left\langle x(t+\tau)\overline{y(t)} \right\rangle_t = \int S_{xy}(\tau) e^{i\Upsilon f \tau} \mathrm{d}\tau \tag{II.24}$$

$$= \lim_{T\to\infty} \frac{\hat{x}_T(f)\overline{\hat{y}_T(f)}}{T} \quad . \tag{II.25}$$

where the last expression with duration restriction and limit, required in the frequency domain, can be considered as a practical shortcut to the computation of the delayed correlation $R_{xy}(\tau)$ in the time domain. Without the normalization by the duration T, we would respectively obtain the energy spectral density $\hat{x}(f)\overline{\hat{y}(f)}$ and the cross-correlation integral, suitable for recordings of time-localized energy.

Notice the necessity of the factor $|f|$ in Eq.(II.23): it recalls that a power log-frequency density integrates to the power with respect to the integrator $\frac{\mathrm{d}f}{|f|}$, i.e. with respect to $\mathrm{d}\log f$ for positive frequencies, contrary to a power spectral density, integrated with respect to $\mathrm{d}f$. In other words, the power is directly read from the area under the density curves, either $S_x(f)|f|$ on a log-frequency axis or $S_x(f)$ on a linear one. The logarithmic sampling is indeed natural for the wavelet transform because of its constant quality factor at all scales. To match the power spectral density, the factor $|f|$ is commonly absorbed in the definition Eq.(I.6) of the wavelet transform by choosing the other convention $p = 2$; this hides the log-frequency interpretation that we find important.

The power log-frequency density can in fact be computed directly from $S_{xy}(f)$ (i.e. from the product of Fourier spectra) as its log-frequency smoothing with a wavelet kernel:

$$S_{xy}(f;Q)|f| = \int S_{xy}(f')|\hat{\psi}_Q(f'/f)|^2 \mathrm{d}f' \frac{Q}{\sqrt{\pi}} \quad . \tag{II.26}$$

2.3 Wigner-Ville distribution: a glimpse out of the wavelet framework

We could try to go farther into the temporal and spectral localization of energy or power densities. If we omit both the time and log-frequency integrations, the bare intensity product $X(t,f;Q)\overline{Y(t,f;Q)}\frac{Q}{\sqrt{\pi}}$ remains; sometimes called the cross-wavelet transform or (cross-) scalogram. But we cannot take both the instantaneous limit $Q \to 0$, as in Eq.(II.22), and the converse one $Q \to \infty$, as in Eq.(II.24), without getting rid of the wavelet ψ_Q. Nevertheless, this seems possible from Eq.(II.24) by omitting the time average, which is called (when centring the time with respect to the delay) the Wigner-Ville distribution [Ville 1948; Wigner 1932]:

$$W_{xy}(t,f) = \int x(t+\tfrac{\tau}{2})\overline{y(t-\tfrac{\tau}{2})} e^{-i\Upsilon f \tau} \mathrm{d}\tau = \int \hat{x}(f+\tfrac{\eta}{2})\overline{\hat{y}(f-\tfrac{\eta}{2})} e^{i\Upsilon \eta t} \mathrm{d}\eta \quad . \tag{II.27}$$

Notice the strangeness of the second equality from the wavelet's scaling point of view over frequencies.

This quadratic time-frequency representation is the first attempt to define an instantaneous power spectral density for a signal $x = y$, that has found many applications in signal processing and in quantum mechanics (as a phase-space distribution). This transform has many mathematical properties, like the possibility to achieve a perfect time-frequency localization of oscillating modes [Flandrin 1998b], or the recovery, as marginal distributions, of the power spectral density Eq.(II.24) and the product $x(t)\overline{y(t)}$, notably different from the instantaneous power Eq.(II.21) (due to a mixing of negative and positive frequencies). However, it also lacks some practical ones, like the positivity of $W_{xx}(t,f)$, only guaranteed for Gaussian signals [Flandrin 1998a], and the presence of interferences (cross-terms) between components at different

time-frequency localizations, that complicate its interpretation. These difficulties have motivated their regularization by various filtering techniques, and may be interpreted as different manifestations of the uncertainty principle.

One of them yields the product of wavelet transforms (see proof in Appendix 2):

$$X(t,f;Q)\overline{Y(t,f;Q)} = \iint W_{xy}(t',f')\overline{W_{\psi_Q\psi_Q}(f(t'-t),f'/f)}|f|\mathrm{d}t'\mathrm{d}f' \quad , \tag{II.28}$$

the filtering kernel being the Wigner-Ville distribution of the wavelet ψ_Q, introducing an explicit extent for the time-frequency uncertainty. In fact, the Wigner-Ville is like a bare frame for quadratic time-frequency expressions, that is "dressed" by such kernel, conferring it more regularity and particularity. Another such "dressing" produces the Gabor transforms product, that treats time and frequency alike, as shown in Appendix 2.

In conclusion, $X(t,f;Q)\overline{Y(t,f;Q)}\frac{Q}{\sqrt{\pi}}$ is the most localized quadratic expression (intensity) within the wavelet approach. The Wigner-Ville distribution is an external viewpoint, suitable for a comparison with the Gabor approach. It underlies quadratic time-frequency representations in both case, "dressed" (convolved) with a kernel that constitutes an explicit atom of time-frequency uncertainty, $W_{\psi_Q\psi_Q}(u,v)$ for the wavelet approach.

Summary of sections 1 and 2

The time-frequency representation of a recording x, based on the wavelet transform with quality factor Q, is given a simplified notation $X(t,f;Q)$, illustrated and commented for a model signal and a natural recording. The associated time-frequency plane has a hyperbolic geometry, conditioning from the behaviour of phase vortices to its efficient numerical implementation. The temporal and spectral intuitions associated the recording signal $x(t)$ or spectrum $\hat{x}(f)$ are transferred to the time and frequency variables of this linear representation. Both the signal and the spectrum can be retrieved by integrating over frequency or time respectively, and the spectral (respectively the temporal) interpretation can be favoured by choosing a high (respectively low) value of the quality factor: $Q \to 0, \infty$.

This time-frequency intuition extends to quadratic quantities, involving two possibly different recordings, x, y. The integration of the intensity $X(t,f;Q)\overline{Y(t,f;Q)}$ over the time and/or the frequency variable(s) leads to the notions of energy of the signals, its temporal density and its log-frequency density. Again, the low and high quality factor limits step in to identify the energy temporal density to the power $P_{xy}(t)$, while the log-frequency density is related to the power spectral density $S_{xy}(f)$ by a factor $|f|$.

These limits of maximal time or frequency localization cannot be achieved simultaneously, without breaching the wavelet framework: the Wigner-Ville distribution appears as the "skeleton" of quadratic representation, "dressed" into the wavelet intensity product from a wavelet-kernel convolution in time and frequency.

3 Quadratic wavelet-estimators for stochastic processes

Irregular fluctuations in data are classically approached in a probabilistic way, so that continuous recorded signals are often modelled as a combination of deterministic and stochastic processes. Therefore, statistical aspects of the recorded fluctuations are important, such as expectations, and ideally their probability distribution. In the context of empirical measurements (in finite number), the main difficulty of the probabilistic approach consists in defining a suitable statistical ensemble, especially when the properties of the underlying physical system are

yet unknown. A collection of realizations of the same stochastic process is needed to estimate empirically its properties. Fixing a statistical ensemble thus consists in specifying, from physical assumptions, which parts of the recorded data will be considered as distinct realizations of the process.

This question is crucial in two aspects: first in order to reduce the complexity of the data, or to the contrary, to explore it at its full extent, and second to control the significance of the observations we can make of it. This polarity is a balance: the finite amount of information in the data is shared between depicting more (or less) of its complexity, and acquiring less (or more) certainty over its statistical properties.

In the absence of prior information on the recorded signals, we follow the development of the previous section in a bottom-up approach: we start with the strongest and least realistic assumptions, reducing radically the complexity, to the weakest and most parsimonious ones allowing to uncover it. We introduce below the minimum required material and notations to interpret the previous section in a statistical manner.

3.1 Assumptions of stationarity and ergodicity: expectation and distribution

Let the signal $x(t)$ be the recorded activity of the system under measurement, modelled as a stochastic process. Consider that the statistical variability of the signal has no dependence on the time parameter t, which measures the recording time relative to an arbitrary initial time: thus, we assume its *stationarity*. Its ensemble average $\mathbb{E}$ (or expectation) and its probability density function $p(x)$, related for any function $h(x)$ of the signal by

$$\mathbb{E}\left[h(x)\right] = \int h(x)p(x)\mathrm{d}x \quad , \tag{II.29}$$

may be estimated from the average over the recorded duration introduced in Eqs.(II.19, II.20):

$$\left\langle h(x(t))\right\rangle_t = \int h(x)\tilde{p}(x)\mathrm{d}x \ , \quad \tilde{p}(x) = \left\langle \delta(x - x(t))\right\rangle_t \quad . \tag{II.30}$$

The empirical probability distribution $\tilde{p}(x)$ is a normalized histogram, where the binning function is taken infinitely thin (a Dirac delta) in the limit of an infinite duration. A process $x(t)$ for which the following equality holds for any h:

$$\mathbb{E}\left[h(x)\right] = \left\langle h(x(t))\right\rangle_t \quad , \tag{II.31}$$

is said to be *ergodic*. Equivalently, the normalized histogram equals to the probability density function, $\tilde{p}(x) = p(x)$. Ergodicity is an abstract and strong additional hypothesis, which can be weakened by assuming it only for the first moments: for instance for the mean $\mathbb{E}\left[x\right] = \left\langle x(t)\right\rangle_t$.

In practice, the duration of the recording is finite as well as the sampling frequency, so that, even when the assumptions are correct, an equality such as Eq.(II.31) turns into a statistical estimation: $\mathbb{E}\left[h(x)\right] \approx \left\langle h(x(t))\right\rangle_t$ for a duration $T < \infty$.

A certain aspect of stationarity, namely the possibility of correlation in the process between different times, is not visible is this simplified introduction, and is approached from the second order statistics in the following.

3.2 Statistical variability

Now, we discuss the assumption of stationarity for a stochastic process, extending it to correlations between different times, and we show how the quadratic representations of the signal, exposed in the previous section 2, directly constitute estimators of the second order statistics

for the stochastic process. Following section 2, a pair of signals $x(t)$ and $y(t)$ is sufficient to treat quadratic estimator in the case of multi-dimensional recordings; for the sake of simplicity, we keep once more the one-dimensional terminology.

Eventually, we also assume that the signal has zero mean, i.e. we consider the centred process $x(t) - \mathbb{E}[x(t)]$. In practice, this is equivalent to ignoring its global trend, or its constant component (vanishing frequency) for a stationary process.

3.2.1 Variance

The variance between the pair $x(t)$ and $y(t)$ of zero-mean stochastic processes (then a kind of covariance) is defined as:

$$\sigma^2_{xy}(t) = \mathbb{E}[x(t)\overline{y(t)}] \quad , \tag{II.32}$$

which depends on time in the general non-stationarity case. The conjugation is important for complex signals, which additionally have a distinct pseudo-variance $\mathbb{E}[x(t)y(t)]$.

The (jointly) stationary case is independent on the time variable, so that the power P_{xy} in Eq.(II.19) is a natural estimator of the stationary variance σ^2_{xy}. By extension, $P_{xy}(t;Q)$ defined in Eq.(II.21) appears as a natural wavelet-estimator for the non-stationary variance $\sigma^2_{xy}(t)$, parametrized by the quality factor and distinct from an averaging over a short interval of time.

Let us stick to the assumption of stationarity for the moment and examine the more general covariance, here referring to the signals compared at different times:

$$\sigma^2_{xy}(t',t) = \mathbb{E}[x(t')\overline{y(t)}] \quad . \tag{II.33}$$

For this quantity to be stationary, it must not depend on a relative time variable, but it can depend on an absolute time such as the delay $\tau = t' - t$. Therefore, a stationary covariance satisfies at any time t:

$$\sigma^2_{xy}(t+\tau,t) = \sigma^2_{xy}(\tau,0) \quad , \tag{II.34}$$

estimated by the delayed correlation $R_{xy}(\tau)$ in Eq.(II.24).

3.2.2 Power densities again

Under the stationary and ergodic hypothesis, the covariance and the power spectral density are related by a Fourier transform:

$$S_{xy}(f) = \int \sigma^2_{xy}(\tau,0)e^{i\Upsilon f\tau}\mathrm{d}\tau \quad . \tag{II.35}$$

This result is known as the Wiener-Khinchin theorem, that establishes Eq.(II.24) in a rigorous stochastic setting. In particular, the variance is the integral of the power spectral density: $\sigma^2_{xy} = \int S_{xy}(f)\mathrm{d}f$. Since $\hat{x}(f)\overline{\hat{y}(f)}$ is stochastic, the definition of the power spectral density directly from the spectrum of a single trajectory of the process, as in Eq.(II.25), must be adapted with an ensemble average:

$$S_{xy}(f) = \lim_{T\to\infty} \frac{1}{T}\,\mathbb{E}\left[\hat{x}_T(f)\overline{\hat{y}_T(f)}\right] \quad . \tag{II.36}$$

The wavelet-estimator $S_{xy}(f;Q)$ Eq.(II.23) comes here into play: it averages over the time variable of the wavelet transform, at the cost of introducing a certain frequency uncertainty, explicit when expressed as a log-frequency smoothing in Eq.(II.26). In the wavelet framework,

it constitutes a natural and scale-free alternative to the common estimations based on sliding windows of fixed duration, such as Welch's method. These yield a different spectral smoothing with a constant linear frequency resolution (as in the Gabor approach), more resolved but more variable for larger windows. In the wavelet approach the higher the quality factor, the better the log-frequency resolution, but the higher the variability of the estimator. Here is the main difference in the results: contrary to the wavelet approach, low frequencies are sacrificed in the Welch-Gabor approach because of the windowing of the signal, but high frequencies are better resolved than in the wavelet case at fixed total duration T. This is an advantage only if fast components are very periodic with no slow component.

3.2.3 Non-stationary estimations, a first approach

Now how should we define and estimate a time-dependent power density? The room in Eq.(II.35) for a non-stationary covariance Eq.(II.33) primarily leads to the following definition for the time-dependent power spectral density:

$$S_{xy}(t,f) = \int \sigma^2_{xy}(t+\tfrac{\tau}{2}, t-\tfrac{\tau}{2}) e^{i\tau f \tau} \mathrm{d}\tau = \mathbb{E}[W_{xy}(t,f)] \quad , \tag{II.37}$$

which is the statistical expectation of the Wigner-Ville distribution of the process introduced in Eq.(II.27). Among the difficulties associated to using the bare Wigner-Ville distribution as an estimator, there is the possibility of negative values for $S_{xx}(t,f)$ [Flandrin 1986] as well as its infinite variance [Stanković and Stanković 1993]. Therefore, we turn towards its wavelet filtering, and the corresponding time-dependent power log-frequency density:

$$S_{xy}(t,f;Q)|f| = \mathbb{E}\left[X(t,f;Q)\overline{Y(t,f;Q)}\right] \frac{Q}{\sqrt{\pi}} \quad , \tag{II.38}$$

where no limit of the quality factor can lead to Eq.(II.37), as explained in chapter II, section 2.3.

The non-averaged but smooth quantity $X(t,f;Q)\overline{Y(t,f;Q)}\frac{Q}{\sqrt{\pi}}$ could be interpreted as a naive estimator of $S_{xy}(t,f;Q)|f|$. However, it cannot be considered as a good estimator since it is computed from a single realization of the process. We can anticipate that the single atom of uncertainty covered by the wavelet is just enough to mitigate the lack of regularity of Eq.(II.37), but still insufficient to reduce variability as a statistical estimator. This puzzle is left for section 5, where it becomes crucial.

Instead, we discuss the not-less-puzzling issue of estimating the general covariance $\sigma^2_{xy}(t',t)$. Eq.(II.37) suggests using the inverse Fourier transform of with respect to f: $\int S_{xy}(t,f;Q)e^{i\tau f\tau}\mathrm{d}f$. We argue that the following estimator is more natural in a time-frequency perspective, that does not require the Fourier transform any more, to jump from frequency to time delay:

$$R_{xy}(t',t;Q) = \int X(t',f;Q)\overline{Y(t,f;Q)}\frac{\mathrm{d}f}{|f|}\frac{Q}{\sqrt{\pi}} \quad . \tag{II.39}$$

Although non-equivalent, both alternatives agree for a vanishing delay, that yields the power $P_{xy}(t;Q) = R_{xy}(t,t;Q)$, estimator of the variance. Here is the main reason for this choice: while they have the same asymptotics at large quality factor $Q \to \infty$ (which is stationary), the instantaneous limit is ambiguous in the alternative Fourier definition but well-defined in Eq.(II.39):

$$R_{xy}(t',t;Q\to 0) = x^{+}(t')\overline{y^{+}(t)} + x^{-}(t')\overline{y^{-}(t)} \quad , \tag{II.40}$$

generalizing the instantaneous power Eq.(II.22). These observations are detailed in Appendix 7.

In the context of the ergodic hypothesis, the above integral across scales could be understood as assuming scale-invariance rather than stationarity. A logarithmic or exponential change of variable (time and log-frequency) is indeed the only formal difference between these assumptions; this correspondence is called the Lamperti transform [Borgnat, Flandrin, and Amblard 2002; Flandrin, Borgnat, and Amblard 2003].

3.3 Example of self-similar Gaussian processes: noises and their colour

Consider a 1-dimensional, real and zero-mean stochastic process $x(t)$. Assume that this process has a Gaussian distribution, then it is unambiguously defined by specifying of the covariance, $\sigma^2_{xx}(t',t)$. In the stationary case, the power spectral density $S_{xx}(f)$ is equivalently sufficient to characterize the process from Eq.(II.35).

3.3.1 Fractional Brownian motion

The fractional Brownian motion is an example of Gaussian process, useful to model self-similar noises $x(t) = \xi_H(t)$, so that $|a|^{-H}\xi_H(at)$ has the same distribution for any scaling a, with a regularity parameter H called the Hurst exponent in this stochastic context. This process is characterized by the covariance [Meyer, Sellan, and Taqqu 1999]:

$$\sigma^2_{xx}(t',t) = \frac{\sigma^2}{2H}\left(|t'|^{2H} + |t|^{2H} - |t'-t|^{2H}\right) \quad , \tag{II.41}$$

for any exponent $0 < H < 1$; notice our choice to normalize by H. The Brownian motion or Wiener process is the intermediary case $H = \frac{1}{2}$, which can be rewritten $\sigma^2_{xx}(t',t) = \sigma^2 \min(t',t)$ for t', $t' > 0$, where the usual diffusion coefficient in physics corresponds to $D = \frac{\sigma^2}{2}$. Because of the two first terms, this process is not stationary, so that we cannot apparently define the power spectral density. But the wavelet transform happens to be stationary for any $f \neq 0$:

$$\begin{aligned}
S_{xx}(t,f;Q)|f| &= \iint \mathbb{E}[\xi_H(t_1)\xi_H(t_2)]\overline{\psi_Q(f(t_1-t))}\psi_Q(f(t_2-t))f^2\mathrm{d}t_1\mathrm{d}t_2 \\
&= \tfrac{\sigma^2}{2H}\Re\left\{\int |t'|^{2H}\overline{\psi_Q(f(t'-t))}f\mathrm{d}t' \int \psi_Q(u)\mathrm{d}u\right\} \\
&\qquad - \tfrac{\sigma^2}{2H}\iint |t'-t|^{2H}\overline{\psi_Q(f(t_1-t))}\psi_Q(f(t_2-t))f^2\mathrm{d}t_1\mathrm{d}t_2 \\
&= -\frac{\sigma^2}{2H}|f|^{-2H}\iint |u_1-u_2|^{2H}\overline{\psi_Q(u_1)}\psi_Q(u_2)\mathrm{d}u_1\mathrm{d}u_2 \\
&= -\frac{\sigma^2}{2H}|f|^{-2H}\iint |u|^{2H}|\hat{\psi}_Q(v)|^2 e^{i\Upsilon uv}\mathrm{d}u\mathrm{d}v \quad ,
\end{aligned} \tag{II.42}$$

since the non-stationary terms vanishes ($\int \psi_Q(u)\mathrm{d}u = \hat{\psi}_Q(0) = 0$).

Therefore, we can compute the following wavelet estimate for the power log-frequency density at $f \neq 0$:

$$S_{xx}(f;Q)|f| = \frac{\sigma^2 \sin(H\pi)\Gamma(1+2H)}{H\Upsilon^{1+2H}}|f|^{-2H}e^{(H/Q)^2} \quad , \tag{II.43}$$

where the factor $e^{(H/Q)^2}$ corresponds to the bias of the estimator and to a frequency dilation $\lambda^{-\frac{H}{2}}$. It goes to 1 in the limit of an infinite quality factor $Q \to \infty$, providing an expression for the power density $S_{xx}(f)|f|$. The fact that we can define a power density for any non-zero frequency can be understood as a sort of stationarity of the process, except for a "random trend" (in the neighbourhood of $f = 0$). This trend is removed when we consider instead the increment $\xi_H(t+\tau) - \xi_H(t)$, which is really stationary.

Coloured noises with any Hurst exponent can be though of as successive integrations ($H \to H+1$) or differentiations ($H \to H-1$) of the fractional Brownian motion of Hurst exponent H. In particular, the white noise $H = -\frac{1}{2}$, with correlation function

$$\sigma_{xx}(t,t') = \sigma^2\delta(t-t') \quad , \tag{II.44}$$

is the "derivative" of the Brownian motion with $H = \frac{1}{2}$. Although a general definition is delicate in the time domain, any real value of H can be conceived in the log-frequency domain from the power density Eq.(II.43) once $f = 0$ is excluded.

3.3.2 Log-correlated noise: pink is the new white

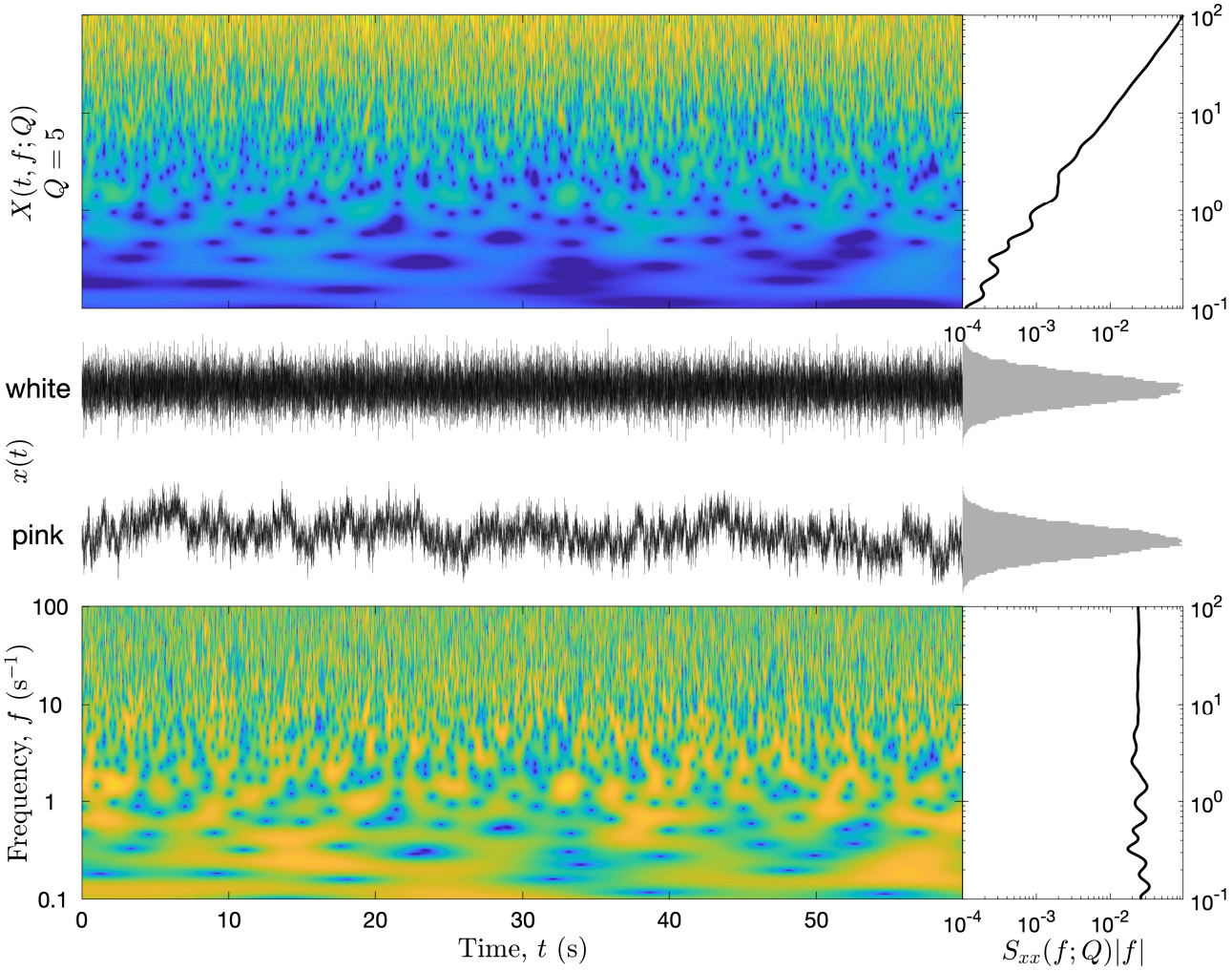


Figure II.3: *Comparison of the white (top) and pink (bottom) Gaussian noises. Signals and their histogram (centre), wavelet transform's amplitude (left) and estimated power density (right), with $Q = 5$. From the white noise spectrum $\hat{x}(f)$, the pink noise is obtained as $\hat{x}(f)|f|^{-\frac{1}{2}}$. Their power log-frequency density scales as $S_{xx}(f;Q)|f| \sim |f|^{-2H}$, with Hurst exponents $H = -\frac{1}{2}$ and 0 respectively.*

In Fig.II.3, the white noise is compared to the case $H = 0$, called the *pink* noise $x(t) = \xi_0(t)$, which is special to the wavelet approach [Abry, Gonçalvès, and Flandrin 1995], since it has a constant power log-frequency density:

$$\lim_{H\to 0} S_{xx}(f;Q)|f| = \frac{\sigma^2}{2} \quad , \tag{II.45}$$

well-defined for any $f \neq 0$ in the limit $H \to 0$, thanks to the normalization by H in Eq.(II.41). It also holds for any quality factor: the estimator is unbiased in this case. Therefore, the power per octave is constant and equal to $\frac{\sigma^2}{2}\log 2$, where $\log 2$ is the "per octave" unit (change of base for the logarithm from e to 2).

Said otherwise, this models the "1/f noise", named after the behaviour of its power spectral density $S_{xx}(f;Q) = D|f|^{-1}$. The divergence for $f \to 0$ in this linear view over the frequency domain explains the difficulties in the time domain. For instance, the covariance Eq.(II.41) is undefined in the limit $H \to 0$. However, there is only one possible finite term, up to an additive

constant, which is stationary (thus remaining in Eq.(II.42)):

$$\lim_{H\to 0} \mathbb{E}[|X(t,f;Q)|^2] = \sigma^2 \iint \log \frac{T}{|t_1-t_2|} \overline{\psi_Q(ft_1)} \psi_Q(ft_2) f^2 \mathrm{d}t_1 \mathrm{d}t_2 \quad , \tag{II.46}$$

which can be obtained from the limit: $\frac{|\tau|^{2H}-1}{2H} \to \log|\tau|$, and T is an arbitrary timescale (allowed because the wavelet has zero mean). For this reason, this noise is also called the log-correlated Gaussian process, presented as a tempered distribution in [Duplantier et al. 2017], where the zero-mean test function is the role played by the wavelet ψ_Q. This ideally self-similar and log-correlated process is crucially involved in the description of continuous multiplicative random cascades [Arneodo, Manneville, and Muzy 1998; Muzy, Baïle, and Bacry 2013]. In these models, the diverging term or equivalently the reference timescale T in the covariance are clearly interpreted as the largest timescale of the system: the duration of the signal.

Although this is not required in the wavelet formalism, this fact about the covariance can be straightforwardly modelled from Eq.(II.45), stopping the integral bound $f \to 0$ in its computation at the smallest practical frequency T^{-1}:

$$\begin{aligned} \frac{\sigma^2_{xx}(\tau,0)}{\sigma^2} &\approx \frac{1}{2}\left(\int_{-\infty}^{-T^{-1}} e^{i\Upsilon f\tau}\frac{\mathrm{d}f}{|f|} + \int_{T^{-1}}^{+\infty} e^{i\Upsilon f\tau}\frac{\mathrm{d}f}{|f|}\right) \\ &= \int_{T^{-1}}^{+\infty} \cos(\Upsilon f\tau)\mathrm{d}\log f = -\gamma + \log\frac{T}{\Upsilon\tau} + \mathrm{Cin}\left(\Upsilon\frac{\tau}{T}\right) \quad . \end{aligned} \tag{II.47}$$

Here, $\gamma \approx 0.5772$ is the Euler-Mascheroni constant, that often appears when describing a logarithmic divergence, and the cosine integral Cin is a special function that gives oscillating polynomial corrections at a finite large scale T, see Eq.(A.59) in the Appendix. Since $\mathrm{Cin}(0) = 0$, it can be neglected. Note that this derivation is stationary, because the only non-stationarity, inherited from the definition Eq.(II.41), pertains to the limit $f \to 0$. Finite frequencies are as far from this vanishing limit than to the diverging one $f \to \infty$, as ensured by the fast decay of the log-normal wavelet on both sides. Non-stationary (causal) versions of the pink noise can be found in [Arneodo, Audit, et al. 1998; Muzy, Baïle, and Bacry 2013].

The pink (log-correlated) Gaussian noise is to log-frequency just what the white Gaussian noise is to linear frequency: a constant power density over the domain. But its description from the time domain (i.e. from the covariance) is cluttered by formal discrepancies between the additive (time) and multiplicative (log-frequency) approaches, sometimes conferring an aura of mystery to the "1/f noise". We prefer to call it the pink noise, whose simplicity is clearer from the wavelet multiplicative viewpoint Eq.(II.45), turning the white noise into an "f noise" as we can see in Fig.II.3.

4 Correlation and Coherence

Elaborating upon the previously introduced covariance and the power densities, we show how their normalized versions, the correlation and the coherence coefficients, characterize the similitude between a pair of signals $x(t)$ and $y(t)$, independently of the intensity of their fluctuations. We still model these signals as complex stochastic processes with zero mean.

4.1 Correlation coefficient

The covariance Eq.(II.33) normalized by the geometric average of the variances provides a robust way to compare the fluctuations in a pair of signals, potentially of very different nature.

Quantifying the strength of the linear relationship between x and y, it is called the (Pearson) correlation coefficient, whose general definition (possibly delayed and time-dependent) is:

$$\rho_{xy}(t',t) = \frac{\sigma^2_{xy}(t',t)}{\sigma_{xx}(t',t')\sigma_{yy}(t,t)} \quad , \tag{II.48}$$

equivalent to the covariance of signals standardized to a unit variance (e.g. $\frac{x}{\sigma_{xx}}$). As previously, the stationary case can be written $\rho_{xy}(\tau,0)$, where $\tau = t' - t$ is a delay between the signals. The delay is set to zero in its simplest form.

4.1.1 Stationary case and interpretation

The natural ergodic estimator for $\rho_{xy}(\tau,0)$ is

$$\Gamma_{xy}(\tau) = \frac{R_{xy}(\tau)}{\sqrt{P_{xx}P_{yy}}} \quad , \tag{II.49}$$

assuming joint stationarity.

For real signals, this coefficient takes values in the interval $[-1,1]$. A value of 1 (respectively -1) indicates a full (anti-)correlation, while a zero value indicates the absence of (linear) correlation. For complex signals, the correlation coefficient takes values in the unit disk: the strength of the linear relationship is given by the modulus $|\Gamma_{xy}| \in [0,1]$, whereas the sign of the correlation is extended to a continuum of angular phase (the complex argument of Γ_{xy}). Positive correlation correspond to the phase 0, and anti-correlation to the phase $\pm\pi$.

By definition, a perfect correlation is reached for $x = y$ and $\tau = 0$: $\Gamma_{xx}(0) = 1$. However, a non-trivial correlation $\Gamma_{x\overline{x}}(0)$ is possible between a general complex signal and its conjugate.

In the real case, we argue that the correlation coefficient of the analytic signals $\rho_{x^+y^+}(\tau)$, which is complex, is much more instructive than its real counterpart, although they contain the same information. They are related as follows:

$$\rho_{xy}(\tau) = \Re\left\{\rho_{x^+y^+}(\tau)\right\} \quad , \tag{II.50}$$

and the same goes for the complex estimator $\Gamma_{x^+y^+}(\tau)$. This comes from the Hermitian symmetry of the power spectral density $S_{xy}(-f) = \overline{S_{xy}(f)}$, that relates to the correlation coefficient as:

$$\rho_{xy}(\tau,0) = \frac{\int S_{xy}(f)e^{i\Upsilon f\tau}\mathrm{d}f}{\sqrt{\int S_{xx}(f)\mathrm{d}f \int S_{yy}(f)\mathrm{d}f}} \ , \quad \rho_{x^+y^+}(\tau,0) = \frac{\int_0^\infty S_{xy}(f)e^{i\Upsilon f\tau}\mathrm{d}f}{\sqrt{\int_0^\infty S_{xx}(f)\mathrm{d}f \int_0^\infty S_{yy}(f)\mathrm{d}f}} \quad . \tag{II.51}$$

Slightly anticipating the next sections, we may call these complex quantity in the case $\tau = 0$ the *global coherence* $\rho_{x^+y^+} = \rho_{x^+y^+}(0,0)$, estimated as $\Gamma_{x^+y^+}(0)$.

While the real interpretation remains straightforwardly available, the angular information is useful to characterize phase quadrature ($\pm\frac{\pi}{2}$) phenomena such as the one between a component and its derivative: $x(t) = \frac{\mathrm{d}}{\mathrm{d}t}y(t) \ \Rightarrow \ S_{xy}(f) = i\Upsilon f S_{yy}(f) \ \Rightarrow \ \rho_{x^+y^+}(0,0) = i = e^{i\frac{\pi}{2}}$. This extends the idea of full correlation to the extent of the modulus, maximal in this example: $|\rho_{x^+y^+}| = 1$, while the angle indicates whether there is a positive correlation, an anti-correlation (phase opposition $\pm\pi$) a phase quadrature ($\pm\frac{\pi}{2}$) as in this example, or any intermediate case.

4.1.2 Estimation in the non-stationary case

The fully general correlation coefficient $\rho_{xy}(t',t)$ defined in Eq.(II.48) is estimated by:

$$\Gamma_{xy}(t',t;Q) = \frac{R_{xy}(t',t;Q)}{\sqrt{P_{xx}(t';Q)P_{yy}(t;Q)}} \quad , \tag{II.52}$$

where the non-stationary covariance estimator $R_{xy}(t',t;Q)$ has been defined in Eq.(II.39) from a log-frequency integral. See Appendix 7 for a discussion about its definition and properties.

In the real non-stationary case, the property Eq.(II.50) is also satisfied, and applies to the estimator:

$$\Gamma_{xy}(t',t;Q) = \Re\left\{\Gamma_{x^+y^+}(t',t;Q)\right\} \quad , \tag{II.53}$$

thanks to the Hermitian symmetry of the wavelet transform. In the case $t' = t$, we call the complex version $\Gamma_{x^+y^+}(t,t;Q)$ the estimator for the *temporal coherence* $\rho_{x^+y^+}(t,t)$, proposed in [Gurley, Kijewski, and Kareem 2003] as a "scale-averaged" coherence.

The local stationarity is assumed relatively to each scale: in contrast to a supposed characteristic duration of steadiness, we posit at each distinct scale a steady statistical behaviour for a certain number $f\delta t = \frac{Q}{\sqrt{2\uparrow}}$ of oscillations, see Eq.(I.41). The higher the quality factor, the more we assume stationarity. Assuming no stationarity at all corresponds to the instantaneous limit, for which the estimator of temporal coherence does not reduce the variability any more in this single trial estimation. In particular, we always obtain:

$$|\Gamma_{x^+y^+}(t',t;Q\to 0)| = 1 \tag{II.54}$$

in this limit, even though $|\rho_{x^+y^+}(t',t)| \neq 1$ in general.

Overall, Eq.(II.53) constitutes a useful tool to scrutinize linear relationship in complex fluctuating signals in a scale-free manner. Without fixing any arbitrary timescale in the way, the quality factor Q advantageously controls the balance between the amounts of remaining variability and assumed stationarity. Pursuing the distinction of positive and negative frequency domains with analytic and anti-analytic signals $x^{\pm}(t)$, the contributions to the correlation at different scales remain to be distinguished.

4.2 Coherence: distinguishing correlations at different scales

We introduce here the concept of coherence, well-documented [Gardner 1992] and with many signal processing applications [Halliday et al. 1995]. In particular, it has been used as a way to estimate delays [Carter 1987]. The reason for this is its close relation to the Fourier spectrum of the stationary correlation coefficient:

$$\int \rho_{xy}(\tau,0)e^{-i\uparrow f\tau}\mathrm{d}\tau = \frac{S_{xy}(f)}{\sqrt{\int S_{xx}(f)\mathrm{d}f \int S_{yy}(f)\mathrm{d}f}} \quad , \tag{II.55}$$

whose integral across frequencies is the usual (Pearson) correlation coefficient, see Eq.(II.51). The contributions to correlation can thus be detailed at each scale, and their normalization defines a continuum which takes the same values and has the same interpretation as a complex correlation coefficient at each frequency.

4.2.1 Spectral coherence: the stationary case

This frequency-wise normalization leads to the *spectral coherence*:

$$\gamma_{xy}(f) = \frac{S_{xy}(f)}{\sqrt{S_{xx}(f)S_{yy}(f)}} \quad , \tag{II.56}$$

between two jointly stationary zero-mean random processes, a natural extension of the correlation analysis for a pair of multi-scale or multi-component signals. It takes again complex values in the unit disk, and its squared modulus $|\gamma_{xy}(f)|^2 \leq 1$ (original definition of the coherence)

represents the proportion of the power density at frequency f of one signal that can be linearly predicted from the other signal.

The presence of a delay or a global phase (as for anti-correlation or quadrature) is contained in the angle $\phi_{xy}(f)$ of the spectral coherence: $\gamma_{xy}(f) = |\gamma_{xy}(f)|e^{i\phi_{xy}(f)}$. Indeed, consider that the signals are shifted copy of each other, $x(t) = y(t-\tau)$, then their coherence is $\gamma_{xy}(f) = e^{-i\Upsilon f\tau}$ so that $\phi(f) = \Upsilon f\tau$: x is delayed by τ compared to y (i.e. y is early of τ as compared to x). Such interpretation of $\phi(f)$ is as exact as the coherence modulus $|\gamma_{xy}(f)|$ is close to 1. See reference [Carter 1987] for an extended statistical analysis.

The properties of the coherence Eq.(II.56) crucially depend on the ensemble average, implicit in Eq.(II.35) or explicit in Eq.(II.36). Indeed, for a single realization of the processes, the wavelet-estimator is obtained from Eq.(II.23):

$$\gamma_{xy}(f;Q) = \frac{S_{xy}(f;Q)}{\sqrt{S_{xx}(f;Q)S_{yy}(f;Q)}} = \frac{\langle X(t,f;Q)\overline{Y(t,f;Q)}\rangle_t}{\sqrt{\langle|X(t,f;Q)|^2\rangle_t\langle|Y(t,f;Q)|^2\rangle_t}} \quad . \tag{II.57}$$

As in Eq.(II.54) for the correlation coefficient, the most resolved version is the worst estimator:

$$|\gamma_{xy}(f;Q\to\infty)| = 1 \tag{II.58}$$

since the modulus is always 1, only the angle varies. But the converse limit is not any better as the spectral distinction is lost and we obtain the global coherence:

$$\gamma_{xy}(f;Q\to 0) = \Gamma_{x^+y^+}(0) \quad , \tag{II.59}$$

for $f > 0$, or $\Gamma_{x^-y^-}(0)$ for $f < 0$.

4.2.2 Non-stationary case: outlook on the problem

The time-frequency coherence quantifies the strength of correlation between non-stationary signals at different locations of the time-frequency plane. Its generic definition depends on time-dependent power densities:

$$\gamma_{xy}(t,f) = \frac{S_{xy}(t,f)}{\sqrt{S_{xx}(t,f)S_{yy}(t,f)}} \quad . \tag{II.60}$$

The use of Wigner-Ville spectra Eq.(II.37) is in general excluded unless carefully filtered [Matz and Hlawatsch 2000; Orini, Bailon, Mainardi, Minchole, et al. 2009; Orini, Bailon, Mainardi, Laguna, et al. 2012; White and Boashash 1990]. Indeed, its interference terms must be managed to keep a coherence value in the unit disk. Such filtering is achieved by the wavelet approach, i.e. by using the power density $S_{xy}(t,f;Q)|f|$ defined in Eq.(II.38):

$$\gamma_{xy}(t,f;Q) = \frac{\mathbb{E}[X(t,f;Q)\overline{Y(t,f;Q)}]}{\sqrt{\mathbb{E}[|X(t,f;Q)|^2]\mathbb{E}[|Y(t,f;Q)|^2]}} \quad . \tag{II.61}$$

For a single realization of the processes, we could estimate it by omitting the ensemble averages, as first done in [Liu 1994]. In spite of all the smoothness of the wavelet transform, this would cause the same "worst estimator" phenomenon as in Eqs.(II.54, II.58): a modulus always equal to one for any signals and values of the quality factor!

We are led back to the problem of finding a good statistical estimation of the ensemble average in time and frequency, from single trajectories. In a non-stationary context, this issue can only be circumvented when many samples of the same experiment are available or many simultaneous equivalent but independent measurements are available, see e.g. [Le and Argoul 2016; Zhan et al. 2006].

Summary of sections 3 and 4

The probabilistic interpretation of fluctuations in recordings is introduced, by treating them as stochastic processes. This naturally confers to the quadratic expressions derived from the wavelet transform to role of statistical estimator for average quantities of the stochastic process: the (time-dependent) variance, covariance and power spectral density. The mean of the process is associated to the non-oscillating part of the linear time-frequency representation, corresponding to the limit $f \to 0$. The spectral estimation from a time average corresponds to an ergodic hypothesis, while the temporal estimation can be associated to an analogous self-similarity hypothesis.
The self-similar Gaussian noise is taken as an example, both computational and illustrated. We emphasize the case of the pink noise, characterized by a constant power log-frequency density, thus taking the role of the white noise in the multi-scale context of the wavelet analysis.
The correlation coefficient ρ_{xy} is then introduced as the normalized covariance, describes synthetically the strength and the sign of the linear similitude between two, possibly delayed, fluctuating signals. This global quantity can be estimated locally in time with the wavelet estimator, and extended to a complex coefficient through the use of the analytic version of the signals, extending the interpretation of the sign of correlation to a phase of global or temporal coherence. Its frequency counterpart, the spectral coherence γ_{xy}, follows naturally as the normalized power spectral density.
The question of estimating statistical quantities in the time-frequency remains unclear, although they can be properly defined from an ensemble average.

5 Wavelet-estimators in time and frequency: two paths, one destination

The time-frequency uncertainty must increase in order to reduce the variance of a statistical estimator [Pitton 2000], and to discriminate true coherence from spurious one. The statistics of the spurious coherence is also essential to assess the significance of the estimation. Two types of estimation exists: one involves orthogonal sequences called multi-tapers often linked to the Gabor approach (spectrogram estimation), the other uses a local average, often chosen in the wavelet approach. The first approach, discrete, directly provides the number of degrees of freedom of the statistical estimator, while the second one, continuous, offers an explicit control over its time-frequency localization. Neither of these strategies are limited to their favourite time-frequency representation, as is apparent from their direct applications to the Wigner-Ville distribution [Matz and Hlawatsch 2000; Orini, Bailon, Mainardi, Minchole, et al. 2009; Orini, Bailon, Mainardi, Laguna, et al. 2012; White and Boashash 1990] (that underlies both representations).

In the wavelet framework, we propose to harvest the best of both techniques in an explicit manner, resuming and pursuing the work of [Cohen and Walden 2010a,b]. By doing so, we reach the conclusion that a statistical estimation in time and frequency accounts for separating the quality factor used in time and the one used in frequency: their ratio is the number of degrees of freedom. This solves the puzzle of estimating a time-dependent spectral density and the associated coherence from single stochastic trajectories, in a canonical and synthetic way in the log-normal wavelet framework.

5.1 From multi-taper to multi-wavelet method

To improve the limitations of the practical estimation of the power spectral density $S_{xy}(f)$, also called the periodogram, multi-taper techniques have been introduced [Babadi and Brown 2014; Percival and Walden 1993]. The taper refers to the window function used to filter the data before performing the spectral decomposition, such as the Gabor window. The shape of the taper (sharp or smooth edges) may improve the spectral discrimination of relevant frequency bands from background noise, and off-peak frequencies. However, analyses of shortened segments increase the variance of the spectral estimators. Multi-taper spectral method was shown to have better spectral separation power than the single taper method [Thomson 1982]. The multi-taper method mimics independent realizations of the process by averaging together multiple spectra computed from multiple orthogonal taper functions applied to the same signal. Formalizing the optimization of the spectral concentration into an eigen-decomposition problem lead to an optimal solution called discrete prolate spheroidal sequence [Slepian 1978], resulting in an improved localization and reduced variance of the estimator. The number of tapers used is involved in the statistical distribution of the estimator as its number of degrees of freedom [Walden 2000], chosen smaller than the product of duration and bandwidth.

The multi-taper method can also be applied to time-dependent spectral estimations [Babadi and Brown 2014; Prerau et al. 2017], so that multi-taper-based time-frequency coherence estimator are available [Daly et al. 2004; Lovett and Ropella 1997; Yan Xu, Haykin, and Racine 1999]. Similarly, the variance of the wavelet-estimators could also be improved from the use of orthogonal sequences of wavelets, called "multi-wavelet" method or "eigenscalogram", implemented from the Morse family of orthogonal wavelets [Bayram and Baraniuk 2000; Cohen and Walden 2010b; Daubechies 1988]. This corresponds in our framework to the use of the orthogonal wavelets $\varphi_Q^{(k)}$ Eq.(I.46), normalized to a unit energy in the definition of $X(t,f;Q,k)$. Therefore, the time-dependent power spectral density defined in Eq.(II.38) can be estimated by:

$$S_{xy}(t,f;Q,K)|f| = \sum_{k=0}^{K-1} \frac{1}{K} X(t,f;Q,k)\overline{Y(t,f;Q,k)} \quad , \tag{II.62}$$

corresponding to a uniform average over the K first wavelets of the orthonormal basis.

While the variance of this estimator is reduced for a large number of degrees of freedom $n_d = K$, the time-frequency localization is necessarily degraded: the atom of uncertainty spreads, with few control on it. Replacing $\frac{1}{K}$ by positive weights w_k in Eq.(II.62) could help to tune it. The effective number of degrees of freedom is defined as $n_d = (\sum_k w_k)^2 / \sum_k w_k^2$ [Walden 2000]. However, the first order is sufficient: indeed, wavelets are well-localized so that the information aggregated to a certain location (t,f) by higher order wavelets is also available near (t,f) at zero order (see the change of wavelet formula in Appendix 3). This motivates the alternative strategy based on smoothing.

5.2 Degrees of freedom and uncertainty atoms with local averages

The reduction of the variance of the estimator can also be achieved from local average operations of the spectrum. This smoothing strategy allows to control the loss of resolution, contrary to the multi-taper approach. Many references can be found on wavelet-based coherence, see for instance [Cazelles et al. 2008; Chang and Glover 2010; Grinsted, Moore, and Jevrejeva 2004; Gurley, Kijewski, and Kareem 2003; Torrence and Compo 1998; Yaesoubi et al. 2015], but the diverse forms of smoothing are scarcely discussed. The spread of the smoothing kernel determines both the resolution of the analysis and the level of spurious coherence (variance of the coherence estimator).

We start by recalling the unique type of smoothing that is wavelet-compliant. Given a stochastic intensity field $S(t,f)$, its ensemble average can be estimated from the following "affine" convolution:

$$\langle S(t,f)\rangle_K = \iint S(t',f')K(f(t'-t),f'/f)\mathrm{d}t'\mathrm{d}f' \quad , \tag{II.63}$$

consistently with the wavelet time-frequency resolution, as in Eq.(II.28).

5.2.1 Smoothing in both domains

For estimating $S_{xy}(t,f)$, a smoothing in both dimensions of the product of wavelet transforms has been proposed [Torrence and Compo 1998]. This corresponds to using any positive kernel $K(u,v) = \chi(u,v)$ of unit weight ($\langle 1\rangle_\chi = 1$) to smooth $S(t,f) = X(t,f;Q)\overline{Y(t,f;Q)}\frac{Q}{\sqrt{\pi}|f|}$, providing alternative estimators to Eq.(II.63). Notice that such kernel generalizes all previous global integrals over time or frequency: a free (non-integrated) variable corresponds to a fully local kernel in this domain (a Dirac delta), while the kernel is fully global (constant) in the integrated (or averaged) one.

This estimator can be simplified by smoothing along a single dimension. Unless there is some need for averaging over a sophisticated time-frequency domain, we even argue that smoothing in a second dimension is useless for most applications, since the same effect is obtained by changing the quality factor of the wavelet beforehand. Smoothing in time assumes more stationarity (thus a weaker localization), while more scale-invariance is assumed when smoothing in frequency.

We can expect that the number of degrees of freedom has something to do with the number of uncertainty atom covered by the kernel K. A constant kernel supported by a time-frequency domain Ω, for instance of widths $nf\delta t$ and $m\delta\log f$ (using the practical wavelet resolution, see Eq.(I.41)), covers an area $\|\Omega\| = \iint_\Omega \mathrm{d}t\mathrm{d}f$, of the order nm. In the one-dimensional case, the kernel K does not spread over a surface but over a length, and we still expect it to cover about n atoms for the time smoothing alone or m atoms for the frequency smoothing alone. Thus, the wavelet is necessarily involved in the resulting estimator's degree of freedom.

5.2.2 Number of degrees of freedom from multi-taper

The kernel of a spectral estimator is usually thought as an operator acting directly on the signals or the spectra to produce the estimator:

$$xKy^\dagger = \iint x(t_1)\overline{y(t_2)}\check{K}(ft_1,ft_2)|f|\mathrm{d}t_1\mathrm{d}t_2 = \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\hat{K}(f_1/f,f_2/f)/|f|\mathrm{d}f_1\mathrm{d}f_2 \quad , \tag{II.64}$$

where K depends on a frequency variable f and the time variable has been set to $t=0$ in the stationary case. In the discrete and finite context of multi-taper estimators, the kernel is a Hermitian symmetric positive-definite matrix, whose eigenvectors are the multiple tapers and the effective number of degrees of freedom is readily computed from the eigenvalues [Walden 2000]. This computation is equivalently expressed directly from the kernel as a beautiful trace formula:

$$n_d = \frac{(\mathrm{Tr}\,K)^2}{\mathrm{Tr}\,KK^\dagger} = \frac{\left(\int \check{K}(u,u)\mathrm{d}u\right)^2}{\iint |\check{K}(u',u)|^2\mathrm{d}u'\mathrm{d}u} = \frac{\left(\int \hat{K}(v,v)\mathrm{d}v\right)^2}{\iint |\hat{K}(v',v)|^2\mathrm{d}v'\mathrm{d}v} \quad , \tag{II.65}$$

where the second and third equalities clarify the correspondence between the continuous integrals and symmetry $\overline{\check{K}(u,u')} = \check{K}(u',u)$, and the linear algebraic trace Tr and product $KK^\dagger$

with conjugated transposed matrix. Its consistency with the continuous multi-wavelet estimator Eq.(II.62) is straightforward (n_d is the number of orthonormal wavelets).

It remains to extend the computation of n_d to the wavelet-estimators and to interpret it as a number of time-frequency uncertainty atoms.

5.2.3 Entropic uncertainty from Wigner-Ville distribution

The inclusion of the wavelet into the kernel and the need for a rewriting in the form of time-frequency integrals lead to use the Wigner-Ville distribution $S(t,f) = W_{xy}(t,f)$ in Eq.(II.63). By comparison with Eq.(II.64), we identify:

$$\check{K}(u',u) = \int K(\tfrac{u'+u}{2}, v) e^{i\Upsilon v(u'-u)} \mathrm{d}v \quad ; \quad \hat{K}(v',v) = \int K(u, \tfrac{v'+v}{2}) e^{-i\Upsilon(v'-v)u} \mathrm{d}u \quad , \quad \text{(II.66)}$$

whose Hermitian symmetry translates into a real kernel $K(u,v) = \overline{K(u,v)}$.

This leads to the following information theoretic perspective for the logarithm of the number of degrees of freedom when expressed in time and frequency:

$$\log n_d = -\log \frac{\iint K(u,v)^2 \mathrm{d}u\mathrm{d}v}{\left(\iint K(u,v)\mathrm{d}u\mathrm{d}v\right)^2} = H_2[K] \quad . \quad \text{(II.67)}$$

This is the order 2 Rényi entropy, defined in Eq.(A.45) (extended to the bivariate case), of the time-frequency kernel K. Since an exponential entropy counts an effective number of "states" in the distribution $K(u,v)$, we reach the following interpretation: to gain statistical significance, n_d atoms of time-frequency uncertainty are integrated in the smoothing operation, and their distinction are lost to the time-frequency localization.

Note that the series of formulas Eqs.(II.64 - II.67) are "paradigm-agnostic", i.e. independent of the choice of the wavelet or Gabor approach. The case of the wavelet transform alone corresponds to the kernel $K(u,v) = W_{\psi_Q\psi_Q}(u,v)$, or equivalently $\hat{K}(v',v) = \overline{\hat{\psi}_Q(v')}\hat{\psi}_Q(v)$, that yields $n_d = 1$: the smoothing of the wavelets (a separated kernel) covers a single uncertainty atom. The same holds for the Gabor approach, whose atom was called a "logon" [Gabor 1946]. Using the bare Wigner-Ville distribution as an estimator corresponds to the degenerate case $n_d = 0$: no smoothing covers no uncertainty atom, directly related to its infinite variance [Pitton 2000; Stanković and Stanković 1993].

Although the time-frequency integral expression Eq.(II.67) for n_d is in the reach of several analyses on the topic [Cohen and Walden 2010a; Pitton 2000], it seems its usefulness remained unrecognized and never put into practice. As far as the entropic interpretation is concerned, similar expressions based on the Rényi entropy are used to measure the amount of information in time-frequency representations of the full signal [Baraniuk et al. 2001; Flandrin, Baraniuk, and Michel 1994; Jones and Parks 1990; Stanković 2001; Williams, Brown, and Hero III 1991], instead of the one available through the kernel K.

5.2.4 Explicit control of locality: natural versus practical

Now that we know how to compute the degrees of freedom associated to an estimator of the power density or the coherence, we come back to smoothing $X(t,f;Q)\overline{Y(t,f;Q)}\frac{Q}{\sqrt{\pi}|f|}$ with a kernel χ, possibly the simplest (one-dimensional). After examining the form of the smoothing in the wavelet-estimators $S_{xy}(f;Q)$ and $P_{xy}(t;Q)$, we propose to transpose their wavelet-smoothing into the kernel χ to define natural estimators for $S_{xy}(t,f;Q)$. We summarize our findings in Table II.1.

smoothing	*localization*	*degrees of freedom* n_d	*approximation*
$S_{xy}(f;Q)=\int \hat{x}_T(f')\overline{\hat{y}_T(f')}\lvert\hat{\psi}_Q(f'/f)\rvert^2\mathrm{d}f'\frac{Q}{\sqrt{\pi}\lvert f\rvert T}$			
f	Q	∞	$T=t_{\max}-t_{\min}\to\infty$
$S_{xy}(f;Q,T)=\langle X(t,f;Q)\overline{Y(t,f;Q)}\rangle_t\frac{Q}{\sqrt{\pi}\lvert f\rvert}$			
f	Q	$\sqrt{\frac{\lambda}{6}}\frac{\Upsilon f T}{Q}$	$1\ll Q\sim f\delta t\ll\lvert f\rvert T$
$P_{xy}(t;Q)=\iint x(t_1)\overline{y(t_2)}\int\psi_Q(f(t_1-t))\overline{\psi_Q(f(t_2-t))}\lvert f\rvert\mathrm{d}f\mathrm{d}t_1\mathrm{d}t_2\frac{Q}{\sqrt{\pi}}$			
t	Q	∞	$B=f_{\max}/f_{\min}\to\infty$
$P_{xy}(t;Q,B)=\int_{f_{\min}}^{f_{\max}}X(t,f;Q)\overline{Y(t,f;Q)}\frac{\mathrm{d}f}{\lvert f\rvert}\frac{Q}{\sqrt{\pi}}$			
t	Q	$\sim\frac{Q}{\sqrt{\Upsilon}}\log B$	guess
$S_{xy}(t,f;Q,Q)=X(t,f;Q)\overline{Y(t,f;Q)}\frac{Q}{\sqrt{\pi}\lvert f\rvert}$			
	Q	1	exact
$S_{xy}(t,f;Q_+,Q_-)=\iint\hat{x}(f_1)\overline{\hat{y}(f_2)}\sqrt{\hat{\psi}_{Q_+}(f_1/f_2)\hat{\psi}_{Q_-}(f_1f_2/f^2)}e^{i\Upsilon(f_1-f_2)t}\mathrm{d}f_1\mathrm{d}f_2\frac{Q_-}{\sqrt{\pi}\lvert f\rvert}$			
t,f	Q_+,Q_-	Q_+/Q_-	exact
$S_{xy}(t,f;nQ,Q)=\int X(t',f;Q)\overline{Y(t',f;Q)}e^{-(\Upsilon f(t'-t)/nQ)^2}\mathrm{d}t'\frac{2}{n}$			
t	$\sim nQ$	$\sqrt{1+n^2\lambda}$	$(nQ)^2\lambda\gg1$

Table II.1: *Comparison of wavelet-estimators for the power spectral density $S_{xy}(f)$, the variance $\sigma^2_{xy}(t)$ and the time-dependent power spectral density $S_{xy}(t,f;Q)$. The first column is the domain of smoothing, followed by the resulting localization in this domain expressed as a quality factor, see Eq.*(I.41)*. Next, the number of degrees of freedom n_d is given, with its potential approximation in the last column. When an integral in Eq.*(II.65) *is not tractable, it is approximated from the Laplace's method, see Appendix 8. We use the notation $\lambda=e^{\frac{1}{Q^2}}$.*

The computation of $S_{xy}(f;Q)$ in Eq.(II.26) is a natural example of smoothing multiplicatively along frequency, with the squared wavelet spectrum $|\hat{\psi}_Q(f'/f)|^2$ in the role of the kernel. We can choose a different parameter for the quality factor:

$$\chi_{\epsilon_-}(v)=|\hat{\psi}_{\epsilon_-}(v)|^2\quad,\quad\epsilon_-^{-2}=Q_-^{-2}-Q^{-2}\quad, \tag{II.68}$$

as in the formula for changing of quality factor Eq.(II.12); the frequency resolution is decreased to a quality factor $Q_-<Q$, letting room for exactly

$$n_d=\sqrt{1+(Q/\epsilon_-)^2}=\frac{Q}{Q_-} \tag{II.69}$$

statistical degrees of freedom, as computed from Gaussian integrals. The resulting estimator may be denoted $S_{xy}(t,f;Q,Q_-)$. Its numerical computation may, however, be complicated by the fact that efficient implementations of the wavelet transform do not use the same time-sampling at different scales.

In this respect, a time-smoothing with $\chi(f(t'-t))|f|$ is very convenient, and the statistical properties of its discrete implementation has been studied in comparison to the multi-taper approach [Cohen and Walden 2010a; Walden and Cohen 2012]. This is the strategy we apply for most estimations of the time-frequency coherence. Unluckily, the computation of $P_{xy}(t;Q)$

in Eq.(II.21) does not provide a wavelet expression corresponding to a one-dimensional kernel. Furthermore, in the absence of simple expressions, neither for $|\psi_Q(f(t'-t))|^2 f$ (as may be intuited), nor for its spectrum, we use a Gaussian window that best approximates it:

$$\chi_n(u) = \frac{\sqrt{2\Upsilon}}{nQ} e^{-(\Upsilon u/nQ)^2} \tag{II.70}$$

$$S_{xy}(t,f;nQ,Q) = \int X(t',f;Q)\overline{Y(t',f;Q)}\chi_n(f(t'-t))\mathrm{d}t'\frac{Q}{\sqrt{\pi}} \quad , \tag{II.71}$$

of time-width intended to be n resolution units, $nf\delta t = \frac{nQ}{\sqrt{p\Upsilon}}$, with $p = 2$ for a quadratic expression. Although we cannot compute exactly the corresponding degrees of freedom, its approximation from the Laplace's method:

$$n_d \approx \sqrt{1+n^2\lambda} \quad , \tag{II.72}$$

is indeed of the order n (for nQ large enough), see computation in Appendix 8.

Then, a natural (wavelet-based) time-smoothing seems missing at the place of its linear counterpart Eq.(II.13), increasing the quality factor to $Q_+ > Q$. This is clarified by examining the form of the smoothing in the time-dependent power $P_{xy}(t;Q)$. In fact, it is a two-dimensional kernel, equivalent to wavelet-transforming separately in times t_1 and t_2 before integrating over the resulting (equal) frequency variable. It can also be expressed in terms of the covariance estimator Eq.(II.39):

$$\begin{aligned} S_{xy}(t,f;Q_+,Q) &= \iint X(t_1,f;Q)\overline{Y(t_2,f;Q)}\int \overline{\psi_{\epsilon_+}(f'(t_1-t))}\psi_{\epsilon_+}(f'(t_2-t))|f'|\mathrm{d}f'\mathrm{d}t_1\mathrm{d}t_2\frac{Q\epsilon_+}{\pi|f|} \\ &= \iint R_{xy}(t_1,t_2;\epsilon_+)\overline{\psi_Q(f(t_1-t))}\psi_Q(f(t_2-t))|f|\mathrm{d}t_1\mathrm{d}t_2\frac{Q}{\sqrt{\pi}} \ , \ \epsilon_+^2 = Q_+^2 - Q^2 \ . \end{aligned} \tag{II.73}$$

This is equivalent to binding the separated frequencies in the wavelet filters $\overline{\hat{\psi}_Q(f_1/f)}\hat{\psi}_Q(f_2/f)$ with the factor $\sqrt{\hat{\psi}_{\epsilon_+}(f_1/f_2)} = \hat{\psi}_{\frac{\epsilon_+}{\sqrt{2}}}(f_1/f_2)$ (from the integration over f' in the first line), indeed increasing the quality factor in time to $Q_+ > Q$. The resulting number of degrees of freedom is simply and exactly:

$$n_d = \sqrt{1+(\epsilon_+/Q)^2} = \frac{Q_+}{Q} \quad , \tag{II.74}$$

which is precisely what we are looking for in Eqs.(II.70 - II.72).

5.2.5 Canonical kernel for the log-normal wavelet framework

The number of degrees of freedom for time or frequency smoothing Eqs.(II.69, II.74) are similar. This is more than a coincidence, since $Q = Q_+$ in Eq.(II.68) and $Q = Q_-$ in Eq.(II.73) yield the exact same estimator $S_{xy}(t,f;Q_+,Q_-)$. The identification of the kernel $\hat{K}(f_1/f, f_2/f)$ Eq.(II.64) in their definition Eqs.(II.68, II.73) confirms it:

$$\hat{K}(v',v) = \sqrt{\hat{\psi}_{Q_+}(v'/v)\hat{\psi}_{Q_-}(v'v)}\,\frac{Q_-}{\sqrt{\pi}} \quad , \qquad n_d = \frac{Q_+}{Q_-} > 1 \quad . \tag{II.75}$$

As a result, there is a single natural estimator for $S_{xy}(t,f;Q)$: it requires two different quality factors Q_+ in time and Q_- in frequency, and the associated number of degrees of freedom is simply $\frac{Q_+}{Q_-}$. We call $S_{xy}(t,f;Q_+,Q_-)$ the canonical wavelet-estimator in the log-normal

framework, which controls analytically both the time-frequency resolution (contrary to multi-wavelet estimators) and the statistical variability of the quadratic estimator (contrary to other smoothing estimators).

Eventually, we temper on the practical side the analytic impact of this estimator when confronted to its efficient implementation. Indeed, our current numerical estimation is based on the common two-step strategy of computing the wavelet transforms first and then to smooth their product. This is only compatible with the frequency smoothing computation of $S_{xy}(t, f; Q_+, Q_-)$ from Eq.(II.68). However, the decrease of the time sampling at increasing frequencies, which is a key for an efficient computation of the wavelet transforms in the first step, complicates the smoothing in second step. This practical limitation of our implementation of the wavelet transform suggests possible room for algorithmic improvements. We tried a one-step computation exploiting the sparsity of the kernel Eq.(II.75), with no gain in efficiency. This improvement may come from the use of a discrete wavelet transform as a first step, followed by a change to the log-normal wavelet in the second averaging step.

For lack of it, we mainly apply the practical time-smoothing Eq.(II.70), complying with our current implementation with a good efficiency, yet very close to the canonical estimation, with a precisely controlled localization, $nQ \approx Q_+$ and $Q = Q_-$, and degrees of freedom, $n_d \approx n$ (for Q and n large enough).

5.3 Significance of the coherence wavelet-estimator

The variability of the estimator can be responsible for a misleading interpretation, especially for the correlation and coherence estimators, whose square is as close to one as there are few degrees of freedom in the estimation (equal to one for $n_d = 1$), see Eqs.(II.54, II.58) and the section 4.2.2. This phenomenon is called spurious correlation or spurious coherence.

While the significance of the coherence estimation is often assessed *a posteriori*, using well-constructed surrogates (see [Chavez and Cazelles 2019] in the non-stationary case), it can be estimated beforehand from the typical statistics associated to the estimator.

In the absence of information on the characteristics of the correlated signals and associated noises, their estimated coherence can be compared to the one between two independent, jointly stationary and Gaussian noises. We interpret it as giving the minimum expected level of spurious coherence. Its statistics is very close to a single-parameter beta distribution, as verified from numerical simulations, from which explicit levels of significance are straightforwardly derived.

5.3.1 Distribution of the estimators

Beyond a number of degrees of freedom, the study of the bias and variance of an estimator, and further, its statistical distribution is of great interest. If known, it allows assessing analytically the significance of the estimation. In this context of inference, it is appropriate to study the distribution of the estimator applied to Gaussian noises. Indeed, a Gaussian process is the simplest one determined by the second moment that we are estimating (the power spectral density and the coherence), and often the only one that provides analytic results.

The distribution of wavelet-estimators, both for the multi-wavelet and the time-smoothing methods, has been investigated under the additional assumptions of stationarity of the Gaussian processes in [Cohen and Walden 2010a,b]. This assumption is appropriate to describe the influence of the ambient or instrumental noises present in the recordings. The result is as follows: the estimator for the power spectral density matrix is distributed according to a Wishart distribution (matrix χ^2) and the squared coherence estimator is distributed according to a so-called Goodman distribution [Goodman 1963]. They both depend on two parameters: the first one is the number of degrees of freedom n_d of the estimator, the second is the average of the

estimator (for the power spectrum or squared spectral coherence). The condition of stationarity can be weakened to a certain class of processes whose power spectra are modulated in time, but their spectral coherence is time-independent [Walden and Cohen 2012].

We focus on the spurious coherence associated to the previously introduced power density estimators, $\gamma_{\rm sp} = \gamma_{xy}(t, f; Q_+, Q_-) = S_{xy}/\sqrt{S_{xx}S_{yy}}$. A specificity of these wavelet-estimators is their constant number of degrees of freedom, whatever the time or frequency, so that the spurious coherence has a homogeneous distribution in the time-frequency plane that only depends on $n_d = \frac{Q_+}{Q_-}$. In contrast, a smoothing kernel of constant duration at all frequencies would have much higher spurious correlations at low frequencies than at higher ones, as described in [Gurley, Kijewski, and Kareem 2003; Torrence and Compo 1998].

The wavelet-estimator only corresponds to the true spectral coherence of the stationary Gaussian noises asymptotically, for $1 \ll Q_- \ll Q_+$. We do not know how to compute the expected spurious coherence $\mathbb{E}[|\gamma_{\rm sp}|^2]$ as a function of $\gamma_{xy}(f)$, Q_+ and Q_-. In practice, the background noises in each recording are arguably incoherent, so that we assume: $\gamma_{xy}(f) = 0$. In this case, $\mathbb{E}[\gamma_{\rm sp}] = 0$, and the angle of the spurious coherence $\gamma_{\rm sp}$ is uniformly distributed. We also know that $\mathbb{E}[|\gamma_{\rm sp}|^2]$, as the variance of the estimator $\gamma_{\rm sp}$, vanishes in the limit $n_d \to \infty$.

In this limit, the Goodman distribution for the spurious coherence $|\gamma_{\rm sp}|^2$ cannot be distinguished from the simpler beta distribution $\mathcal{B}(1, \beta)$ with a single parameter, of mean and cumulative distribution:

$$\mathbb{E}[|\gamma_{\rm sp}|^2] = \frac{1}{\beta + 1} \tag{II.76}$$

$$\Pr(|\gamma_{\rm sp}|^2 \leq \gamma^2) = 1 - (1 - \gamma^2)^\beta \quad , \tag{II.77}$$

where the parameter β takes the role of the number of degrees of freedom. In [Cohen and Walden 2010b], β corresponds to $n_d - 1$. However, the one that best approximates the Goodman distribution of the spurious coherence might be slightly different, since the second Goodman parameter has been neglected (set to 0).

Assuming a beta distributed spurious coherence $|\gamma_{\rm sp}|^2$ between incoherent noises is verified in practice for a certain parameter β, then we can build significance level for the estimated coherence $|\gamma_{xy}(t, f; Q_+, Q_-)|^2$ between two natural recordings. Significance levels correspond to p-values derived from the beta distribution Eq.(II.77):

$$p(\gamma^2) \equiv \Pr(|\gamma_{\rm sp}|^2 > \gamma^2) = (1 - \gamma^2)^\beta \tag{II.78}$$

$$\gamma^2(p) = 1 - p^{\frac{1}{\beta}}. \tag{II.79}$$

Thus, the threshold $\gamma^2(p)$ for observing a significant time-frequency coherence γ_{xy} between some signals $x(t)$, $y(t)$ increases when β and the number of degrees of freedom $n_d = \frac{Q_+}{Q_-}$ decreases. This means that a low time-frequency resolution ($Q_- \ll Q_+$) is required to observe small coherences values reliably. Conversely, strong coherences can still be observed significantly when the time-frequency resolution is high (and $Q_- < Q_+$).

The beta distribution was already proposed to evaluate the significance of the time-frequency coherence in the context of a multi-wavelet estimator [Brittain et al. 2007], as an extension of the significance for the spectral coherence. Indeed, the spectral coherence also follows the Goodman distribution [Carter 1987], hence the beta distribution [Gish and Cochran 1988] when no coherence is expected. Its origin lies in the characterization of the significance for the sample estimator of the correlation coefficient; for two independent random Gaussian vectors x, y of length N, we have the spurious correlation $|\rho_{\rm sp}|^2 = |\Gamma_{xy}|^2 \sim \mathcal{B}(\frac{1}{2}, \frac{N}{2} - 1)$, and its complex extension $|\Gamma_{x^+y^+}|^2 \sim \mathcal{B}(1, \frac{N}{2} - 2)$, interpreting the random vectors as sampled signal and using their analytic version.

The relevance of the beta distribution to evaluate the statistical significance and the correspondence between n_d and β is then assessed numerically.

5.3.2 Numerical verification

To verify the relevance of the previous formulas and their robustness to the discretization of continuous expressions in numerical computations, we simulate the spurious coherence to check its statistical distribution in the time-frequency plane. We use the practical coherence estimator $\gamma_{\rm sp}(t,f) = \gamma_{xy}(t,f;nQ,Q)$ defined from a time-smoothing in Eq.(II.71), very close to the canonical estimation with $nQ \approx Q_+$, $Q = Q_-$, for which the number of degrees of freedom, close to n, is approximated by Eq.(II.72): $n_d \approx \sqrt{1+n^2\lambda}$. Obtained from the Laplace's method, see Appendix 8, the precision of this approximation is also examined from the numerical integration of the problematic integral Eq.(A.65).

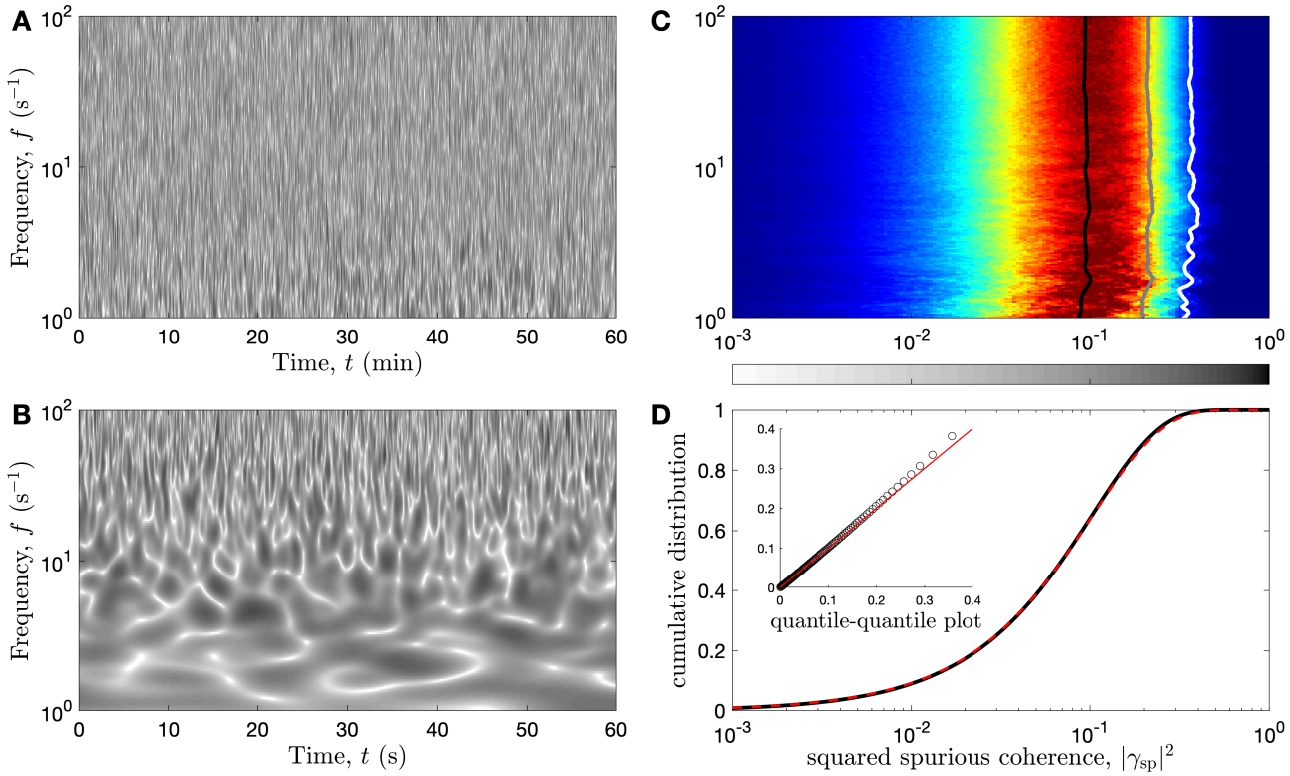


Figure II.4: *Simulation of the spurious coherence from two independent pink (log-correlated Gaussian) noises x, y: $\gamma_{sp}(t,f) = \gamma_{xy}(t,f;nQ,Q)$, with the quality factor $Q = 5$, and the time-smoothing parameter $n = 10$. (A) Realization of the squared spurious coherence $|\gamma_{sp}(t,f)|^2$. (B) Zoom on a shorter timescale, showing the characteristics of the time-smoothing. The colour scale for the squared coherence in (A) and (B) is aligned with the x-axes of (C) and (D). (C) Density map that represents the distribution of $|\gamma_{sp}|^2$ in time for each fixed frequency; a horizontal slice thus corresponds to a histogram: red indicates a high density and blue a low density. The black line indicates its mean for each frequency, the grey line is the quantile at 0.9 and the white line is the quantile at 0.99. (D) Sample cumulative distribution of the simulated squared spurious coherence (black line), and beta distribution $\mathcal{B}(1,\beta)$ Eq.(II.77) with estimated parameter $\beta = 9.6$ (red dashed line). Quantile-quantile plot (inset) of the sample versus estimated beta distributions (black circle for each percentile), indicating a close to ideal alignment (red line).*

The spurious coherence $|\gamma_{\rm sp}(t,f)|^2$ is simulated from two independent Gaussian noises $x(t), y(t)$ (real and stationary) of length 2^{21}, which lasts more than 1 hour when sampled at 500 Hz. We vary the colour of the noises (both pink or white), the quality factor ($Q = 5, 10$) and the value of the time-smoothing parameter from $n = 2$ to 100. In order to avoid border effects, that affect the value of the spurious coherence, we only retain the frequency range from 1 Hz to

100 Hz and use periodic boundary conditions in time. The discretized time step is taken more than 20 times smaller than the time resolution. Eq.(II.76) provides a simple relation between the parameter β and the spurious coherence: $\beta + 1 = \mathbb{E}[|\gamma_{\rm sp}|^2]^{-1}$. Estimating the ensemble average, we can compare the empirical distribution obtained from the simulation with the beta distribution.

The case of the pink noise for $Q = 5$ and $n = 10$ is illustrated in Fig.II.4. An obvious property of our time-frequency coherence estimator is its homogeneous level of spurious coherence everywhere in the time-frequency plane, see Fig.II.4 (A, B). In particular, it does not change with frequency, as can be observed in the density map of Fig.II.4 (C) obtained from histograms of $|\gamma_{\rm sp}(t,f)|^2$ for each frequency. This is due to the use of the wavelet-compliant time-smoothing Eq.(II.63), adapted to each scale.

Exploiting this property, the data aggregated from all frequencies can be merged to estimate $\beta = \beta(Q,n)$ from the time-frequency average of $|\gamma_{\rm sp}(t,f)|^2$. A sufficient time resolution (more than 10 per wavelet resolution δt) and suitable frequency integration depending on the frequency sampling, $\iint \mathrm{d}f\mathrm{d}t = \iint f\mathrm{d}t\mathrm{d}\log f$, are critical aspects of this estimation. We obtain $\langle|\gamma_{\rm sp}|^2\rangle_{t,f}^{-1} = \beta + 1 \approx 10.6$, which is intermediate to $n+1 = 11$ and to $n = 10$ or $n_d \approx 10.25$. The accuracy of the beta distribution is strikingly illustrated in Fig.II.4 (D), where the sample cumulative distribution and the beta distribution with parameter $\beta \approx 9.6$ are precisely superimposed. The quantile-quantile plot in the inset helps visualize the slight departure from a true beta distribution in the last percentiles. This precision improves at higher n and degrades in the limit $n \to 1$. It means that the p-value $p(\gamma^2)$ associated to a strong coherence are slightly overestimated (its high significance is slightly underestimated), which is unimportant compared to the actual coherence value γ^2. The coherence value loses its relevance at a low significance.

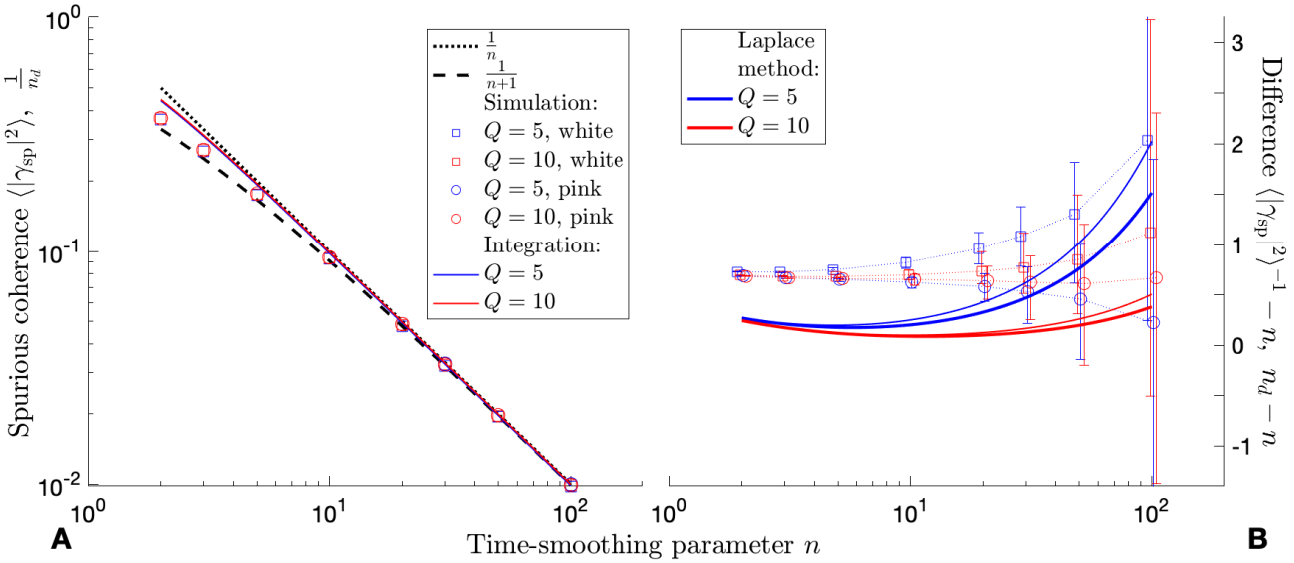


Figure II.5: *Relation between the spurious level* $\langle|\gamma_{sp}|^2\rangle = \frac{1}{\beta+1}$ *and the number of degrees of freedom* n_d *of the coherence estimator* $\gamma_{xy}(t,f;nQ,Q)$*, Eq.*(II.71)*, when varying* n*,* Q *and the colour of the noises* x, y*. (A) Spurious level of coherence, simulated (markers), estimated as* $\frac{1}{n_d}$ *from numerical integration (thin lines), or approximated by* $\frac{1}{n}$ *(black dotted line) or* $\frac{1}{n+1}$ *(black dashed line). (B) Details of the result:* n *is subtracted from* $\langle|\gamma_{sp}|^2\rangle^{-1}$ *and* n_d*. Markers and error bars indicate the medians and dispersions from the 0.05 to 0.95 quantiles. The precise estimation of* n_d *from numerical integration of Eq.*(II.65) *is compared to its approximation from the Laplace's method,* $\sqrt{1+n^2\lambda}$ *(thick lines). Quality factors:* $Q = 5$ *(blue) or* $Q = 10$ *(red). Gaussian noises and Hurst exponents: white* $H = -\frac{1}{2}$ *(square) and pink* $H = 0$ *(circle).*

The decreasing level of spurious coherence for increasing values of n is shown in Fig.II.5

(A), with few variations when changing the colour of the noise or the quality factor. At small numbers of statistical degrees of freedom, the level of spurious coherence is overestimated by $\frac{1}{n_d}$ (thin lines), where n_d is precisely computed from numerical integration. Both estimates from numerical simulation and integration appear bounded by $\frac{1}{n+1}$ and $\frac{1}{n}$.

More details are visible in Fig.II.5 (B), where we plot the difference between n_d or $\beta+1 = \langle|\gamma_{\text{sp}}|^2\rangle_{t,f}$ and n. The simulations are reproduced 100 times for each marker, and the 0.05 and 0.95 quantiles are used to construct an error bar around the median. They show an increasing dispersion caused by a decreasing number of effectively independent time steps (of duration $n\delta t$). In spite of this dispersion, we can clearly observe an increasing discrepancy between the simulations from pink noises and white noises, stronger for $Q = 5$ than for $Q = 10$. For instance, for $Q = 5$ and $n = 10$ we have $\beta = 9.82 \pm 0.05$ for the white noise ($H = -\frac{1}{2}$), while $\beta = 9.62 \pm 0.05$ for the pink noise ($H = 0$). While this effect is small, no such difference between Gaussian processes was predicted, possibly pointing at limitations in our numerical implementation.

In Fig.II.5 (B), the Laplace approximation of n_d (thick lines) is also compared to its estimates from numerical integration (thin lines): the Laplace's method slightly underestimates n_d, especially at increasing degrees of freedom. A comparison of $Q = 5$ and $Q = 10$ suggests that this underestimation amounts to a factor $\lambda^{-\frac{1}{8}}$. The difference between n_d and n is also more important for a small quality factor since $\sqrt{1+n^2\lambda} \sim n\lambda^{\frac{1}{2}} = ne^{\frac{1}{2Q^2}}$.

Although they are close, no precise relation can be given between β and n_d: $\beta \approx n_d - 0.5$ is acceptable for $n < 10$, and $\beta \approx n_d - 1$ is compatible with the simulations at higher n in spite of an important dispersion (± 1) for various quality factors and noise colours. This is a limitation of approximating the Goodman distribution as a beta distribution. Nevertheless, this precision is sufficient to evaluate the significance levels for our numerical implementation of the coherence estimator.

For instance, consider a time-frequency coherence γ_{xy} between two recordings $x(t)$, $y(t)$ with $n_d = 10$ degrees of freedom, implemented as $Q = 5$ and $n = 10$. Using $\beta = 9.7 \pm 0.5$, the threshold for observing a significant coherence with p-value $p = 0.1$ (90% level of significance) is $|\gamma_{xy}|^2 > 1 - 0.1^{\frac{1}{\beta}} = 0.21 \mp 0.01$, which is precise enough. A coherence analysis with $n_d = 10$ has a quite precise time-frequency resolution, but only detects significant coherence of modulus $|\gamma_{xy}| > \sqrt{0.21} \approx 0.46$. In contrast, 5 times lower time-frequency resolution ($n_d = 50 \approx \beta + 1$) would distinguish a coherence of modulus near 0.2 as significant ($p < 0.1$), and a coherence of modulus 0.5 with a very high significance, $p(0.5^2) < 10^{-6}$.

We conclude that the beta distribution $\mathcal{B}(1, \beta)$ is sufficient to evaluate *a priori* the significance levels of the time-frequency coherence estimator in many practical situations. Although the underlying assumptions of independent, jointly stationary and Gaussian noises can be difficult to test and even incorrect in many cases. In the absence of more information, this remains an effective method that we can interpret as giving the minimum expected level of spurious coherence. The relevance of this approach can be confirmed by comparing spurious coherence results with the coherence of surrogates. The main limitation for this approach is the situation of a very small number of degrees of freedom, in which case the Goodman distribution becomes very different from the beta distribution and concentrates near $|\gamma_{\text{sp}}|^2 = 1$. Border effects are another limitations, causing a local drop of the significance at numerical borders, such as initial and final times (when periodic boundary condition does not hold), and the maximal (Nyquist) frequency (half of the sampling frequency f_s of the signals) in the range of the coherence resolutions $n\delta t$ and $\delta \log f$.

5.4 Cross-talk between physiological sensors

Here, we anticipate the applications in part B of the analytical tools described in part A of this thesis, in order to illustrate the time-frequency coherence of natural recordings and their significance. The recordings are an electroencephalogram (EEG), an electrocardiogram (ECG) and an airflow sensor signal (AF), respectively measuring the neural, cardiac and respiratory activities of a sleeping person during a full night. These clinical measures are commonly found in a polysomnography, that contains many such simultaneous recordings. Their precise description is left for part B. We show how the coherence puts into evidence cross-talks between these physiological sensors.

Based on the evaluation the statistical significance, a synthetic representation for the time-frequency coherence is constructed. The relevance of this approach can be confirmed by introducing a surrogate signal in the analysis.

5.4.1 Neural, cardiac and respiratory activity in time and frequency

Fig.II.6 (A, B, C) presents their amplitude in time and frequency from a wavelet transform, $X(t, f; Q)$, first step to compute their time-frequency coherence. The cardiac and respiratory rhythmic oscillations are represented as horizontal lines in the ECG and AF (B, C), of high amplitude and fluctuating frequency (the heart and breathing rates), with vertical lines at higher frequency for the ECG reflecting its short and intense pulses. In contrast, the EEG (A) contains a mixture of horizontal and vertical lines, and a wide-band intensity similar to a coloured noise, heterogeneously distributed in time and frequency. At the top, a *hypnogram* indicates the wake and sleep stages of the person, impacting the physiological (especially the EEG) activity. Fig.II.6 (D) shows the effect of shuffling uniformly the phase of the EEG spectrum $\hat{x}(f)$: the initially non-stationary signal is "stationarized" while the global power spectral density is conserved. This yields a phase-randomized surrogate for the EEG, that will serve as an alternate method to assess the significance of coherence [Lancaster et al. 2018].

Computing the wavelet transform in such a large time-frequency domain is extensive in memory, and managed by a careful time-frequency sampling that approximates its scale-free resolution. The frequency domain is divided into frequency bands of an octave in which the wavelet transform is estimated as distinct matrices. Each matrix is a time-frequency image with a geometric frequency sampling and a linear time sampling, whose steps are a fraction (for smoothness) of the practical resolutions δt and $\delta \log f$ Eq.(I.41). Starting from the highest octave: each time we compute the next matrix an octave below, we down-sample it in time by a factor 2 (akin to the discrete orthogonal wavelet transform). This accounts for the frequency dependence of the time resolution $\delta t = \frac{Q}{f\sqrt{\dagger}}$. In this way, no memory is wasted on unnecessary precision in time or frequency, while controlling the degree of smoothness of the representation. A sufficient time sampling (10 steps per δt or more) will be critical for a precise estimation of the coherence from the time-smoothing method.

5.4.2 Time-frequency coherence between polysomnography recordings

In the time-frequency coherence computed between two recordings, regions of significant coherence provides a way to locate the oscillations or fluctuations that are jointly collected by both measuring apparatus, regardless of their intensity in each recording. They indicate a connectivity or "cross-talk" between the sensors, of physiological or instrumental origin, which are preferably minimized for an optimal specificity of each measure. Such cross-talks can be observed between the EEG, ECG and AF signals, even though they are very different.

The time-frequency coherence $\gamma_{xy}(t, f; nQ, Q)$ between pairs of recordings x, y, computed from the practical time-smoothing estimator, is represented in Fig.II.7 (A, C, E, G), while the

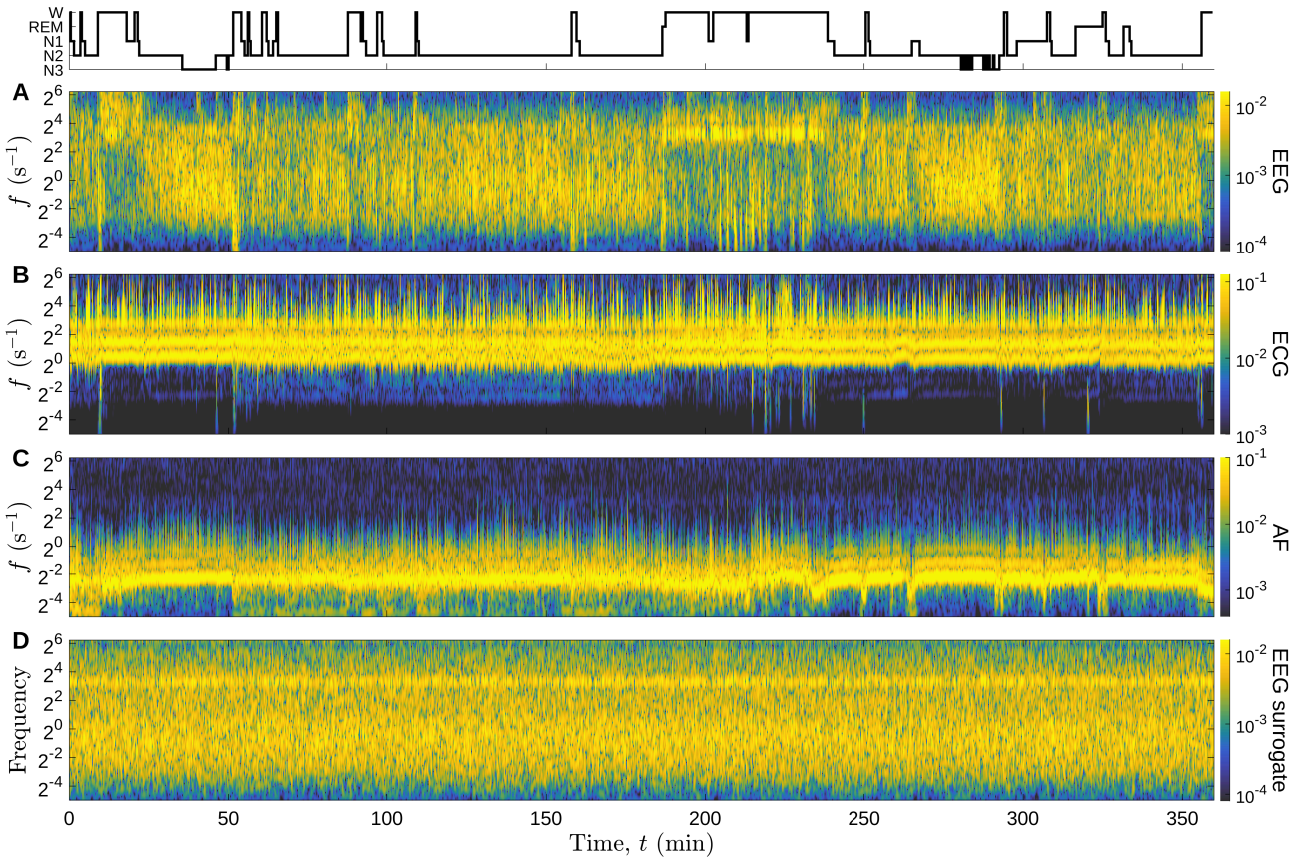


Figure II.6: *Polysomnography recordings of subject 04 from the database slpdb, in time and frequency. (A) neural activity from the EEG (C3-O1), (B) cardiac rhythm from the ECG, (C) breathing rhythm from the AF, (D) phase-randomized EEG surrogate. The amplitude (twice the modulus) of the wavelet transform is colour-coded on a logarithmic scale, and computed with the quality factor $Q = 5$. The amplitudes have the physical unit of the signals: all are in millivolt (mV), except for the airflow (AF) which is in litre per second ($l{\cdot}s^{-1}$). The lowest amplitude in the colour bar corresponds to the resolution of the signals.*

spectral coherence $\gamma_{xy}(f; Q)$ (colour-coded phase and plotted modulus) is shown in panels (B, D, F, H). The coherence modulus of this wavelet estimator is compared to its Gabor counterpart, for a simple window of 5 min (Welch's method). Both coherence moduli are consistent, apart from their different resolution, obvious at high frequencies.

For both the time-frequency and the spectral coherences, the phase difference measured as its angle (complex argument) is, once again, colour-coded with the hues in the chromatic circle. As is Fig.II.1 (D), the modulus of the coherence is coded into the colour saturation. This hue-saturation coding is ideal for coherence maps, see [Yaesoubi et al. 2015] for a similar use for fMRI signals. The modulus and its significance are difficult to read in such a synthetic representation. We improved its interpretation by discretizing the levels of coherence and associated saturation. In order to distinguish an important range of coherence values, we have chosen $n = 50$, i.e. a relatively large number of degrees of freedom: $n_d \approx 51 \approx \beta + 1$ according to Fig.II.5 (B) for $Q = 5$. We use Eq.(II.78) to evaluate their significance. The first threshold of significance is set to a p-value 10^{-1}, i.e. a coherence modulus $\gamma(10^{-1}) \approx 0.21$, below which the time-frequency region is coloured in white (considered incoherent). Up to the next threshold $\gamma(10^{-3}) \approx 0.36$, the regions of small but significant coherence modulus have the least saturated colour. For $\gamma_{xy} > 0.36$, the coherence is very significant and the saturation increases at each threshold $\gamma = 0.5, 0.7$ (maximum above 0.7).

The most coherent region lies around the breathing frequency, 0.2 Hz, between ECG and

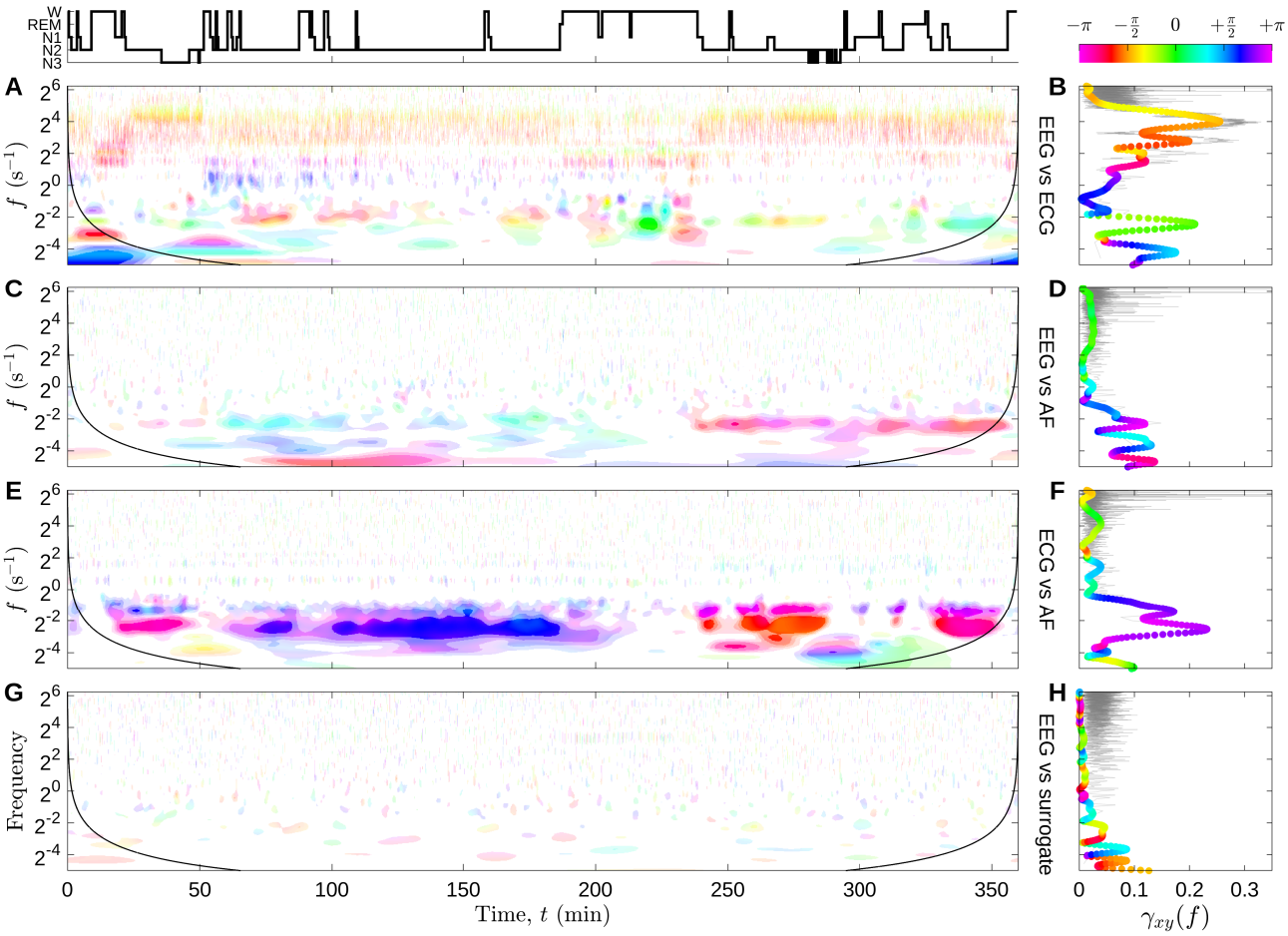


Figure II.7: *Time-frequency coherence $\gamma_{xy}(t, f; nQ, Q)$ (A, C, E, G) and spectral coherence $\gamma_{xy}(f; Q)$ (B, D, F, H) computed between pairs of physiological recordings (see Fig.II.6), x versus y, with parameters $Q = 5$ and $n = 50$. (A, B) EEG versus ECG, (C, D) EEG versus AF, (E, F) ECG versus AF, and (G, H) EEG versus its phase-randomized surrogate. The phase of coherence is colour-coded with hues of the chromatic circle. In (A, C, E, G), colour saturations code for ranges of coherence moduli $|\gamma_{xy}|$, delimited by the lower thresholds $\gamma(10^{-1}) \approx 0.21, \gamma(10^{-3}) \approx 0.36, 0.5, 0.7$. Black lines materialize a distance $n\delta t$ from the initial and final times, beyond which border effects are possible. In (B, D, F, H) the modulus of the spectral coherence, estimated from Welch's method with 5 min windows (thin grey line), is provided for comparison with the wavelet estimator (coloured spots).*

AF, Fig.II.7 (E, F). This strong coherence is the imprint of the breathing rhythm in the cardiac recording, although of very weak amplitude in the ECG, see Fig.II.6 (B). The time-frequency coherence Fig.II.7 (E) shows that this strong coherence ($|\gamma_{xy}| > 0.7$) is intermittent and its phase varies close to opposition: $\pm\pi \pm \frac{\pi}{3}$. This explains the smaller value of the spectral coherence modulus in Fig.II.7 (F). This cardio-respiratory connectivity seems specific to the sleep stage N2.

The same cross-talk appears with less coherence between EEG and AF, in Fig.II.7 (C), especially in the second part of the night and again in phase opposition: the breathing rhythm is also present in the neural recording, although hardly visible in the EEG, Fig.II.6 (A). Corresponding to even weaker EEG oscillations, another coherent region ($|\gamma_{xy}| \sim 0.5$) in phase opposition (magenta) can also be noticed below the breathing frequency, that will be linked to the occurrence of sleep apnea in chapter V.

As a result, some coherence is also observed between EEG and ECG at the breathing frequency, together with additional coherent regions, for instance when the person is awake

(187 to 237 min), characterized by shorter duration, broader frequency range and various phase. We interpret them as motions artefacts, sudden and intense events associated to body motions, and visible as thin vertical structures in Fig.II.6 (A, B).

Contrary to other the pairs, EEG and ECG in Fig.II.7 (A) are also coherent at higher frequencies, in a broad band ranging from 1 to 30 Hz, with low to high significance (up to $|\gamma_{xy}| \sim 0.5$). This is caused by the presence of the cardiac rhythm in the EEG, under the form of a contaminating ECG activity. The coherence range from 6 to 30 Hz is due to the short cardiac pulses. They must have a very low amplitude in the EEG, since they are invisible from Fig.II.6 (A) or from a direct inspection of the EEG signal. In this particular range, the coherence phase $\phi_{xy} = \phi_{\text{EEG}} - \phi_{\text{ECG}}$ is close to quadrature $-\frac{\pi}{2}$, and a dependence to frequency is noticeable. The phase trend between 6 and 30 Hz can be more precisely described as: $\phi_{xy} \approx \phi_0 + \Upsilon f \tau$, where $\tau \approx 7$ ms and $\phi_0 \approx -\frac{2\pi}{3}$. Therefore, the apparent phase quadrature, between the ECG pulses and the corresponding ones in the EEG, is better modelled as a greater phase shift ϕ_0 compensated by a delay τ. This is a first example of delay analysis from the coherence phase.

The level of spurious coherence is illustrated in Fig.II.7 (G), computed between the EEG and its phase-randomized surrogate [Lancaster et al. 2018]. Thus, the estimation of the significance for low coherence values can be controlled and visualized: the surrogate coherence only exhibits scattered spots of significance $10^{-3} < p < 10^{-1}$ (consistent with a density of false positives of about 10% in the time-frequency plane) with a random phase. The size of these spots is representative of the resolution of the coherence analysis in the time-frequency plane: their area is of the order of $n = 50$ time-frequency atoms. The same incoherence or lack of significance is found in (C, E) at high frequency. The introduction of a surrogate signal confirms the relevance of our approach, based on independent, jointly stationary and Gaussian background noises, to predict the level of spurious coherence.

5.4.3 A note on linear correction and partial coherence

Eventually, the time-frequency coherence and power densities between two recordings can serve to correct numerically their cross-talk, even when appearing with non-stationary and multi-scale characteristics. This can be particularly useful if a cross-talk between sensors could not be avoided instrumentally. A recording x corrected "linearly" from the influence of the recording z, can be constructed from the time-frequency representation:

$$X_{/z}(t,f;Q) = X(t,f;Q) - \frac{\mathbb{E}[X(t,f;Q)\overline{Z(t,f;Q)}]}{\mathbb{E}[|Z(t,f;Q)|^2]} Z(t,f;Q) \tag{II.80}$$

$$= X(t,f;Q) - \gamma_{xz}(t,f;Q)\Lambda_{xz}(t,f;Q)Z(t,f;Q) \quad ; \quad \Lambda_{xz} = \sqrt{\frac{S_{xx}}{S_{zz}}} \quad . \tag{II.81}$$

The coherence γ_{xz} gives a weight and a phase shift to the correction term, and Λ_{xz} scales it. The signal $x_{/z}(t)$ is retrieved by integrating over log-frequencies, see Eqs.(II.9, II.11). By construction, it is incoherent with $z(t)$. A similar correction operation could be achieved from the spectral coherence (or the temporal coherence) only, but the resulting linear correction would be insensitive to non-stationary (respectively multi-scale) characteristics. Thus, a linear correction in time and frequency is much more precise than in one domain only. In practice, we estimate the ensemble averages over a sufficient number of statistical degrees of freedom, $n_d \approx n$ in the following practical estimator:

$$X_{/z}(t,f;nQ,Q) = X(t,f;Q) - \frac{S_{xz}(t,f;nQ,Q)}{S_{zz}(t,f;nQ,Q)} Z(t,f;Q) \quad . \tag{II.82}$$

The influence of a recording can also be corrected from a coherence analysis, by comparing two recordings both corrected for the influence of the third one, $x_{/z}, y_{/z}$. The result is a *partial* coherence:

$$\gamma_{xy/z} = \frac{\gamma_{xy} - \gamma_{xz}\overline{\gamma_{yz}}}{\sqrt{(1-|\gamma_{xz}|^2)(1-|\gamma_{yz}|^2)}} \quad , \tag{II.83}$$

where the ideal terms are replaced by time-frequency estimators in practice. This is the first step of a multivariate (multi-signal) extension for the statistical analysis of correlations, called partial and multiple correlations [Dwyer 1940; Kendall, Stuart, and Ord 1987]. The time-frequency coherence estimator is compliant with this theory, able to exploit these relations between the matrix entries γ_{xy}, hence opening perspectives for a time-frequency analysis of multiple complex recordings beyond both the stationary and the bivariate analysis [Aguiar-Conraria, Soares, and Sousa 2018; Aguiar-Conraria and Soares 2014]. The principal axes of maximum covariance have also been proposed to approach power patterns in multiple time series, by applying a singular value decomposition to the wavelet cross-spectra S_{xy} [Rouyer et al. 2008].

Summary of section 5

In the log-normal wavelet framework, there is a canonical time-frequency estimator for the power density $S_{xy}(t,f;Q_+,Q_-)$, characterized by a simple and explicit relation between statistical and time-frequency uncertainties. The number of statistical degrees of freedom:

$$n_d = \frac{Q_+}{Q_-} \quad , \tag{II.84}$$

is a ratio of quality factors, controlling respectively the time and the frequency resolutions. In addition to the trade-off between the time and frequency resolutions controlled by a single quality factor Q, this second balance between statistical significance and resolution is the reason for distinguishing a higher quality factor Q_+ for the time-resolution and a lower one Q_- for the frequency-resolution.

The canonical estimator can be approximated in practice from a time-smoothing operation over n wavelet units of duration, so that $Q_+ \approx nQ$, $Q_- = Q$ and $n_d \approx \sqrt{1+n^2\lambda}$.

For the analysis of the time-frequency coherence $\gamma_{xy}(t,f;Q_+,Q_-)$, this means that the higher n_d, the more significantly we can distinguish low coherence values from the spurious one expected in the background, but the lower is the time-frequency resolution. Conversely, the lower n_d, the better we can localize coherence between two recordings, but the lower its statistical significance.

The time-frequency coherence estimated from a pair of independent and jointly stationary Gaussian noises can be interpreted as the minimum expected level of spurious coherence. The distribution of the squared spurious coherence is closely approximated by a beta distribution of parameter $\beta \approx n$, which allows to predict the significance of the time-frequency coherence observed between two natural signals.

Coherence thresholds corresponding to levels of significance can thus be evaluated from the pair of quality factors, helping to choose them accordingly. By calibrating the hue-saturation colour coding of this complex-valued map, we construct a synthetic visualization of the significant time-frequency coherence.

This framework makes a bridge between two different perspectives on complex recordings: oscillations and fluctuations. It clarifies the relation between regular features, localized in time or in frequency, and irregular ones relying on a statistical ensemble.

Chapter III

Estimation of rates modulation for a fluctuating harmony

In this chapter, we aim at constructing practical and generic wavelet estimators for the characteristic and time-dependent frequency of an arbitrary signal that exhibits rhythmic oscillations. Such signals, oscillating with a certain (but imprecise) regularity, are ubiquitous in nature. A very familiar example is sound [Daubechies and Maes 1996; Kronland-Martinet, Morlet, and Grossmann 1987], in particular the one of our voice that will serve as an illustration.

We first present existing concepts built on the time-derivative of the wavelet transform, namely ridges [Carmona, Hwang, and Torrésani 1997] and frequency reassignment [Auger and Flandrin 1995]. These ideas are very generic, but their practical use often requires additional techniques to extract rhythmic features, such as frequency and amplitude modulations. We introduce a set of rudimentary operations on the wavelet transform that achieve this goal for the practical applications of the next chapters. With an increasing level of specificity and associated control parameters, the resulting family of estimators is inferred from the formalism of the previous chapters, spanning from the instantaneous frequency of the analytic signal at the beginning of chapter I, to the time-frequency coherence at the end of chapter II.

1 Cyclic dynamics

Some notions associated to oscillatory systems are recalled in this section. A voice sample, recorded from a female singer, serves to illustrate how the temporal and spectral intuitions are conserved and combined in the wavelet representation, see Figs. III.1 and III.2.

1.1 Time-periodicity and harmonic spectrum

From a dynamical perspective, a system is oscillatory if its trajectory is cyclic in a certain state space. In particular, a *dynamical* phase $\phi(t) \in [0, \Upsilon)$ can be associated to it as a map from time to the (angular) position in the cyclic orbit, that grows monotonically with time and steadily within each cycle [Pikovsky, Rosenblum, and Kurths 2001]. Given a signal $x(t)$ that measures the state of the oscillatory system, the real and imaginary parts of its analytic version $x^+(t)$ Eq.(I.2) can be used as state space coordinates [Kralemann, Cimponeriu, et al. 2008]. The trajectory of the voice signal Fig.III.1 (A) in this state space is represented in (B). We can observe a well-defined cyclic orbit, which is far from circular, so that the angular argument of $x^+(t)$ does not coincide with the dynamical phase.

A periodic signal, $x(t+\tau) = x(t)$, with period τ, is the particular case where the rate of the dynamical phase is constant: $\dot{\phi}(t)/\Upsilon = f_1 = \tau^{-1}$. The repeated pattern is called the *waveform*. A slow modulation of the period, $\tau(t)$, does not affect the waveform but slightly dilates it in

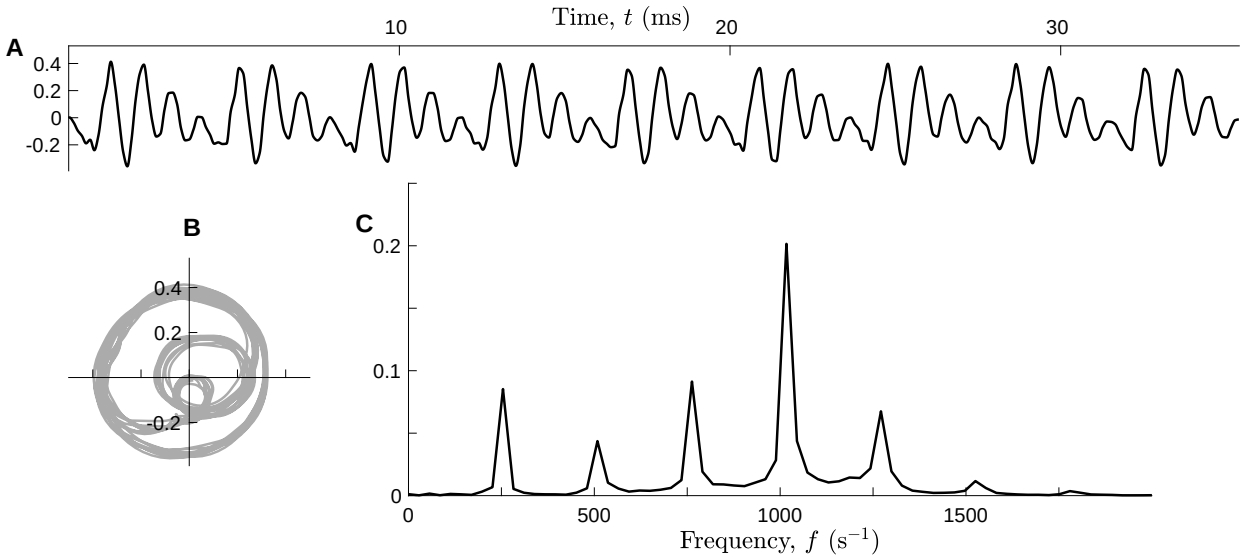


Figure III.1: *Voice recording: "A" vowel sung by a female singer, short selection. (A) Recorded signal $x(t)$. (B) Trajectory of the analytic signal $x^+(t)$ in the complex plane. To be aligned with the real signal, $x(t) = 2\Re\{x^+(t)\}$ is plotted on the Y-axis, while $2\Im\{x^+(t)\}$ is on the X-axis. (C) Amplitude spectrum $2|\hat{x}_T(f)|/T$ from the Fourier transform $\hat{x}_T(f)$ of the selection of duration $T = 35ms$. All amplitudes have the signal's (arbitrary) unit.*

time. Similarly, a slow modulation of the oscillation amplitude $A(t)$ dilates the waveform in the signal's space, and the addition of another term $r(t)$ in the signal, such as a constant, a trend or a noise, shifts the original waveform in a deterministic or random direction. This is summarized by the following model:

$$x^+(t) = A(t)H(e^{i\phi(t)}) + r(t) \quad , \tag{III.1}$$

where the function $H(z)$ represents the waveform. The simplest wave is called *harmonic*, $H(z) = z$, and corresponds to the slowly modulated pure wave introduced in chapter I, section 1.3, which has a circular orbit: $z(t) = e^{i\phi(t)}$. In general, $H(z)$ is not a linear function; for the voice signal shown in Fig.III.1, it encodes its timbre, here the vowel "A" that is sung. The corresponding orbit, although non-circular, but it can be decomposed from a Taylor series as a sum of harmonic (circular) components:

$$H(z) = \sum_{n=1}^{\infty} c_n z^n \quad , \tag{III.2}$$

also called *harmonics* of integer orders n, with frequency $f_n = nf_1$. The superposition of all harmonic components, with complex coefficient c_n, encodes the waveform that repeats at each period $\tau = f_1^{-1}$. As the period $\tau = \tau(t)$, the waveform can also be slowly modulated, $c_n = c_n(t)$.

The Taylor coefficients are precisely the ones of a Fourier series when the analytic signal $x^+(t)$ is exactly periodic (with no modulation). Obtained from the Fourier transform defined in Eq.(I.1), its one-sided spectrum (for positive frequency) is:

$$\hat{x}(f) = \sum_{n=1}^{\infty} c_n \delta(f - nf_1) \quad . \tag{III.3}$$

Each spectral component is concentrated on a harmonic frequency, integer multiple of the *fundamental* frequency $f_1 = \tau^{-1}$. For instance, it is clear from the voice spectrum in Fig.III.1 (C), that the fundamental frequency (the voice pitch) is slightly above 250 Hz, and the 4^{th}

harmonics is the one with the highest amplitude. In practice, the harmonic peaks are not as concentrated as the ideal Dirac deltas, since the signal is finite. They can also be broadened by modulations of the oscillations, which are more difficult to recognize from the spectrum.

This combination of spectral and temporal information is put into evidence in a time-frequency representation such as the Gabor and wavelet transforms.

1.2 Quality factor, a cursor separating waves from beats

As a continuum of band-limited analytic signals, the wavelet transform $X(t,f;Q)$ defined in Eq.(II.1) naturally extends and combines the above spectral and temporal approaches. At any fixed frequency, the signal is analytic and contained in a frequency band of width $\delta \log f$; and at any fixed time, the spectrum characterizes a temporal extent of $f\delta t$ oscillations (whatever the considered frequency).

Introduced in Eq.(I.41), the practical time and frequency resolutions of $X(t,f;Q)$ are expressed in terms of the quality factor Q as:

$$f\delta t = \frac{Q}{\sqrt{\curlyvee}} \quad ; \quad \delta \log f = \frac{\sqrt{\curlyvee}}{Q} \quad , \tag{III.4}$$

for the norm 1 (amplitude). In Fig.III.2, we compare a small and a high value of the quality factor Q. The frequency resolution is so low for $Q = 5$ (A, B) that we cannot distinguish the harmonic structure of the spectrum, except for its fundamental component ($f_1 \approx 256$ Hz) and perhaps the first harmonics, in contrast to $Q = 25$ (C, D) where harmonics are numerous. Instead, the high time resolution at $Q = 5$ let us distinguish evenly spaced vertical structures indicating the regular repetition of the waveform.

Based on the practical frequency resolution, the order n harmonics can be distinguished from the higher orders for at most $\delta \log f = \log \frac{n+1}{n}$. Therefore, at least $Q \sim \frac{\sqrt{\curlyvee}}{\log 2} \approx 3.6$ is required to separate the fundamental ($n = 1$) component from the harmonics $n = 2$, as observed at $Q = 5$ in panels (A, B), but the second gets separated from the third from $Q \sim \frac{\sqrt{\curlyvee}}{\log 3/2} \approx 6.2$ onwards.

A similar reasoning from the time domain goes as follows: the wavelet should oscillate at least $f\delta t = n$ times to resolve the order n harmonics of the waveform, so that we start to resolve the second harmonics at about $Q \sim 2\sqrt{\curlyvee} \approx 5$, and the tenth harmonics at $Q \approx 25$. Unresolved "harmonics" interfere, resulting in a regular beating at the fundamental frequency, as confirmed in both illustrated cases (A, C).

These spectral and temporal perspectives are consistently interlaced from the following harmonic ordering:

$$n \; < \; \left(\log \frac{n+1}{n} \right)^{-1} \; < \; n+1 \quad , \qquad \forall n > 0 \tag{III.5}$$

as verified by $(\log \frac{n+1}{n})^{-1} \sim n + \frac{1}{2} - \frac{1}{12n} + ...$ This relies on the fact that the practical resolutions Eq.(III.4) are inverses of each other, so that they form a unit area of time-frequency uncertainty. The relevance of the factor $\sqrt{\curlyvee}$ is confirmed by observations, which discards any other option in Table I.1 for this application. The number $n = \frac{Q}{\sqrt{\curlyvee}}$ can be considered as an effective order below which harmonic components are resolved as horizontal lines (waves) and above which they interfere, forming vertical lines (beats). Put another way, for $n < \frac{Q}{\sqrt{\curlyvee}}$, two successive harmonic components are considered additively, as in $\cos(\curlyvee n f t) + \cos(\curlyvee (n+1) f t)$, while for $n > \frac{Q}{\sqrt{\curlyvee}}$, the beating phenomenon occurs and they are considered as a single modulated component, $2\cos(\pi f t)\cos(\curlyvee (n+\frac{1}{2}) f t)$, in the wavelet transform.

It is useful to recall the key differences with the Gabor transform in this context of cyclic dynamics. Any oscillation whose waveform is shorter than the Gabor window is decomposed by

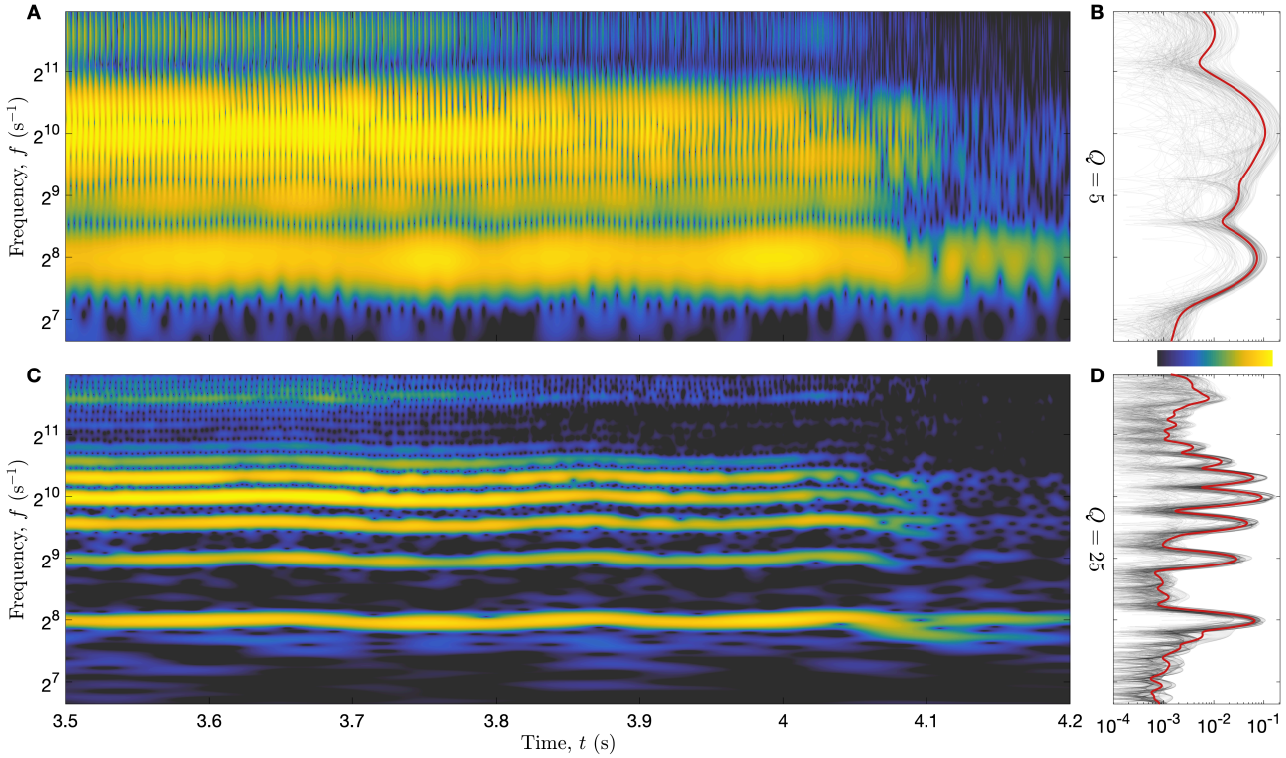


Figure III.2: *Time-frequency representation of the voice recording: "A" vowel sung by a female singer. (A, C) Wavelet transform for $Q = 5$ and $Q = 25$, the colour codes for the amplitude (twice the modulus). (B, D) Amplitude wavelet spectra, for each time in the selected interval (light grey lines) and root-mean-squared value (thick red line). The colour scale is aligned with the amplitude of the spectra, in signal unit (arbitrary).*

the transform into harmonic horizontal structures, without restriction on the harmonic order. Only modulations that are slower than the window are resolved in time. When the Gabor window does not last long enough to analyse the full waveform, then all harmonics are lost by lack of frequency resolution, and the transform has rhythmic (beating or vertical) structures only. Therefore, the form of the Gabor transform crucially depends on how the window's duration T compare to the oscillation period τ_1 and modulation timescale τ_2: the spectral interpretation requires $T \gg \tau_1$ and the temporal one $T \ll \tau_2$.

With the wavelet transform, the harmonic and rhythmic behaviours are simultaneously present for any oscillating signal, whatever its period: the transform is scale-invariant. The time-frequency trade-off is controlled by the quality factor Q: the time-resolution improves with the harmonic order, limited to about $\frac{Q}{\sqrt{7}}$, beyond which the frequency resolution is insufficient to separate harmonic lines, instead producing beats.

Summary of section 1

Oscillations in a voice recording are slowly modulated, with a quite steady waveform. This natural rhythmic oscillation has a cyclic dynamics, $x^+(t)$ has a time period $\tau = f_1^{-1}$, and a harmonic structure of the spectrum $\hat{x}(f)$, with peaks at frequencies $f_1, 2f_1, 3f_1, ...$ These features are transferred in the time-frequency representation as horizontal lines —the fundamental and harmonic components— up to a certain order n. At frequencies $f > nf_1$, the wavelet duration δt is shorter than the period, and successive harmonics are indistinguishable by the bandwidth $\delta \log f$, yielding a beating phenomenon with the same period τ as the waveform. The definition of both practical time and frequency resolutions proves consistent for the prediction of the threshold order: $n = \frac{Q}{\sqrt{\pi}}$. The quality factor of the wavelet analysis determines the number of distinct harmonic components that can be observed for a rhythmic oscillation, as well as the time resolution of their modulations.

2 Time derivative, phase ridges, and oscillating components

The wavelet transform with respect to the log-normal wavelet inherits the same exceptional regularity properties as its wavelet, so that we can differentiate it almost everywhere, even for a stochastic trajectory. In particular, the derivative of the phase can serve to characterize precisely the instantaneous frequencies in the voice recording, see Fig.III.3 (A, B).

Furthermore, the log-normal wavelet has the very convenient property of being invariant under time-differentiation [Altes 1976], whereby most other wavelets would change their shape. As shown in Eq.(II.14), it takes the form:

$$\dot{X}(t, f; Q) = \partial_t X(t, f; Q) = i\Upsilon f \sqrt{\lambda} X(t, \lambda f; Q) \quad , \tag{III.6}$$

where the factor $i\Upsilon f$ is the Fourier representation of the time-derivative, and a small dilation factor $\lambda = e^{\frac{1}{Q^2}}$ appears.

2.1 Instantaneous rate and frequency

Starting with a real signal $x(t)$, the concept of instantaneous rate arises from the interpretation of its analytic version $x^+(t)$ as a modulated complex wave $\frac{A(t)}{2} e^{i\phi(t)}$. This particular form is an accurate model when the signal is a slowly modulated harmonic wave. Its instantaneous frequency corresponds to $\dot{\phi}(t)/\Upsilon$, given by the real part of the following expression:

$$F_x(t) = \frac{\dot{x}^+(t)}{i\Upsilon x^+(t)} \quad . \tag{III.7}$$

For the slowly modulated harmonic wave, $F_x(t)$ is mostly real, with a small imaginary part related to the rate of modulation of the magnitude $m(t) = \log A(t)$:

$$\partial_t \log |x^+(t)| = \Re \left\{ \frac{\dot{x}^+(t)}{x^+(t)} \right\} = -\Im\{\Upsilon F_x(t)\} \quad . \tag{III.8}$$

In contrast, for a real signal (of instantaneous power $P_x(t) = 2|x^+(t)|^2$) which is wide-band (such as an irregular stochastic trajectory) or contains several components, $F_x(t)$ can be much more complex and difficult to interpret. In the general case of a complex signal, the instantaneous rates from the analytic and anti-analytic signals $x^\pm(t)$ can be distinguished: $F_x^\pm(t)$.

The expression of the instantaneous rate for the signal $x(t)$ naturally extends, with its wavelet transform, to:

$$F_x(t,f;Q) = \frac{\dot{X}(t,f;Q)}{i\Upsilon X(t,f;Q)} \quad , \tag{III.9}$$

introducing, through the quality factor Q, a frequency localization f (scale variable). Thus, the modulations of several oscillating components, or rhythms, at different scales can be distinguished in the time-frequency plane. The result is no more a signal $F_x(t)$ but a time-frequency field $F_x(t,f;Q)$ containing all rates associated to each component.

It is instructive to examine its asymptotic behaviour with respect to the quality factor Q, obtained from the relations Eqs.(II.8, II.10):

$$F_x(t,f;Q\to 0) = F_x(t) \qquad ; \qquad F_x(t,f;Q\to\infty) = f \quad . \tag{III.10}$$

The fully temporal limit Eq.(III.7) is retrieved at a small quality factor (f-localization is lost) and the fully spectral limit reduces to the scale variable f that is examined. This also means that at a high enough quality factor, a noise $x(t)$ examined at the scale f can be interpreted as a slowly harmonic wave. We can verify that the instantaneous rate $F_x(t,f;Q)$ of the noise is mostly real, even when the wavelet has a few oscillations.

In practice, the instantaneous frequency is conveniently computed from Eq.(III.6) as:

$$F_x(t,f;Q) = \sqrt{\lambda} f \frac{X(t,\lambda f;Q)}{X(t,f;Q)} \quad . \tag{III.11}$$

We can visualize it in comparison to f for the voice recording in Fig.III.3 (D). Near singular points described in chapter II, section 1.1.3, its behaviour is deduced easily: let (t_0,f_0) denote the location of such a phase vortex, characterized by a zero amplitude, $X(t_0,f_0;Q)=0$. Since the vortex is of unit charge, the amplitude vanishes linearly with the distance to (t_0,f_0). At this location, F_x diverges because of its denominator, and it vanishes at $(t_0,f_0/\lambda)$ because of its numerator. Singular points are thus easily identified in Fig.III.3 (D) from their dipolar behaviour with vertical polarity.

2.2 Oscillating components as horizontal time-frequency ridges

A first approach consists in focusing on time-frequency lines of interest, $\{t_k(s), f_k(s)\}$, as representations of components in the signal (parametrized by s, indexed by k). These lines, called ridges, can be defined in different ways from partial time and frequency derivatives of the Gabor or the wavelet transforms [Carmona, Hwang, and Torrésani 1997; Delprat et al. 1992; Guillemain and Kronland-Martinet 1996], depending on the nature of the analysed signal. Suitable for the representation of an oscillating component, a horizontal ridge can be written $f(t)$, while a vertical ridge $t(f)$ represents a singular event [Guillemain and Kronland-Martinet 1992]. We present below elementary definitions of horizontal ridges as a way to characterize the instantaneous frequency of oscillating components in the signal from its wavelet transform.

One definition consists in considering only locations for which instantaneous frequency and frequency scale agree by solving:

$$\Re\{F_x(t,f;Q)\} = f \quad , \tag{III.12}$$

as expected for a consistent description of oscillating components. The many lines that satisfy this first condition are called phase ridges [Lilly and Olhede 2010], in contrast to amplitude ridges, defined by an extremal modulus: $\partial_f |X(t,f;Q)| = 0$. These different approaches are illustrated in Fig.III.3 (C, D).

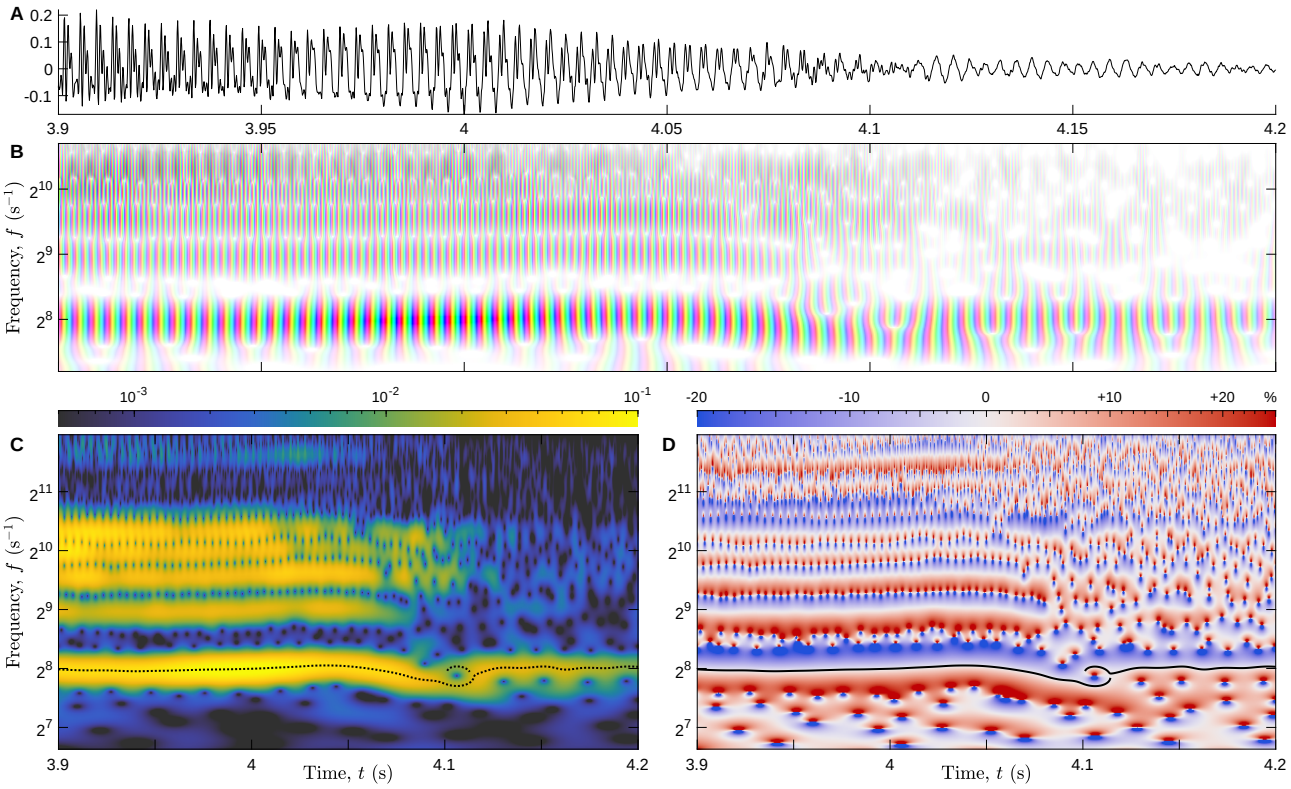


Figure III.3: *Phase and amplitude in the time-frequency plane of the voice recording. (A) Signal $x(t)$, and (B) its wavelet transform $X(t, f; Q)$, of quality factor $Q = 10$. Both the phase and the amplitude are represented, as the angle associated to a hue in the chromatic circle, and the modulus coded by the saturation of the colour. (C) Amplitude only, $2|X(t, f; Q)|$, on a logarithmic scale in signal's unit (arbitrary). (D) Instantaneous frequency, $\Re\{F_x(t, f; Q)\}$, expressed in percentage above or below the frequency variable f. Stable ridges of the fundamental voice component: amplitude ridges as dotted black lines in (C) and phase ridges as plain black lines in (D), in two fragments.*

A second condition is used to identify oscillating components among these lines. For amplitude ridges, it consists in retaining only local maxima of the amplitude, $\partial_f^2 |X(t, f; Q)| < 0$, while for phase ridges, this corresponds to the condition:

$$\partial_f \Re\{F_x(t, f; Q)\} < 1 \quad . \tag{III.13}$$

This relies on the observation of a pure wave, whose instantaneous frequency does not vary in scales (near its ridge). For this reason, we refer to ridges satisfying the second condition as stable ones, and others as unstable ridges. For instance, the fundamental component of the voice corresponds to a stable ridge, plotted in Fig.III.3 in a black plain line from the phase (D) and in black dotted line from the amplitude, with minor difference. All types of ridges are shown in Fig.(III.4) (A), with unstable ridges in white.

Phase and amplitude ridges are observed to be quasi undistinguishable. That is explained for stable ridges because the frequency variable f, equated to the instantaneous frequency for phase ridges, Eq.(III.12), is the peak frequency of the wavelet, that maximizes the amplitude of a pure wave of horizontal ridge. Both for the phase and the amplitude approaches, the unstable ridges span phase vortices / zero of the amplitude, so that they are also very close. The extraction of phase ridges is interrupted at the singular points because of phase indeterminacy, while amplitude ridges are not. Apart from that, the only visible differences can be spotted when these mostly horizontal lines are getting close to vertical, see details of Fig.(III.4) (A).

Physical recording such as the voice recording are subject to some background noise of

instrumental origin. Contrary to the pedagogical signal in Fig.I.1, these signals are not smooth, at least because of this stochastic component of non-zero power density. For this reason, ridges and singularities are observed everywhere in their time-frequency representation (contrary to Fig.I.3). From Fig.(III.4) (A), we can make the following simple observations: ridges are closed loops, or join borders of the time-frequency domain. Stable and unstable ridges meet at vertical tangent points, out of which they can be parametrized by time, $\{t, f_k(t)\}$. Singularities are located on unstable ridges. At any time, a pair of stable ridges $f_k(t) < f_l(t)$ are necessarily separated by at least one simultaneous unstable ridge (and vice versa). Long stable ridges are associated to well resolved regular oscillations, such as the fundamental component of the voice, near $2^8 = 256$ Hz in Fig.III.3 (C, D). Numerous in Fig.(III.4) (A), short ridges can be related to interfering (non-separated) components, such as harmonics of high order, or to the low power background noise, that remains when the voice stops after $t = 4$ s.

Stable ridges allow us to extract corresponding locally harmonic components which are modulated, in phase and amplitude, slowly enough compared to the wavelet size. The oscillating component associated to a stable ridge $f_k(t)$ can be reconstructed by evaluation of the wavelet transform along the ridge [Delprat et al. 1992; Lilly and Olhede 2010]:

$$x_k(t) = 2\Re\{X(t, f_k(t); Q)\} \quad , \tag{III.14}$$

so that $x(t) \approx \sum_k x_k(t)$ up to a residue expected to be small or the signal's trend [Chui and Mhaskar 2016].

This idea of summarizing rhythms as time-frequency lines $f_k(t)$ gets complicated by the presence of an important noise compared to the oscillation amplitude, causing singular points to distorts or interrupts the ridge of the oscillatory component. In the illustration, this only happens when the voice is evanescent near $t = 4.1$ s. In this context of a low signal-to-noise ratio, the extraction and reconstruction of a single oscillatory component as a continuous ridge has been achieved using a variational approach, by enforcing continuity and optimizing the trade-off between the regularity of the ridge and the approximate ridge conditions [Carmona, Hwang, and Torrésani 1997].

Fragmentation also occurs when oscillating components of close frequencies are not separated; its periodic occurrence is related to the beating phenomenon. For this reason, the fragmentation of the ridge highly depends on the quality factor. This is illustrated by harmonic components of order $n > 4$ of the voice recording in Fig.(III.4) (A), due to the quality factor $Q = 10$, i.e. a wavelet duration $f\delta t = \frac{Q}{\sqrt{\pi}} \approx 4$.

2.3 Time-frequency concentration from reassignment

Closely related to the ridge approach, the reassignment procedure can also concentrate the representation of oscillatory components along time-frequency lines. It consists in replacing the original time and frequency variables of the representation for new ones [Auger, Flandrin, et al. 2013; Flandrin, Auger, and Chassande-Mottin 2018]. In particular, the frequency-reassignment of the analytic wavelet transform changes the scale (frequency) variable for the instantaneous frequency Eq.(III.9):

$$(t, f) \quad \leftarrow \quad (t, \Re\{F_x(t, f; Q)\}) \quad . \tag{III.15}$$

By definition, this procedure does not affect phase ridges Eq.(III.12). However, neighbouring locations are attracted towards stable phase ridges and repelled from unstable ones during the reassignment. Let us denote $\mathcal{R}$ the frequency-reassignment operator. Its actions upon the

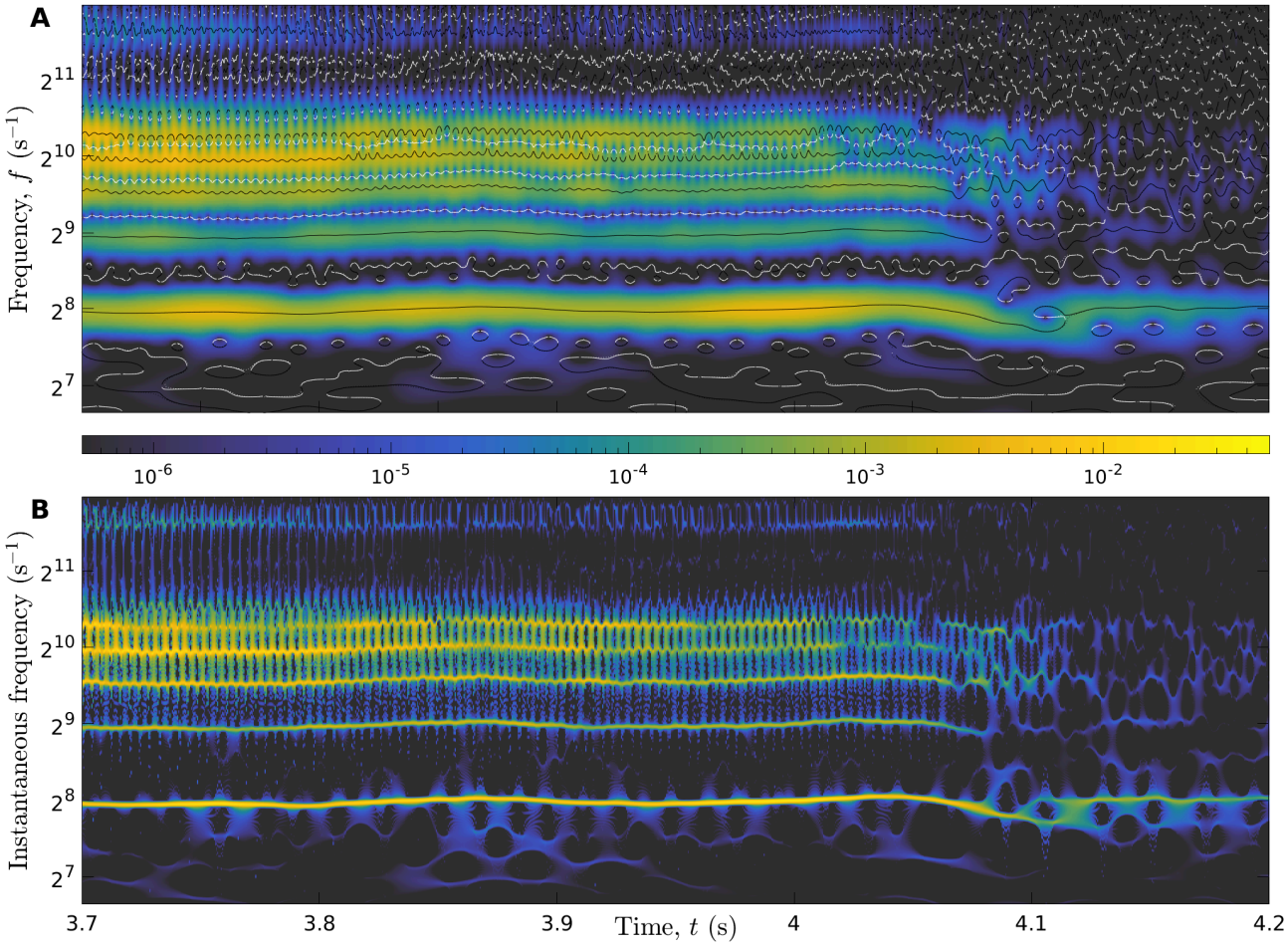


Figure III.4: *Ridges and frequency reassignment of the voice signal. (A) Stable ridges (black lines) and unstable ones (white lines), superimposed to the wavelet transform squared modulus (scalogram) for $Q = 10$. Phase ridges (plain lines) are hardly distinguishable from amplitude ridges (dotted lines). (B) Frequency-reassigned scalogram based on the change of variable Eq.*(III.15)*. The alternative reassignment Eq.*(III.20) *is almost identical.*

wavelet transform and its square (the scalogram) can be expressed explicitly as:

$$\mathcal{R}[X](t,f;Q) = \int X(f',t;Q)\delta(f - \Re\{F_x(t,f';Q)\})|f|\frac{\mathrm{d}f'}{|f'|} \tag{III.16}$$

$$\mathcal{R}[|X|^2](t,f;Q) = \int |X(f',t;Q)|^2\delta(f - \Re\{F_x(t,f';Q)\})|f|\frac{\mathrm{d}f'}{|f'|} \quad , \tag{III.17}$$

where the Dirac delta handles the change of evaluated variable. The first operation has been called synchrosqueezing [Daubechies, Lu, and Wu 2011; Daubechies and Maes 1996]. In contrast but not in contradiction with the original definitions [Auger and Flandrin 1994; Auger and Flandrin 1995], notice the occurrence of $|f|$ and $\frac{\mathrm{d}f'}{|f'|}$. This ensures consistency with the choice of the convention $p = 1$, instead of $p = 2$, in the definition Eq.(I.6) of the wavelet transform, i.e. the same behaviour under frequency integration:

$$\int \mathcal{R}[X](t,f;Q)\frac{\mathrm{d}f}{|f|} = \int X(t,f;Q)\frac{\mathrm{d}f}{|f|} = \frac{\sqrt{\Upsilon}}{Q}x(t) \quad , \tag{III.18}$$

$$\int \mathcal{R}[|X|^2](t,f;Q)\frac{\mathrm{d}f}{|f|} = \int |X(t,f;Q)|^2\frac{\mathrm{d}f}{|f|} = \frac{\sqrt{\pi}}{Q}P_x(t;Q) \quad , \tag{III.19}$$

to retrieve the signal or its time-dependent power $P_x(t;Q) = P_{xx}(t;Q)$ defined in Eq.(II.21).

The resulting reassigned representations are much more concentrated than the original ones, since the time-frequency contributions accumulate near the stable phase ridges. This is illustrated in Fig.III.4 (B). Visually, oscillating components are obtained as time-frequency lines, ideally discrete for slowly modulated and distinct harmonic components, or spread into a less local density under the influence of a strong noise level or a loss of regularity in the oscillation.

Since $\Re\{F_x\} \approx F_x \approx |F_x|$ for the voice harmonics, or any other components when the quality factor is high enough, this interpretation does not vary when reassigning as:

$$(t,f) \quad \leftarrow \quad (t, |F_x(t,f;Q)|) \quad , \tag{III.20}$$

i.e. using $\delta(\log f |F_x(t,f';Q)|^{-1})$ instead of $\delta(f - \Re\{F_x(t,f';Q)\})|f|$ in Eqs.(III.16, III.17). When applied to the voice recording with $Q = 10$, we cannot distinguish the result from Fig.III.4 (B).

Even though it helps to clarify the time-frequency representation of rhythmic signals, the reassignment is not in itself a technique for extracting oscillatory components. It has been used as a preliminary step to the extraction of ridges of maximal amplitude [Su and Wu 2017].

The phase in the wavelet transform seems to give access to sub-resolution information. The precision in frequency of the reassigned representation would only be available to the wavelet transform at a very high quality factor. We can wonder whether it bypasses the Heisenberg uncertainty principle? In our understanding, the time-frequency uncertainty is not modified, even though the frequency resolution appears entirely "squeezed". Instead, the frequency uncertainty is made implicit, rather than explicitly assimilated to the width of frequency peaks in $|X(t,f;Q)|^2$. Indeed, the ability to distinguish continuously a pair of harmonic components with close frequencies does not improve: it is still determined by the quality factor of the wavelet, subject to uncertainty relations (see section 3.2.3 and Appendix 4).

2.4 Horizontal-vertical correspondence: from rhythms to fractals

All the above concepts still apply when swapping the time and frequency variables, with few to no changes except their interpretation: regular oscillations forming the horizontal lines of a rhythm turn into vertical lines pointing to singularities of a fractal.

Instead of obtaining an instantaneous frequency from the time derivative, we get a group delay from the frequency derivative:

$$T_x(t,f;Q) = \frac{\partial_f X(t,f;Q)}{i\Upsilon X(t,f;Q)} = \frac{1}{i\Upsilon}\partial_f \log X(t,f;Q) \quad . \tag{III.21}$$

Vertical phase ridges are obtained at the vanishing delay $\Re\{T_x(t,f;Q)\} = 0$, or alternatively, amplitude ridges correspond to $\partial_t |X(t,f;Q)| = 0$ for each frequency. The stable ones are $\partial_t \Re\{T_x(t,f;Q)\} > 0$ for phase ridges, and $\partial_t^2 |X(t,f;Q)| < 0$ (local maxima) for amplitude ridges. Eventually, the delay can serve to reassign time: $(t,f) \leftarrow (t - \Re\{T_x(t,f;Q)\}, f)$.

We recall that there is no simple relation between time and frequency partial derivatives in the log-normal wavelet framework, as the frequency-derivative accounts for a change of wavelet, see chapter II, section 1.2.5. We argue that a very close behaviour can be achieved with the log-normal wavelet, by taking advantage of the small rescaling in the time derivative, Eq.(III.6), to define a finite differentiation (instead of infinitesimal) in the log-frequency domain:

$$\begin{aligned} \frac{\log X(t,\lambda f;Q) - \log X(t,f;Q)}{\log \lambda} &= Q^2 \log \frac{\partial_t X(t,f;Q)}{i\Upsilon \lambda^{\frac{1}{2}} f X(t,f;Q)} \\ &= Q^2 \log(F_x(t,f;Q)/f) - \tfrac{1}{2} \approx i\Upsilon f T_x(t,f;Q) \quad , \end{aligned} \tag{III.22}$$

where we have identified and replaced the complex rate Eq.(III.9). The approximate equality in the last line, that can be written $F_x \approx f\lambda^{\frac{1}{2}+i\Upsilon f T_x}$, is as precise as $\lambda = e^{\frac{1}{Q^2}}$ is close to 1, for Q large enough. The vertical ridge condition $\Re\{T_x(t,f;Q)\} = 0$ translates into:

$$\Im\{\log F_x(t,f;Q)\} = 0 \quad , \tag{III.23}$$

providing an interpretation of the angle (complex argument) of $F_x(t,f;Q)$ as a delay.

Instead of oscillatory components, stable vertical ridges $t_k(f)$ focus on singular events (indexed by k) across scales, and the time-reassignment can help to locate them in time. The value of the wavelet transform on vertical ridges, called the wavelet skeleton [Muzy, Bacry, and Arneodo 1994], allows to extract the scaling behaviours of these singularities: $|X(t_k(f),f;Q)| \sim |f|^{-h_k}$. For the self-similar Gaussian processes introduced in chapter II, section 3.3, $h_k = H$ is the Hurst regularity exponent. It can be directly estimated as minus the real part of Eq.(III.23), although this estimator fluctuates a lot around the central value H when we vary the time-frequency location (approaching singular points of zero amplitude).

These noise have a single regularity exponent and are called monofractal. However, intermittent signals, such as turbulent velocity [Arneodo, Bacry, Manneville, et al. 1998] or rainfall noise [Venugopal et al. 2006], may exhibit multiple scaling, hence called multifractal. They were originally studied using real wavelets [Argoul, Arneodo, Elezgaray, et al. 1989], that broadly applies to higher-dimensional fractal objects from dendritic aggregates [Argoul, Arneodo, Elezgaray, et al. 1990; Argoul, Arneodo, Grasseau, et al. 1989] to galaxies (such as our Milky Way) and breast tissues [Gerasimova et al. 2014; Gerasimova-Chechkina et al. 2016; Khalil et al. 2006].

The role of the wavelet skeleton, with vertical ridges from the maximum modulus, is twofold in the context of fractals. First it reduces possibly high-dimensional data to a set of essential lines of interest able to localize singularities. Then it allows to reliably apply the multifractal formalism to their wavelet transform by avoiding its zeros, sources of instability. This formalism has a thermodynamical interpretation [Arneodo, Audit, et al. 1997; Arneodo, Bacry, and Muzy 1995; Muzy, Bacry, and Arneodo 1991], whose partition function may be expressed with our notations as:

$$\mathcal{Z}_x(f,q) = \langle |X_\psi(t,f)|^q \rangle_t \quad , \tag{III.24}$$

where the order n_ψ of the wavelet ψ is sufficiently high (see chapter I, section 3.3.2), the q-norm exponent plays the role of an "inverse temperature" and $f \to \infty$ corresponds to the macroscopic limit in statistical mechanics. On zeros of the wavelet transform, numerical instabilities arise when $q < 0$, turning into a mathematical divergence of the partition function when $q \leq -1$. In the thermodynamical limit, the "free energy" is proportional to a (generalized) fractal dimension, related through the Legendre transform to an "entropy" called the singularity spectrum. It measures the set of locations in the signal with a given scaling or regularity exponent ("energy"); there is a single one for a monofractal or a broad continuum of scaling exponents for a multifractal.

Nothing prevents from using analytic wavelets for a multifractal analysis, quite the opposite as ψ_Q benefits from an infinite order Eq.(I.40). This would even be a natural development beyond the quadratic and spectral analysis presented in chapter II.

We mention here an important difference between the analysis of a real signal with a real or with a complex (analytic) wavelets: zeros of the wavelet transform are lines in the real case and points in the second. These singularities are the phase vortices, at the intersection of vertical and horizontal unstable ridges. For the white Gaussian noise, they form a point process of repelling particles, that has been recently characterized [Bardenet, Flamant, and Chainais 2020; Bardenet and Hardy 2021] for different transforms (with the Cauchy wavelet and the Gaussian

Gabor window). The analysis of this "time-frequency gas", is a complementary viewpoint to the one of stable ridges [Flandrin 2015; Koliander et al. 2019].

3 Estimating rate signals

When studying a rhythmic system whose characteristic frequency depends on time, we have just highlighted how the dynamical quantity of interest, $f(t)$, is embedded in the time-frequency representation of the rhythmic signal. It is not clear yet how to extract it as a stand-alone signal, whose dynamics could be analysed with the previously developed tools.

A direct ridge approach suggests to start with an exhaustive identification of time-frequency locations (t, f) satisfying the ridge conditions, Eqs.(III.12, III.13) for the phase ridge. This requires a sufficiently precise time-frequency sampling, improved to some extent by interpolation, which can be computationally intensive for long signals. Then, contiguous locations (t, f) are connected into times series, yielding a set of ridge signals $f_k(t)$, in particular stable ones, some of which represent part or (hopefully) the totality of the oscillating component of interest. Therefore, relevant fragments remain to be sorted, using various criteria about the amplitude of the oscillation, its time support or its frequency range. While the interruption of the ridges is expected when the oscillation is lost, the conditions for this interruption is not specified in this approach.

The reassignment approach is not able in itself to extract the rate signal. But we learn from it that the information on the modulations of an oscillating component is not only located on its stable ridge $f(t)$, but also contained in the surrounding region of attraction delimited by phase vortices, i.e. by the nearest unstable horizontal ridges. The size of this band is determined by the wavelet band-width, and widened when the signal-to-noise ratio increases.

In the following, let us assume we can specify an approximate frequency band, that contains the oscillating component of interest. We present different rate signals, estimated by integrating in this band either a linear or a quadratic expression. Applied to the voice signal, they are represented in Fig.III.5 in comparison to the frequency-reassigned power density. The narrow-band estimation centred on the fundamental voice component provides the pitch of the voice (C, D). In contrast, the wide-band estimation (A, B) takes into account all harmonics, which results in a voice modulation that is distinct from the pitch. This modulation is representative of the relative change of amplitude among the voice harmonics, also called vibrato.

3.1 Narrow to wide band estimations

Based of Eqs.(III.7, III.9), we can define a variety of rate signals estimating the instantaneous frequency of a rhythmic signal, whose oscillating component of interest is contained in a band $B = [f_{\min}\,,\, f_{\max}]$. For simplicity, this band is kept fixed in time, which is sufficient for our illustration. Based on an integration in this band, we first detail a "linear" approach (where linear refers to the integrated expression). To improve its limitations, a second "quadratic" approach is then introduced. Both agree on the narrow-band limit $f_B = f_{\min} = f_{\max}$, which is Eq.(III.9) with a fixed f, but their estimations in the wide-band limit differ.

3.1.1 Linear estimator

The first estimator is based on the signal's instantaneous frequency $F_x(t)$ from Eq.(III.7), but utilizes the analytic signal partially reconstructed from the band B rather than the complete

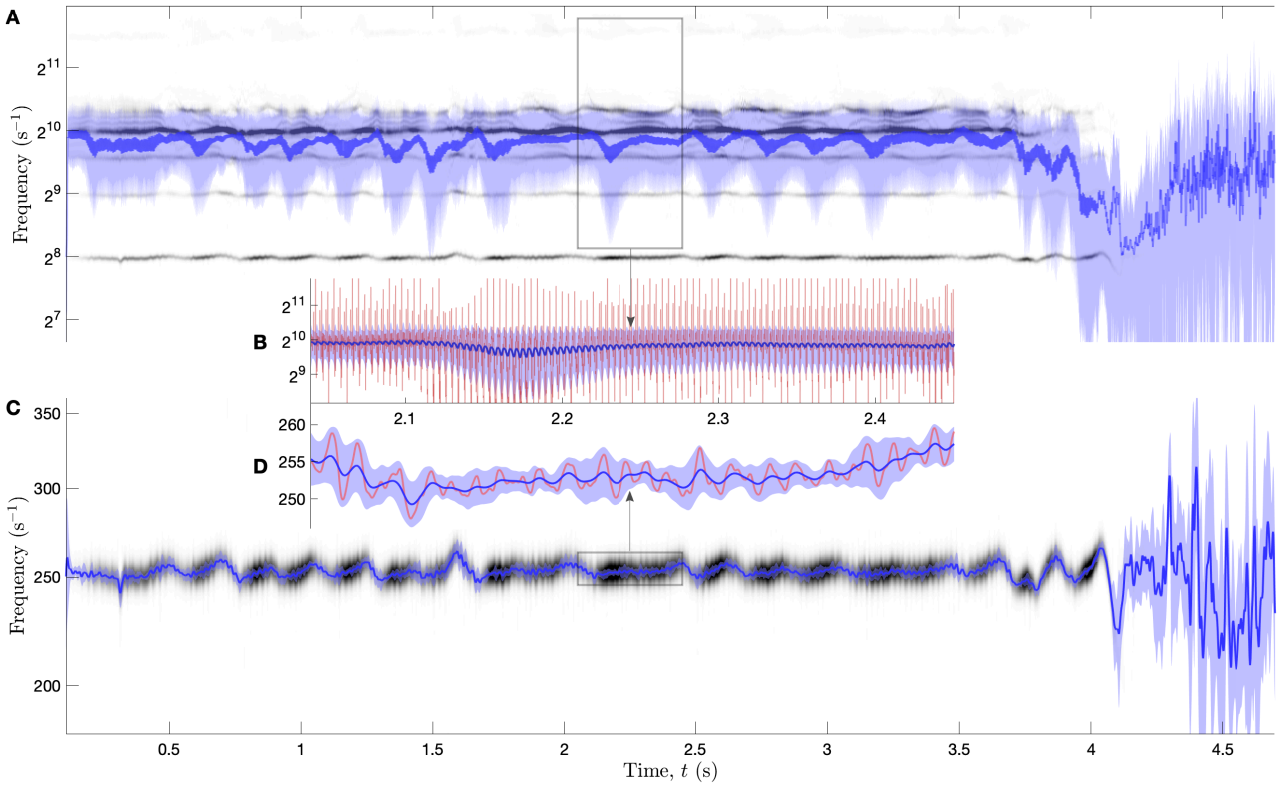


Figure III.5: *Wide and narrow band estimations of the time-dependent rate of the full voice recording. The rate signals (coloured) are plotted above the frequency-reassigned power density (grey scale in panels A and C). The band B of integration, wide in (A) and narrow in (C), is the apparent frequency range. (B, D) Zoom of (A, C). The quadratic estimations Eq.*(III.27) *(blue lines) are surrounded by the corresponding uncertainty $\pm\Delta F_x$ (light blue area). The linear estimations Eq.*(III.25) *(red lines), much more oscillatory for a wide band, are shown in zooms (B, C) only. The narrow band (C) focuses on the fundamental component of the voice (its pitch). The wide band (A) averages over all harmonics: it is sensitive to relative variations of their amplitude (such as a vibrato).*

one. Therefore, the instantaneous frequency is estimated as the real part of:

$$f_1(t;B) = \frac{\int_B \dot{X}(t,f;Q)\mathrm{d}\log f}{i\Upsilon\int_B X(t,f;Q)\mathrm{d}\log f} \quad . \tag{III.25}$$

Eq.(III.7) is recovered in the wide-band limit, $F_x(t) = f_1(t;\mathbb{R}^+)$, whereas the narrow-band limit corresponds to Eq.(III.9): $f_1(t;\{f_B\}) = F_x(t,f_B;Q)$.

In fact, the case of an intermediary band is very similar, in practice, to choosing:

$$F_x(t,f_B;Q_-) \approx f_1(t;B) \quad , \tag{III.26}$$

where $f_B = \sqrt{f_{\max} f_{\min}}$ is the central frequency of the band, and the quality factor $Q_- < Q$ has been decreased so that the frequency resolution $\delta\log f$ matches the width of B: $\frac{\sqrt{\Upsilon}}{Q} \approx \log\frac{f_{\max}}{f_{\min}}$. Indeed, the band integration has the effect of increasing the wavelet band-width $\delta\log f$; see the change of quality factor in Eq.(II.12), where the kernel plays the role of the band B with smooth borders instead of sharp ones.

This linear approach of a band-estimation boils down to adjusting the parameters f and Q in Eq.(III.9), and the condition for a successful estimation has been described previously: the fixed values $f = f_B$ should remain within the region of attraction near the actual phase ridge $f(t)$. The decrease of the quality factor Q can help to widen this region to a certain

extent, limited by the influence of other simultaneous noisy or oscillating components in the signal. Otherwise, the beating phenomenon occurs between these components, and the rate signal oscillates accordingly. The estimation is inaccurate when the line f_B leaves the region of attraction of the ridge $f(t)$, and unstable when it hits a phase vortex.

Some of these points can be observed in the example of the voice signal, in Fig.III.5 (C, D), where the estimated rate signal Eq.(III.25) is plotted as a red line. The narrow-band estimation (D), $\sim F_x(t, f_B; Q)$, follows closely the fundamental voice component, which is well resolved and separated from other components. It yields the pitch frequency. On the contrary, the wide-band estimation, $\sim F_x(t; Q)$, oscillates wildly and at a much higher frequency, because of the numerous harmonics. Their interference produces a beating at the pitch frequency.

3.1.2 Quadratic estimator

The second approach addresses this instability and beating issue. It consists in averaging different values of $F_x(t, f; Q)$ in the band of interest, with respect to the weight $|X(t, f; Q)|^2$, in order to favour regions of higher intensity associated to the oscillating component, and mute unstable ones associated to the phase vortices of vanishing amplitude. The weighted average formula in the band B simplifies to:

$$f_2(t; B) = \frac{\int_B \dot{X}(t, f; Q)\overline{X(t, f; Q)}\mathrm{d}\log f}{i\Upsilon \int_B |X(t, f; Q)|^2 \mathrm{d}\log f} \quad . \tag{III.27}$$

The narrow-band (non-averaged) limit reduces again to $f_2(t; \{f\}) = F_x(t, f; Q)$, while the wide-band limit takes the distinct form $f_2(t; \mathbb{R}^+) = F_x(t; Q)$:

$$F_x(t; Q) = \frac{\int_0^\infty \dot{X}(t, f; Q)\overline{X(t, f; Q)}\mathrm{d}\log f}{i\Upsilon \int_0^\infty |X(t, f; Q)|^2 \mathrm{d}\log f} \quad . \tag{III.28}$$

Less instantaneous as $F_x(t)$ in Eq.(III.7) (except in the limit $F_x(t; Q \to 0) = F_x(t)$), this estimator has also the interpretation of a time-dependent frequency for its real part, in particular it is the mean frequency of the reassigned power density (see derivation in Appendix 9), and its imaginary part relates to $\partial_t \log P_x(t, Q)$, as in Eq.(III.8).

This formulation is especially generic since it does not require to specify a band of interest. Whenever several simultaneous oscillating components can be distinguished in the signal x at quality factor Q, the resulting rate estimation is the power-weighted average of their individual rates. This wide-band rate signal $F_x(t; Q)$ is represented as a blue line in Fig.III.5 (A, B) for the voice: it is the average instantaneous frequency of the voice harmonic spectrum. We could expect it to be proportional to the fundamental frequency and thus follow the pitch modulations, as estimated by the narrow-band version, also a blue line in (C, D). Here, this is not the case, because the voice timbre is subtly modulated by a vibrato, changing the relative amplitude of harmonics. In the zoom of panel (B), we notice that it still oscillates at the fundamental frequency, but at a much smaller extent than its linear counterpart (red line). Indeed, the 4 first harmonics are distinct at $Q = 10$, while higher harmonics at not well separated and produce beats. Therefore, $F_x(t; Q)$ combines much information on the fundamental frequency (pitch) and the non-circularity (timbre) of the oscillating signal (the voice).

Finally, the case of an intermediary band can be reformulated similarly to Eq.(III.26) in the linear approach. The corresponding quadratic situation involves two distinct quality factors $Q_- < Q_+$, as exposed in chapter II, section 5.2. Thus, Eq.(III.27) is practically equivalent to:

$$F_x(t, f; Q_+, Q_-) = \frac{S_{\dot{x}x}(t, f; Q_+, Q_-)}{i\Upsilon S_{xx}(t, f; Q_+, Q_-)} \approx f_2(t; B) \quad , \tag{III.29}$$

for $f = f_B$ fixed, Q_- representative of the relative band-width $f_{\max}/f_{\min}$, and $Q_+ = Q$. Both limit cases, $Q_- = Q_+ = Q$ for narrow band and $Q_- \to 0, Q_+ = Q$ for wide band, return to a single quality factor Q. The only difference is a log-normal frequency weighting instead of the band integration. As shown in Table II.1, a convenient approximation is provided by smoothing time over n wavelet durations, with kernel Eq.(II.70). It only requires to compute the wavelet transform $X(t, f; Q)$ at $f = f_B$ and λf_B, see Eq.(III.6), with $Q = Q_-$ ($Q_+ \approx nQ_-$).

The interpretation of the real part as a time-dependent frequency is conserved, and the imaginary part relates to the spectral density rate:

$$\partial_t \log S_{xx}(t, f; Q_+, Q_-) = -2\Im\{\Upsilon F_x(t, f; Q_+, Q_-)\} \tag{III.30}$$

statistical estimator of $\partial_t \log S_{xx}(t, f; Q)$, and generalizing Eq.(III.8).

Therefore, these quadratic estimation formulas for the rate of a rhythmic signal in a certain time-frequency region can be understood as estimators for the following statistical quantities:

$$F_x(t; Q) \approx \frac{\mathbb{E}[\dot{x}^+\overline{x^+}]}{i\Upsilon\mathbb{E}[|x^+|^2]} = \frac{\sigma^2_{\dot{x}^+x^+}(t)}{i\Upsilon\sigma^2_{x^+x^+}(t)} \tag{III.31}$$

$$F_x(t, f; Q_+, Q_-) \approx \frac{\mathbb{E}[\dot{X}\overline{X}]}{i\Upsilon\mathbb{E}[|X|^2]} = \frac{S_{\dot{x}x}(t, f; Q)}{i\Upsilon S_{xx}(t, f; Q)} \quad , \tag{III.32}$$

As compared to their "linear" counterparts $F_x(t)$ and $F_x(t, f; Q)$, they are associated to several number of degrees of freedom, $n_d = \frac{Q_+}{Q_-} > 1$, that brings a statistical interpretation to their increased stability.

3.2 How precise is the rate estimation?

A horizontal ridge that represents an oscillating component in the signal can be interrupted from a lack of regularity or a low signal-to-noise ratio. However, the above analysis always provides an estimation, even for irregular fluctuations. How reliable is the estimation F_x? In the absence of an oscillating component, we expect it to have a large uncertainty, which decreases with the regularity of the oscillation.

To express this uncertainty, we start with Eq.(III.28), in particular with the fact that $\Re\{F_x(t; Q)\}$ is the mean instantaneous frequency of the reassigned scalogram, as shown in Appendix 9. This time-dependent density function over instantaneous frequencies is highly concentrated on the phase ridge $f(t)$ whenever the signal contains a single slowly modulated harmonic wave. From this probabilistic viewpoint, we can quantify whether it estimates a single and well-defined ridge, or not, by computing the variance of this density. The resulting expression is cumbersome, but luckily, we noticed it simplifies considerably shifting a little from the original frequency-reassignment with the real part to the alternative one with the modulus of F_x, see relations Eqs.(III.15, III.20).

For the voice signal, this modification is justified since it is composed of slowly modulated harmonics, so that $F_x(t, f; Q) \approx |F_x(t, f; Q)| \approx \Re\{F_x(t, f; Q)\}$. Actually, this approximate relation is even good for noisy components, because the quality factor, $Q = 10$, is sufficient to impose a slow amplitude modulation compared to the phase. The variance associated to $F_x(t; Q)$ is therefore defined, in the same spirit as $F_x(t; Q)$ in Eq.(III.28), as follows:

$$\Delta F_x(t; Q)^2 = \frac{\int_0^\infty |F_x(t, f; Q) - F_x(t; Q)|^2 |X(t, f; Q)|^2 \mathrm{d}\log f}{\int_0^\infty |X(t, f; Q)|^2 \mathrm{d}\log f} \tag{III.33}$$

$$= \frac{\int_0^\infty |\dot{X}(t, f; Q)|^2 \mathrm{d}\log f}{\Upsilon^2 \int_0^\infty |X(t, f; Q)|^2 \mathrm{d}\log f} - |F_x(t; Q)|^2 \tag{III.34}$$

$$= \frac{P_{\dot{x}^+}(t; Q)}{\Upsilon^2 P_{x^+}(t; Q)} - |F_x(t; Q)|^2 \quad . \tag{III.35}$$

The integrals can be restricted to the band B of interest, and in both limits $Q \to 0, \infty$, the results vanishes for any kind of signal. The relative uncertainty relates the coefficient of variation of F_x to a correlation estimator:

$$\frac{\Delta F_x(t;Q)^2}{|F_x(t;Q)|^2} = |\Gamma_{\dot{x}^+x^+}(t,t;Q)|^{-2} - 1 \quad , \tag{III.36}$$

where the correlation coefficient estimator and its complex version are introduced in Eqs.(II.52, II.53). This means that the relative uncertainty for the estimation $F_x(t;Q)$ of the ridge frequency $f(t)$ is nothing but a rewriting of the correlation estimator between the signal $x(t)$ and its time-derivative $\dot{x}(t)$. Note that we expect $\rho_{\dot{x}^+x^+}(t,t) = i$ for any signal x (harmonic or stochastic), but its estimator does not reach i in general. It only happens for a pure wave, whose derivative is obviously in phase quadrature: $\Gamma_{\dot{x}^+x^+}(t,t;Q) = i$, the corresponding uncertainty of its instantaneous frequency vanishes.

This result is straightforwardly extended to frequency localized rate estimations with the time-frequency coherence:

$$\frac{\Delta F_x(t,f;Q_+,Q_-)^2}{|F_x(t,f;Q_+,Q_-)|^2} = |\gamma_{\dot{x}x}(t,f;Q_+,Q_-)|^{-2} - 1 \quad . \tag{III.37}$$

For the voice signal, the uncertainty is computed inside each band and plotted in all panels of Fig.III.5 as a light blue area, materializing a distance $\pm\Delta F_x$ from the estimated rate signal (blue line). In the selected time interval of panels (C, D), the wide-band estimate has an average uncertainty of 40%, and 1% for the narrow-band estimate. It is the indication that several components are contained in the wide band while only one of them is contained in the narrow band. This uncertainty rises progressively to above 100% and 20% respectively when the voice fades.

Summary of sections 2 and 3

The fluctuating frequency of modulated harmonic components can be estimated at a much greater precision than the wavelet bandwidth, using the time derivative of their cycling phase or instantaneous frequency. For a single component oscillation x, this is the real part of a complex rate signal $F_x(t)$, while the multi-component case requires extracting particular lines $f = f_k(t)$ in its time-frequency version $F_x(t, f; Q)$, horizontal stable ridges. In addition to frequency modulations, the complex rate also captures amplitude modulations in its imaginary part.

Defined from self-consistent conditions on the amplitude or the phase, the extraction of ridges to produce new modulation signals has practical limitations, such that their possible fragmentation. Frequency-reassignment, a change of variable from frequency f to instantaneous frequency, is able to concentrates the amplitude or the intensity of the time-frequency (linear or quadratic) representation on these stable ridges. While this operation is insufficient to extract modulation signals, we observe basin of attractions conferring a certain flexibility in the selection of frequency range of interest.

Assuming a band of interest, we define a generic family of complex rate signals able to estimate the modulation of a rhythmic oscillation in this band. Compliant with the previously developed framework for linear and quadratic time-frequency quantities, the quadratic case has the specificity of a better stability, in relation with a statistical interpretation. Denoted $F_x(t, f; Q_+, Q_-)$, it is both generic and sophisticated enough to estimate a precise and stable rate signal in a certain frequency band, around f of width controlled by Q_-, at the time resolution corresponding to Q_+. While the narrow-band limit returns to a linear quantity, the wide-band limit provides a more naive but especially generic estimator $F_x(t; Q)$, combining in a weighted average the contributions of all resolved components and unresolved beats. The example of the voice recording distinguishes the outcome of the most generic wide-band estimator, capturing its vibrato, from the more sophisticated one, isolating its pitch modulations. The precision of the rate estimations relates to the estimated coherence between the signal and its derivative.

Appendix A

1 Definition of the Gabor / short time Fourier transform

The Gabor transform is defined as follows:

$$\mathcal{G}_w[x](t,f) = \int x(t')\overline{w(t'-t)}e^{-i\Upsilon f t'}\mathrm{d}t' = \int \hat{x}(f')\overline{\hat{w}(f'-f)}e^{i\Upsilon(f'-f)t}\mathrm{d}f' \tag{A.38}$$

also called short time Fourier transform because the (Gabor) window function w, the canonical one being a Gaussian function, localizes in time and frequency the usual Fourier transform (recovered for $w(t) = 1$). The time and frequency variables are here treated symmetrically, up to a phase factor. Compared to the wavelet transform, the fundamental difference of this time-frequency representation lies in the choice of a fixed scale, the time and frequency widths of the function w. The Gabor window is shifted both in the time and frequency domains, contrary to the wavelet window which is only shifted in time and dilated in both domains (making the wavelet transform scale-free).

2 Wigner-Ville perspective on wavelet and Gabor transforms

Let us recall here the definition Eq.(II.27) of the cross Wigner-Ville distribution:

$$W_{xy}(t,f) = \int x(t+\tfrac{\tau}{2})\overline{y(t-\tfrac{\tau}{2})}e^{-i\Upsilon f\tau}\mathrm{d}\tau = \int \hat{x}(f+\tfrac{\eta}{2})\overline{\hat{y}(f-\tfrac{\eta}{2})}e^{i\Upsilon\eta t}\mathrm{d}\eta \quad . \tag{A.39}$$

For a general wavelet transform (Eq.(I.6) with $p = 1$), we prove the relation between Wigner-Ville distributions and the product of wavelet transforms Eq.(II.28):

$$\begin{aligned}
&\iint W_{xy}(t,f)\overline{W_{\psi\varphi}\left(\frac{t-b}{a}, af\right)}\frac{\mathrm{d}t\mathrm{d}f}{|a|} \\
&\quad = \iiiint x(t+\tfrac{\tau}{2})\overline{y(t-\tfrac{\tau}{2})}\,\overline{\psi\left(\tfrac{t-b}{a}+\tfrac{\nu}{2}\right)}\varphi\left(\tfrac{t-b}{a}-\tfrac{\nu}{2}\right)e^{-i\Upsilon f\tau+i\Upsilon a f\nu}\mathrm{d}\tau\mathrm{d}\nu\frac{\mathrm{d}t\mathrm{d}f}{a} \\
&\quad = \iiint x(t+\tfrac{\tau}{2})\overline{y(t-\tfrac{\tau}{2})}\,\overline{\psi\left(\tfrac{t-b}{a}+\tfrac{\nu}{2}\right)}\varphi\left(\tfrac{t-b}{a}-\tfrac{\nu}{2}\right)\tfrac{1}{a^2}\delta(\nu-\tfrac{\tau}{a})\mathrm{d}\tau\mathrm{d}\nu\mathrm{d}t \\
&\quad = \iint x(t+\tfrac{\tau}{2})\overline{y(t-\tfrac{\tau}{2})}\,\overline{\psi\left(\frac{t+\frac{\tau}{2}-b}{a}\right)}\varphi\left(\frac{t-\frac{\tau}{2}-b}{a}\right)\frac{\mathrm{d}\tau\mathrm{d}t}{a^2} \\
&\quad = \int x(t_1)\overline{\psi\left(\frac{t_1-b}{a}\right)}\frac{\mathrm{d}t_1}{a}\int\overline{y(t_2)}\varphi\left(\frac{t_2-b}{|a|}\right)\frac{\mathrm{d}t_2}{|a|} \\
&\quad = \mathcal{W}_\psi[x](a,b)\overline{\mathcal{W}_\varphi[y](a,b)} \quad .
\end{aligned}$$

The analogous result for the Gabor transform, defined in Eq.(A.38) is as follows:

$$\begin{aligned}
&\iint W_{xy}(t,f)\overline{W_{vw}(t'-t,f'-f)}\mathrm{d}t'\mathrm{d}f' \\
&\quad= \iiiint x(t'+\tfrac{\tau}{2})\overline{y(t'-\tfrac{\tau}{2})}\,\overline{w\left(t'-t+\tfrac{\nu}{2}\right)}v\left(t'-t-\tfrac{\nu}{2}\right)e^{-i\Upsilon f'\tau+i\Upsilon(f'-f)\nu}\mathrm{d}\tau\mathrm{d}\nu\mathrm{d}t'\mathrm{d}f' \\
&\quad= \iiint x(t'+\tfrac{\tau}{2})\overline{y(t'-\tfrac{\tau}{2})}\,\overline{w\left(t'-t+\tfrac{\nu}{2}\right)}v\left(t'-t-\tfrac{\nu}{2}\right)\delta(\nu-\tau)e^{-i\Upsilon f\nu}\mathrm{d}\tau\mathrm{d}\nu\mathrm{d}t' \\
&\quad= \iint x(t'+\tfrac{\tau}{2})\overline{y(t'-\tfrac{\tau}{2})}\,\overline{w\left(t'+\tfrac{\tau}{2}-t\right)}v\left(t'-\tfrac{\tau}{2}-t\right)e^{-i\Upsilon f\tau}\mathrm{d}\tau\mathrm{d}t \\
&\quad= \int x(t_1)\overline{w(t_1-t)}e^{-i\Upsilon f t_1}\mathrm{d}t_1\int\overline{y(t_1)}v(t_2-t)e^{i\Upsilon f t_2}\mathrm{d}t_2 \\
&\quad= \mathcal{G}_w[x](t,f)\overline{\mathcal{G}_v[y](t,f)} \quad .
\end{aligned}$$

3 Change of wavelet formula

We prove here the formula for changing of wavelet:

$$\mathcal{W}_r[s](a,b)=C_{\varphi\psi}^{-1}\iint\mathcal{W}_\psi[s](a',b')\overline{\mathcal{W}_\varphi[r]\left(\frac{a'}{a},\frac{b'-b}{a}\right)}\frac{\mathrm{d}a'\mathrm{d}b'}{a|a'|}\;;\quad C_{\varphi\psi}=\int\hat{\varphi}(v)\overline{\hat{\psi}}(v)\frac{\mathrm{d}v}{|v|}\quad, \tag{A.40}$$

which can be understood as a change of basis in a continuous and overcomplete context. The reconstruction formulas Eq.(I.15) ($r=\delta$) and Eq.(I.14) ($r=\delta=\varphi$) and the reproducing kernel formula $r=\psi$ follow as special cases.

To demonstrate it, we prove first the simpler "Plancherel-Parseval theorem for wavelet transform":

$$C_{\varphi\psi}^{-1}\iint\mathcal{W}_\psi[s](a,b)\overline{\mathcal{W}_\varphi[r](a,b)}\frac{\mathrm{d}a\mathrm{d}b}{|a|}=\int\hat{s}(f)\overline{\hat{r}(f)}\mathrm{d}f=\int s(t)\overline{r(t)}\mathrm{d}t\quad, \tag{A.41}$$

where the usual theorem only refer to the second inequality. We use the definition of the wavelet transform from the frequency domain, simplified using a property of the Dirac delta δ, $\iint h(x)e^{i\Upsilon z(y-x)}\mathrm{d}z\mathrm{d}x=\int h(x)\delta(y-x)\mathrm{d}x=h(y)$, in the following:

$$\begin{aligned}
\iint\mathcal{W}_\psi[s](a,b)\overline{\mathcal{W}_\varphi[r](a,b)}\frac{\mathrm{d}a\mathrm{d}b}{|a|} &= \iint\int\hat{s}(f)\overline{\hat{\psi}(af)}e^{i\Upsilon fb}\mathrm{d}f\int\overline{\hat{r}(f')}\hat{\varphi}(af')e^{-i\Upsilon f'b}\mathrm{d}f'\frac{\mathrm{d}a\mathrm{d}b}{|a|} \\
&= \iiint\hat{s}(f)\overline{\hat{r}(f')}\hat{\varphi}(af')\overline{\hat{\psi}}(af)\left(\int\hat{\psi}(af')e^{i\Upsilon(f-f')b}\mathrm{d}b\right)\mathrm{d}f\mathrm{d}f'\frac{\mathrm{d}a}{|a|} \\
&= \iint\hat{s}(f)\overline{\hat{r}(f')}\int\hat{\varphi}(af')\overline{\hat{\psi}}(af)\frac{\mathrm{d}a}{|a|}\delta(f-f')\mathrm{d}f\mathrm{d}f' \\
&= \int\hat{s}(f)\overline{\hat{r}(f)}\left(\int\hat{\varphi}(af)\overline{\hat{\psi}}(af)\frac{\mathrm{d}a}{|a|}\right)\mathrm{d}f \\
&= \int\hat{s}(f)\overline{\hat{r}(f)}\mathrm{d}f\int\hat{\varphi}(v)\overline{\hat{\psi}(v)}\frac{\mathrm{d}v}{|v|} \quad .
\end{aligned}$$

Notice that the integral $\int\frac{\mathrm{d}a}{|a|}$ is important to get the last equality. Indeed, among measures $\mathrm{d}\mu(a)$, it is the only one that is invariant with respect to varying f real: it is the Haar measure of the group of real numbers with multiplication.

Next we introduce new variables (a,b), differentiating the integrated ones by a dash, that turn $r(t)$ into the scaled and shifted wavelet $r(\frac{t-b}{a})\frac{1}{a}$, i.e. $\hat{r}(af)e^{-i\Upsilon fb}$, to obtain the wavelet transform with respect to this new wavelet in Eq.(A.40). The correct arrangement of the

original (dash) and new variables are identified inside the wavelet φ after a change of variable $af = f'$:

$$\int \hat{r}(af)e^{-i\Upsilon fb}\overline{\hat{\varphi}(a'f)}e^{i\Upsilon fb'}\mathrm{d}f = \int \hat{r}(f')\overline{\hat{\varphi}(a'f'/a)}e^{i\Upsilon f\frac{b'-b}{a}}\mathrm{d}f/|a| \quad , \tag{A.42}$$

the scale is $\frac{a'}{a}$ and the time is $\frac{b'-b}{a}$ which ends the proof.

Eq.(A.40) can be written as the wavelet transform of a wavelet transform (as a signal) with respect to another wavelet transform (as a wavelet), integrated over all scales:

$$\mathcal{W}_r[s](a,b) = C_{\varphi,\psi}^{-1}\int \mathcal{W}_{\mathcal{W}_\varphi[r](a'/a,\cdot)}[\mathcal{W}_\psi[s](a',\cdot)](a,b)\frac{\mathrm{d}a'}{|a'|} \quad ,$$

where the dots indicate the transformed time variable b'.

4 Entropic uncertainty

The well known Heisenberg uncertainty relation of physics is a mathematical inequality that limit the product of the spread of a function with the one of its Fourier transform, see definition in Eq.(I.1), in our case the time-width and the frequency-width, see Eq.(I.31). The bound is saturated by Gaussian functions. It can be deduced from the bounded sum of their Shannon entropies. Originally based on the 2-norm, the uncertainty relation has been generalized [Babenko 1961; Beckner 1975; Hirschman 1957] to other norms, related to the notion of Rényi entropies. This is called the *entropic uncertainty*, based on the sharp Hausdorff-Young or Babenko-Beckner inequality:

$$\left(\sqrt{p}\int |x(t)|^p\mathrm{d}t\right)^{\frac{1}{p}} \geq \left(\sqrt{q}\int |\hat{x}(f)|^q\mathrm{d}f\right)^{\frac{1}{q}} \tag{A.43}$$

holds whenever $\frac{1}{p}+\frac{1}{q}=1$ and $1<p\leq 2\leq q$.

Further assuming that $|x(t)|^2$ and $|\hat{x}(f)|^2$ have a unit energy:

$$E_{xx} = \int |x(t)|^2\mathrm{d}t = \int |\hat{x}(f)|^2\mathrm{d}f = 1 \quad , \tag{A.44}$$

we can interpret them as normalized densities and compute their Rényi entropy, defined for any probability density function $\rho(u)$ as:

$$H_\alpha[\rho] = \frac{1}{1-\alpha}\log\left(\int \rho(u)^\alpha \mathrm{d}u\right) \quad , \tag{A.45}$$

for $\alpha\neq 1$, and the case $\alpha\to 1$ coincide with the Shannon entropy.

$$H_1[\rho] = -\int \rho(u)\log\rho(u)\mathrm{d}u \quad . \tag{A.46}$$

Eq.(A.43) can be turned into a bound for the sum of the Rényi entropies:

$$H_{\frac{p}{2}}\left[|x|^2\right] + H_{\frac{q}{2}}\left[|\hat{x}|^2\right] \geq \frac{\log\frac{p}{2}}{p-2} + \frac{\log\frac{q}{2}}{q-2} - \log 2 \quad , \tag{A.47}$$

so that for $p,q\to 2$:

$$H_1\left[|x|^2\right] + H_1\left[|\hat{x}|^2\right] \geq 1-\log 2 \quad . \tag{A.48}$$

The exponential Shannon entropy bounds the standard deviation Δ_ρ by below:

$$e^{H_1[\rho]} \leq \sqrt{\Upsilon e}\Delta_\rho \quad , \qquad \Delta_\rho^2 = \int u^2\rho(u)\mathrm{d}u - \left(\int u\rho(u)\mathrm{d}u\right)^2 \quad . \tag{A.49}$$

This implies the usual Heisenberg uncertainty relation:

$$\Delta_{|x|^2}\Delta_{|\hat{x}|^2}\Upsilon \geq e^{H_1\left[|x|^2\right]+H_1\left[|\hat{x}|^2\right]-1} \geq \frac{1}{2} \quad . \tag{A.50}$$

5 Orthogonal analytic wavelets: from Morse to log-normal

Following Eq.(I.25), we consider the standardized version of the full Morse family of orthogonal wavelets:

$$\hat{\varphi}^{(k)}_{\beta,\gamma} = \left(ve^{\frac{1-v^\gamma}{\gamma}}\right)^\beta L_k^{(\alpha-1)}\left(\frac{2\beta v^\gamma}{\gamma}\right) \quad , \quad \alpha = \frac{2\beta+1}{\gamma} \quad , \tag{A.51}$$

where generalized Laguerre polynomials $L_k^{(\alpha')}$ are defined in Eq.(I.17).

We prove here their convergence to the Q-parameterized "log-Hermite" orthogonal family of wavelet $\hat{\psi}^{(k)}_Q(v)$ defined in Eq.(I.46), in the scaling limit $\gamma \to 0$ and $\beta \to \infty$ with $Q = \sqrt{\beta\gamma}$ fixed. We recall that this convergence is proved for the order $k = 0$ thanks to the series Eq.(I.35).

In order to prove the high orders, we focus on the orthogonal polynomials. All of them can be studied at once by introducing the generating functions of the (generalized) Laguerre and (physicist's) Hermite polynomials, simply given here:

$$\sum_{k=0}^\infty L_k^{(\alpha-1)}(x)\epsilon^k = \left(\frac{1}{1-\epsilon}\right)^\alpha e^{-x\frac{\epsilon}{1-\epsilon}} \tag{A.52}$$

$$\sum_{k=0}^\infty H_k(y)\frac{\omega^k}{k!} = e^{2y\omega-\omega^2} \quad , \tag{A.53}$$

from which any order k polynomial is recovered as their k^{th} derivative with respect to ϵ, ω evaluated at zero.

The choice of the variable ϵ is intentional since the parameters α, x in the Laguerre generating function diverges in the scaling limit: we will choose $\epsilon \to 0$ vanishing at a suitable rate to compensate for their divergence. From the following Taylor expansions at the second order:

$$\frac{\epsilon}{1-\epsilon} = \epsilon + \epsilon^2 + \mathcal{O}(\epsilon^3) \quad ; \quad -\log(1-\epsilon) = \epsilon + \frac{\epsilon^2}{2} + \mathcal{O}(\epsilon^3) \quad ,$$

we can rewrite the Laguerre generating function as:

$$\sum_{k=0}^\infty L_k^{(\alpha-1)}(x)\epsilon^k = e^{\alpha(\epsilon+\frac{\epsilon^2}{2})-x(\epsilon+\epsilon^2)+\mathcal{O}(\alpha\epsilon^3,\, x\epsilon^3)} \quad .$$

Now we replace, the parameters by their expression for the standardized Morse wavelets, :

$$\begin{aligned}
\log\sum_{k=0}^\infty L_k^{(\alpha-1)}(x)\epsilon^k &= (\alpha-x)\epsilon + \left(\frac{\alpha}{2}-x\right)\epsilon^2 + \mathcal{O}(\alpha\epsilon^3,\, x\epsilon^3) \\
&= \frac{2\beta+1-2\beta v^\gamma}{\gamma}\epsilon + \frac{\beta+\frac{1}{2}-2\beta v^\gamma}{\gamma}\epsilon^2 + \mathcal{O}\left(\frac{\beta}{\gamma}\epsilon^3,\, \frac{\beta v^\gamma}{\gamma}\epsilon^3\right) \\
&= 2\beta\left(\frac{1}{2\beta\gamma} + \frac{1-v^\gamma}{\gamma}\right)\epsilon + \frac{\beta}{\gamma}\left(1+\frac{1}{2\beta}-2v^\gamma\right)\epsilon^2 + \mathcal{O}\left(\frac{\beta}{\gamma}\epsilon^3,\, \frac{\beta v^\gamma}{\gamma}\epsilon^3\right) \\
&\sim 2Q\sqrt{\frac{\beta}{\gamma}}\left(\frac{1}{2Q^2} - \log v\right)\epsilon - \frac{\beta}{\gamma}\epsilon^2 + \mathcal{O}\left(\frac{\beta}{\gamma}\epsilon^3\right) \\
&\longrightarrow 2\left(\frac{1}{2Q} - Q\log v\right)\omega - \omega^2 \quad , \qquad \omega = \sqrt{\frac{\beta}{\gamma}}\epsilon \quad ,
\end{aligned}$$

where we assume v finite and take the scaling limit in the last lines, with $\frac{v^\gamma-1}{\gamma} \to \log v$, $\epsilon \to 0$ so that $\omega = \sqrt{\frac{\beta}{\gamma}}\epsilon$ is finite and $\frac{\beta}{\gamma}\epsilon^3 = \omega\epsilon \to 0$. We obtain the Hermite generating function with the parameter $y = \frac{1}{2Q} - Q\log v$, which explains how Eq.(I.46) is obtained from the Morse wavelet family.

6 Asymptotic transforms for limiting quality factors

Consider a standardized and symmetric analytic wavelet $\psi = \varphi_{\beta,\gamma}$, defined in Eq.(I.25). In certain asymptotic conditions for the wavelet and its quality factor, the wavelet transform of a centred signal $x(t)$ (i.e. assuming $\hat{x}(0) = 0$) can reach fully localized and well-known limits, the analytic signal and the Fourier spectrum:

$$\lim_{Q\to 0} \frac{\mathcal{W}_\psi[x](f^{-1},t)}{\overline{\hat{\psi}(1)}} = x^+(t) \tag{A.54}$$

$$\lim_{Q\to \infty} \frac{\mathcal{W}_\psi[x](f^{-1},t)}{\overline{\psi(0)}} = \hat{x}(f)e^{i\tau ft}|f| \quad . \tag{A.55}$$

Indeed, the wavelet spectrum $\varphi_{\beta,\gamma}$ with $Q = \sqrt{\beta\gamma}$ tends to the Heaviside step function $\Theta(v)$ at a vanishing quality factor, and to the Dirac delta $\delta(v-1)$ at a diverging one [Lilly and Olhede 2009], when suitably normalized. More precisely these are double limits on the Morse parameters: the first case is reached when $\beta = 0 = \gamma$ or when one of them is zero the other one being finite, while the second case is valid when $\beta \to \pm\infty$ with γ finite or diverging with the same sign (but not when γ diverges alone). This subtlety disappears when using the log-normal wavelet $\psi = \psi_Q$, Eq.(I.33).

While the symmetry of the Morse wavelet in the time domain is important to relate the transform to the analytic signal, its standardization is essential to express the asymptotic relation to the spectrum, otherwise complicated by cases of diverging characteristic wavelet frequency. We present below proofs of these equalities.

For the first limit, we notice that any standardized Morse wavelet of finite exponents β, γ, as well as the log-normal case, is positive and finite at all positive frequencies and zero anywhere else (analytic), so that it tends to the Heaviside step function when raised to a vanishing power $\lim_{\epsilon\to 0} \hat{\psi}(f)^\epsilon \to \Theta(f)$, except for the undefined zero frequency. This works with Q for the log-normal wavelet $\hat{\psi}_Q = (\hat{\psi}_1)^{Q^2}$ and with β for the standardized Morse $\hat{\varphi}_{\beta,\gamma} = (\hat{\varphi}_{1,\gamma})^\beta$. Therefore, the associated wavelet transform of a signal tends in this limit to the definition of the analytic signal, Eq.(I.2), where the assumption $\hat{x}(0) = 0$ removes the ambiguity at $f = 0$. For a vanishing parameter γ, the result is the same for a standardized Morse wavelet $\hat{\varphi}_{\beta,\gamma}(v) = (ve^{\frac{1-v^\gamma}{\gamma}})^\beta$, because the expression $\frac{v^\gamma-1}{\gamma}$ is a deformed logarithm, the usual one corresponding to the limit $\gamma \to 0$ for which $ve^{\frac{1-v^\gamma}{\gamma}} \to ve^{-\log v} = 1$ for all positive v. The normalization by $\overline{\hat{\psi}(1)}$ (where 1 is the standardized peak frequency) just ensures that the height of the peak is normalized to 1.

For the second limit, we rely on the interpretation of $\hat{\psi}$ as a density function in the frequencies (or in log-frequencies), normalized to a unit weight by its integral over frequencies $\int_0^\infty \hat{\psi}(v)\mathrm{d}v = \psi(0)$ (respectively by $\int \hat{\psi}(v)\mathrm{d}\log v = C_\psi$ for log-frequencies). In the limit of a diverging quality factor, this normalized density function is increasingly localized at the peak frequency $v_\infty = 1$ so that it approximate increasingly well the Dirac delta, sometimes called a nascent delta function: $\lim_{N\to\infty} \rho(Nx)N\mathrm{d}x = \delta(x)\mathrm{d}x$. It is easy to show it for the log-normal wavelet normalized by C_{ψ_Q}, where $x = \log v$ with the quality factor as the diverging parameter:

$$\hat{\psi}_Q(v)C_{\psi_Q}^{-1}\mathrm{d}\log v = e^{-\frac{1}{2}(Q\log v)^2}\frac{Q\mathrm{d}\log v}{\sqrt{\uparrow}} = \rho(Q\log v)Q\mathrm{d}\log v \quad ,$$

with ρ the standard Gaussian density function. In the definition Eq.(I.7) of the wavelet transform from the frequency domain, $v = af' = \frac{f'}{f}$ and the integrator is $\mathrm{d}f' = f'\mathrm{d}\log f' = f'\mathrm{d}\log v$, so that $\hat{\psi}_Q(\frac{f'}{f})C_{\psi_Q}^{-1}\mathrm{d}f' = f'\rho(Q\log\frac{f'}{f})Q\mathrm{d}\log f'$. In the limit $Q \to \infty$, we replace it by $f'\delta(\log\frac{f'}{f})\mathrm{d}\log f' = f'\delta(f'-f)\mathrm{d}f'$, so that $\hat{x}(f)e^{i\tau ft}f$ remains after the integration (the

Dirac delta evaluates the integrand at $f' = f$). When $f < 0$, then f' is also negative and we use instead $\mathrm{d}f' = |f'|\frac{\mathrm{d}f'}{|f'|}$, which explains the absolute value in $\hat{x}(f)e^{i\Upsilon ft}|f|$ The normalization by $\psi(0)$ instead of C_ψ is equivalent in the limit of a diverging quality factor since the peak frequency is one: $\delta(v-1) = \delta(\log v)$.

7 Comparison of general covariance estimators

We provide details for the discussion around the definition of a wavelet-estimator $R_{xy}(t', t; Q)$ Eq.(II.39) for the general covariance $\sigma^2_{xy}(t', t) = \mathbb{E}[x(t')y(t)]$ of centred stochastic processes $x(t)$ and $y(t)$. More precisely, we compare:

$$R_{xy}(t', t; Q) = \int X(t', f; Q)\overline{Y(t, f; Q)}\frac{\mathrm{d}f}{|f|}\frac{Q}{\sqrt{\pi}} \tag{A.56}$$

$$R^{\mathrm{alt}}_{xy}(t + \tfrac{\tau}{2}, t - \tfrac{\tau}{2}; Q) = \int X(t, f; Q)\overline{Y(t, f; Q)}e^{i\Upsilon f\tau}\frac{\mathrm{d}f}{|f|}\frac{Q}{\sqrt{\pi}} \quad . \tag{A.57}$$

The first observation is their agreement for $\tau = t' - t = 0$: $R_{xy}(t, t; Q) = R^{\mathrm{alt}}_{xy}(t, t; Q) = P_{xy}(t; Q)$.

Then, we examine the asymptotic behaviours for limiting quality factors, applying the results computed in Appendix 6. In particular, Eq.(II.8) yields:

$$R_{xy}(t', t; Q \to \infty) \sim \frac{Q}{\sqrt{\pi}}\frac{\Upsilon}{Q^2}\int \hat{x}(f)\overline{\hat{y}(f)}e^{i\Upsilon f(t'-t)}|f|^2\frac{\mathrm{d}f}{|f|} \sim \frac{\sqrt{2\Upsilon}}{Q}\int \hat{x}(f)\overline{\hat{y}(f)}e^{i\Upsilon f(t'-t)}|f|\mathrm{d}f$$

$$R^{\mathrm{alt}}_{xy}(t + \tfrac{\tau}{2}, t - \tfrac{\tau}{2}; Q \to \infty) \sim \frac{\sqrt{2\Upsilon}}{Q}\int \hat{x}(f)\overline{\hat{y}(f)}e^{i\Upsilon f\tau}|f|\mathrm{d}f \quad .$$

Both approaches agree in this limit, which is not well-defined since the inverse of Q vanishes for $Q \to \infty$.

Notice the asymptotic stationarity. The presence of the factor $|f|$ in the integral is a major difference with the stationary case Eq.(II.24):

$$R_{xy}(\tau) = \int S_{xy}(f)e^{i\Upsilon f\tau}\mathrm{d}f = \lim_{T\to\infty}\int \frac{\hat{x}_T(f)\overline{\hat{y}_T(f)}}{T}e^{i\Upsilon f\tau}\mathrm{d}f \quad .$$

The role of the diverging duration T is played by the diverging quality factor.

We recognize in the prefactor $\frac{\sqrt{2\Upsilon}}{Q}$ the inverse of the dimensionless time-resolution for quadratic expressions: $f\delta t = \frac{Q}{\sqrt{2\Upsilon}}$, see Table I.1. In practice, the finite duration T must contain at least one resolution unit, so that it constrains the choice of a maximum quality factor to the one of a minimum frequency: $\frac{Q_{\max}}{\sqrt{2\Upsilon}} \sim f_{\min}T$; in particular, setting the minimum frequency down to $f_{\min} = T^{-1}$ yields a maximum quality factor as low as $Q_{\max} \sim \sqrt{2\Upsilon}$.

Similarly, applying Eq.(II.10) in the limit of a vanishing quality factor leads to:

$$R_{xy}(t', t; Q \to 0) \sim \frac{Q}{\sqrt{\pi}}\int x^{\mathrm{sgn}(f)}(t')\overline{y^{\mathrm{sgn}(f)}(t)}\frac{\mathrm{d}f}{|f|}$$

$$R^{\mathrm{alt}}_{xy}(t + \tfrac{\tau}{2}, t - \tfrac{\tau}{2}; Q \to 0) \sim \frac{Q}{\sqrt{\pi}}\int x^{\mathrm{sgn}(f)}(t')\overline{y^{\mathrm{sgn}(f)}(t)}e^{i\Upsilon f\tau}\frac{\mathrm{d}f}{|f|} \quad ,$$

where $\mathrm{sgn}(f) = \pm$ is the sign of f. The integral diverges while the quality factor vanishes, so that we need to disambiguate it. The ambiguity dissipates for the first definition, we use

$\hat{\psi}_Q(f_1/f)\hat{\psi}_Q(f_2/f) = \hat{\psi}_{\sqrt{2}Q}\left(\sqrt{f_1 f_2}/f\right)\hat{\psi}_{Q/\sqrt{2}}(f_1/f_2)$ to derive:

$$\begin{aligned}
R_{xy}(t',t;Q) &= \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\int \overline{\hat{\psi}_Q(f_1/f)}\hat{\psi}_Q(f_2/f)\frac{\mathrm{d}f}{|f|}e^{i\Upsilon(f_1t'-f_2t)}\mathrm{d}f_1\mathrm{d}f_2\frac{Q}{\sqrt{\pi}}\\
&= \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\int \hat{\psi}_{\sqrt{2}Q}\left(\sqrt{f_1 f_2}/f\right)\frac{\mathrm{d}f}{|f|}\hat{\psi}_{Q/\sqrt{2}}(f_1/f_2)e^{i\Upsilon(f_1t'-f_2t)}\mathrm{d}f_1\mathrm{d}f_2\frac{Q}{\sqrt{\pi}}\\
&= \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\hat{\psi}_{Q/\sqrt{2}}(f_1/f_2)e^{i\Upsilon(f_1t'-f_2t)}\mathrm{d}f_1\mathrm{d}f_2\int \hat{\psi}_{\sqrt{2}Q}(v)\frac{\sqrt{2}Q}{\sqrt{\Upsilon}}\mathrm{d}\log v\\
&= \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\hat{\psi}_{Q/\sqrt{2}}(f_1/f_2)e^{i\Upsilon(f_1t'-f_2t)}\mathrm{d}f_1\mathrm{d}f_2\\
R_{xy}(t',t;Q\to 0) &= \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\theta(f_1/f_2)e^{i\Upsilon(f_1t'-f_2t)}\mathrm{d}f_1\mathrm{d}f_2\\
&= x^+(t')\overline{y^+(t)} + x^-(t')\overline{y^-(t)}
\end{aligned}$$

replacing the limit wavelet spectrum by the Heaviside step function θ. This proves Eqs.(II.22, II.40).

For the alternative definition, however, the ambiguity remains, as a sign of the discrepancy between this instantaneous limit and the spectral approach:

$$\begin{aligned}
R^{\mathrm{alt}}_{xy}(t+\tfrac{\tau}{2},t-\tfrac{\tau}{2};Q) &= \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\int \overline{\hat{\psi}_Q(f_1/f)}\hat{\psi}_Q(f_2/f)e^{i\Upsilon f\tau}\frac{\mathrm{d}f}{|f|}e^{i\Upsilon(f_1-f_2)t}\mathrm{d}f_1\mathrm{d}f_2\frac{Q}{\sqrt{\pi}}\\
R^{\mathrm{alt}}_{xy}(t+\tfrac{\tau}{2},t-\tfrac{\tau}{2};Q\to 0) &= \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\int \theta(f_1/f)\theta(f_2/f)e^{i\Upsilon f\tau}\frac{\mathrm{d}f}{|f|}e^{i\Upsilon(f_1-f_2)t}\mathrm{d}f_1\mathrm{d}f_2\frac{Q}{\sqrt{\pi}}\\
&= \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\theta(f_1/f_2)\int_0^\infty e^{\mathrm{sgn}(f_1)i\Upsilon f\tau}\mathrm{dlog}\, f\, e^{i\Upsilon(f_1-f_2)t}\mathrm{d}f_1\mathrm{d}f_2\frac{Q}{\sqrt{\pi}} \quad .
\end{aligned}$$

The integral with respect to f presents a logarithmic divergence, as in Eq.(II.47). Therefore, we follow the same strategy of regularization by introducing a minimum frequency T^{-1}:

$$\begin{aligned}
\int_{T^{-1}}^\infty e^{i\Upsilon f\tau}\mathrm{dlog}\, f &= \int_{T^{-1}}^\infty \cos(\Upsilon f\tau) + i\sin(\Upsilon f\tau)\mathrm{dlog}\, f\\
&= \gamma + \log\frac{T}{\Upsilon\tau} + \int_{T^{-1}}^\infty 1-\cos(\Upsilon f\tau)\mathrm{dlog}\, f + i\mathrm{sgn}(\tau)\frac{\pi}{2} - i\int_0^{T^{-1}}\sin(\Upsilon f\tau)\mathrm{dlog}\, f\\
&= \gamma + \log\frac{T}{\Upsilon\tau} + \mathrm{Cin}\left(\frac{T}{\Upsilon\tau}\right) + \mathrm{sgn}(\tau)i\frac{\pi}{2} - i\mathrm{Si}\left(\frac{T}{\Upsilon\tau}\right) \quad , \qquad \text{(A.58)}
\end{aligned}$$

where we have introduced γ the Euler-Mascheroni constant and special functions called the trigonometric integrals:

$$\mathrm{Cin}(y) = \int_y^\infty 1-\cos x\,\mathrm{dlog}\, x = \sum_{n=1}^\infty (-1)^n\frac{y^{2n}}{(2n)!\,2n} \qquad \text{(A.59)}$$

$$\mathrm{Si}(y) = \int_0^y \sin x\,\mathrm{dlog}\, x = \sum_{n=1}^\infty (-1)^n\frac{y^{2n+1}}{(2n+1)!\,(2n+1)} \quad . \qquad \text{(A.60)}$$

They describe alternating polynomial corrections to this logarithmic divergence, related to the Gibbs phenomenon, an oscillation that appears when truncating at a finite cut-off frequency. Assuming $\tau \neq 0$ (i.e. not the special case of the variance or power), we can neglect them when $T\to\infty$ since $\mathrm{Cin}(0) = \mathrm{Si}(0) = 0$. This gives a more precise view of the problem:

$$\begin{aligned}
R^{\mathrm{alt}}_{xy}(t+\tfrac{\tau}{2},t-\tfrac{\tau}{2};Q\to 0) &\sim \iint \hat{x}(f_1)\overline{\hat{y}(f_2)}\theta(f_1/f_2)\left(\gamma + \log\frac{T}{\Upsilon\tau} + \mathrm{sgn}(f_1\tau)i\frac{\pi}{2}\right)\frac{Q}{\sqrt{\pi}}\,e^{i\Upsilon(f_1-f_2)t}\mathrm{d}f_1\mathrm{d}f_2\\
&\sim \frac{Q}{\sqrt{\pi}}\log\frac{T}{\Upsilon\tau}\left(x^+(t)\overline{y^+(t)} + x^-(t)\overline{y^-(t)}\right) \quad ,
\end{aligned}$$

where constant terms have been neglected for Q small enough.

Once again, we recognize in the prefactor $\frac{Q}{\sqrt{\pi}}$ the inverse of the log-frequency resolution for quadratic expressions: $\delta \log f = \frac{\sqrt{\pi}}{\sqrt{2}Q}$, see Table I.1. It should not exceed the ratio between the largest and the smallest scale; in particular, the result $R_{xy}(t', t; Q \to 0)$ can only be recovered from the alternative Fourier definition for $\delta \log f = \log \frac{T}{\Upsilon_\tau}$ and $\tau \to 0$. This explains our choice Eq.(A.56) for a wavelet-estimator of the general covariance.

8 Laplace's method applied to time-smoothing estimator

The Laplace's method aims at approximating certain integrals of a positive function by the integral of a Gaussian which has the same global maximum up to second order (peak value and curvature). Provided that the integrated function is indeed twice differentiable at its unique maximum, the precision of the approximation relies on the existence of a large exponent α, ensuring that essential contributions to the integral come from the neighbourhood of the maximum value of the integrated function. We introduce the 2-dimensional form of this approximation for a double integral of a bivariate function:

$$G(\alpha) = \iint e^{\alpha g(x,y)} \mathrm{d}x\mathrm{d}y \tag{A.61}$$

$$\approx \frac{\Upsilon}{\alpha} \det[-\partial_x \partial_y g(x_0, y_0)]^{-\frac{1}{2}} e^{\alpha g(x_0,y_0)} \quad , \tag{A.62}$$

where (x_0, y_0) is the position of the global maximum of the function $g(x, y)$, $\partial_x \partial_y g(x_0, y_0)$ is its Hessian matrix evaluated at the maximum, and det refers to its determinant.

We now apply it to the estimation of the significance of the practical time-smoothing estimator defined in Eqs.(II.70, II.71). This estimator is characterized by a kernel that takes the following form in the frequency domain:

$$\hat{K}(v, v') = \hat{\psi}_Q(v)\overline{\hat{\psi}_Q(v')}\hat{\chi}_n(v - v')\frac{Q}{\sqrt{\pi}} = e^{-\frac{Q^2}{2}(\log v)^2 - \frac{Q^2}{2}(\log v')^2 - (\frac{nQ}{2})^2(v-v')^2} \quad . \tag{A.63}$$

It is associated to a number of statistical degrees of freedom n_d is defined in Eq.(II.65) from the ratio of the integral expressions:

$$\left(\int \hat{K}(v, v)\mathrm{d}v\right)^2 = \left(\int |\hat{\psi}_Q(v)|^2 \mathrm{d}v \frac{Q}{\sqrt{\pi}}\right)^2 = E^2_{\psi_Q\psi_Q}\frac{Q^2}{\pi} = \lambda^{\frac{1}{2}} \tag{A.64}$$

$$\iint |\hat{K}(v, v')|^2 \mathrm{d}v\mathrm{d}v' = \iint e^{-(Q\log v)^2 - (Q\log v')^2 - \frac{(nQ)^2}{2}(v-v')^2}\Theta(v)\Theta(v')\mathrm{d}v\mathrm{d}v'\frac{Q^2}{\pi} \quad , \tag{A.65}$$

with $\lambda = e^{\frac{1}{Q^2}}$. We use the Laplace's method to approximate the second double integration, where we identify:

$$\lambda^{\frac{1}{2}}\beta^2 \iint e^{\alpha g(x,y)}\mathrm{d}x\mathrm{d}y = \iint e^{\frac{1}{2Q^2} - (Q\log v - \frac{1}{2Q})^2 - (Q\log v' - \frac{1}{2Q})^2 - \frac{(nQ)^2}{2}(e^{\log v} - e^{\log v'})^2} \mathrm{d}\log v \mathrm{d}\log v'$$

$$\alpha = \frac{(nQ)^2}{2}\lambda \tag{A.66}$$

$$(x, y) = (\log v - \tfrac{1}{2Q^2}, \log v' - \tfrac{1}{2Q^2})/\beta \quad ; \quad \beta = \frac{n}{\sqrt{2}}\lambda^{\frac{1}{2}} \tag{A.67}$$

$$g(x, y; \beta) = -x^2 - y^2 - (e^{\beta x} - e^{\beta y})^2 \quad . \tag{A.68}$$

The maximum is 0 at $(x_0, y_0) = (0, 0)$ and the Hessian matrix at the maximum is

$$-\partial_x \partial_y g(x_0, y_0) = \begin{bmatrix} 2+2\beta^2 & -2\beta^2 \\ -2\beta^2 & 2+2\beta^2 \end{bmatrix}$$

with determinant $4 + 8\beta^2$. Therefore, we obtain Eq.(II.72):

$$\iint |\hat{K}(v, v')|^2 \mathrm{d}v \mathrm{d}v' \approx \lambda^{\frac{1}{2}} \beta^2 G(\alpha; \beta) = \lambda^{\frac{1}{2}} (1 + n^2 \lambda)^{-\frac{1}{2}}$$
$$n_d \approx (1 + n^2 \lambda)^{\frac{1}{2}} \quad , \tag{A.69}$$

whenever $\frac{(nQ)^2}{2}\lambda$ is large.

9 Mean instantaneous frequency from the reassigned power density

Here we consider a real and oscillating signal $x(t)$, and we relate the generic quadratic estimation for its time-dependent rate, $F_x(t; Q)$ defined in Eq.(III.28), with the frequency-reassigned scalogram Eq.(III.17). In particular, the real part is its mean instantaneous frequency.

Since the signal is real, the scalogram is symmetric with respect to frequency: $|X(t, -f; Q)|^2 = |X(t, f; Q)|^2$, as well as the frequency-reassigned scalogram. The mean frequency of the real signal would always be zero, without restriction to positive frequency. For this reason, we consider the frequency-reassigned scalogram of the analytic version of the signal $x^+(t)$, which only reassign values of the half-plane $(t, f > 0)$ to $(t, \Re\{F_x(t, f; Q)\})$.

For more clarity, let us normalized the frequency-reassigned power density Eq.(III.17) of the analytic signal $x^+(t)$ as:

$$\rho_{x^+}(f|t; Q) = \mathcal{R}[|X|^2](t, f; Q) \frac{Q}{\sqrt{\pi}} P_{x^+}(t; Q)^{-1} \quad , \tag{A.70}$$

so that it takes the form of a probability density function for the instantaneous frequency, at each time. The normalization:

$$\int \rho_{x^+}(f|t; Q) \frac{\mathrm{d}f}{|f|} = 1 \quad , \tag{A.71}$$

is given by Eq.(III.19), except that $P_{x^+}(t; Q) = \frac{1}{2} P_x(t; Q)$.

This density function is highly concentrated on a phase ridge $f(t)$ whenever the signal contains a single slowly modulated harmonic component.

Now we use the definition of the frequency-reassignment and the wavelet-based instanta-

neous frequency that underlies it, to prove that $\Re\{F_x(t;Q)\}$ is the arithmetic mean of $\rho_x(f|t;Q)$:

$$\langle f \,|\, t;Q\rangle_x = \int f\, \rho_x(f|t;Q)\frac{\mathrm{d}f}{|f|} \tag{A.72}$$

$$\begin{aligned}
&= \frac{\int f \int_0^\infty \delta(f-\Re\{F_x(t,f';Q)\})|f||X(t,f';Q)|^2 \mathrm{d}\log f' \frac{\mathrm{d}f}{|f|}}{\int_0^\infty \delta(f-\Re\{F_x(t,f';Q)\})|f||X(t,f';Q)|^2 \mathrm{d}\log f \frac{\mathrm{d}f}{|f|}} \\
&= \frac{\int_0^\infty \Re\{F_x(t,f';Q)\}|X(t,f';Q)|^2 \mathrm{d}\log f'}{\int_0^\infty |X(t,f';Q)|^2 \mathrm{d}\log f'} \\
&= \frac{\int_0^\infty \Re\{\frac{\dot{X}(t,f';Q)}{i\curlywedge X(t,f';Q)}|X(t,f';Q)|^2\}\mathrm{d}\log f'}{\frac{1}{2}P_x(t;Q)} \\
&= \Re\left\{\frac{\int_0^\infty \dot{X}(t,f';Q)\overline{X(t,f';Q)}\mathrm{d}\log f'}{i\curlywedge \int_0^\infty |X(t,f';Q)|^2 \mathrm{d}\log f'}\right\} \\
&= \Re\{F_x(t;Q)\} \quad .
\end{aligned} \tag{A.73}$$

We find back the generic quadratic estimation formula Eq.(III.28).

Part B

Physiological rhythms: from recordings to interactions

Chapter IV

Cardio-respiratory rhythms and their modulations

Physiology, in particular the one of the human body, is our biophysical system of interest in this chapter and the following. What better way to observe the dynamics of a living organism than by starting with its vital rhythms? Based on several existing datasets, detailed in Appendix 1, the phenomenology of the cardiac and respiratory intertwined activities is revisited from a wavelet perspective.

The breath and the heart beat forms a duo that is studied from the extraction of their rates form the recording of their activity, followed by their comparison, first as a time-frequency coherence of their modulations, then as a frequency ratio. In practice, these extraction and comparison operations can take diverse forms and multiple steps, whose reconciliation into a common and synthetic language seems a hopeless task; anyway, we endeavour to express them with the log-normal wavelet framework as a syntax. Of course, there already exists numerous methods to extract relevant biological observables from the physical measures, some of them being readily available and optimized in specific situations, such as getting the heart rate from an electrocardiogram. However, several limitations motivates their complete recasting into a wavelet framework. The jump from physical measure to biological observable is processed algorithmically, usually from a set of rules, highly specific to the measuring apparatus, and from classical computational methods beyond their original hypothesis. From an analytic viewpoint, this is a "black box" that obscures theoretical developments in Biophysics. From a practical viewpoint, this constitutes a proliferation of context-dependent criteria and adjustable parameters. We propose to translate these expert skills into the mathematical syntax developed in part A, from the processing of raw recordings to their analysis. We focus on essential physical control parameters, such as time-frequency domains and quality factors, distinct from contingent parameters of the numerical implementation, such as the sampling of the wavelet transform, constrained by the physics and optimized for speed. This constitutes a progress towards a sound and synthetic observational framework to decipher the complex dynamics of living systems.

The extraction of rhythmic modulations from generic estimators, developed in chapter III, is applied in the first section to the estimation of the heart rate variability. Different levels of sophistications are presented, with a single to multiple control parameters which are discussed, and their precision is compared statistically to a conventional heart rate estimation from the PhysioNet cardiovascular signal toolbox [Vest, Da Poian, et al. 2018; Vest, Poian, et al. 2019] on a PPG and ECG database [Karlen et al. 2013].

In the next section, we study cardio-respiratory interactions as a coherence (see chapter II) between modulations of the cardiac frequency and fluctuations of the respiratory rhythm [Faes, Pinna, et al. 2004; Saul, Berger, Chen, et al. 1989; Thomas, Mietus, Peng, and Goldberger 2005]. Two kinds of respiratory signals are considered: either directly the recorded breathing

oscillations, or its estimated frequency and amplitude modulations; the latter is shown to circumvent important limitations of the former. The statistics over the large *Sleep Heart Health Study* polysomnography database [Quan et al. 1997; Zhang et al. 2018] highlight three different timescales of interaction with specific phase relations.

Finally, the coordination and possible synchronization between the cardiac and respiratory rhythms is reformulated in terms of ratios of their rate, also called "pulse respiration quotient" [Scholkmann and Wolf 2019], that we measure from the recording as a time-dependent distribution of the instantaneous frequency ratio. As described in [Bartsch, Liu, Ma, et al. 2014; Bartsch, Schumann, et al. 2012], its variability undergoes sudden transitions between different sleep states, illustrated on two individual polysomnographies.

1 The Heart beat

The heart has a central role in the circulatory system of vertebrate, pulsing blood in a rhythmic way and irrigating all organs in the body. Among tetrapods, breathing orchestrate rhythmically the exchange of gas between the blood and the environment. Therefore, the cardiac and respiratory rhythms are intertwined, finely regulated and interacting together and with other functions of the organism, in particular with the nervous system. While of quite well-defined characteristic timescales, these quasi-periodic activities are fluctuating, in frequency and in amplitude, in a non-stationary and multi-scale way. We examine the fluctuations of the cardiac rhythm, especially the heart rate, before studying its interaction with the respiratory and neural rhythms contained in physiological recordings.

We discuss its recording techniques, the importance of uncovering the variability of the heart rate, the difficulty of a sound physical interpretation of the heart rate and its realization in the wavelet formalism, followed by its concrete estimation compared to conventional techniques.

1.1 Recording the cardiac rhythm

Many techniques are available for recording the cardiac rhythm, relying on different types of physical measurement. The shape of the recorded cardiac oscillations is highly dependent on the choice of the recording method (see Fig.IV.4), and contains different information. While an electrocardiography measures the electric activity of the heart, initiated in the pacemaker cells of the sinoatrial node and conducted to muscular fibres, a photoplethysmography optically measures the oxygenation of the blood, via the absorbance of haemoglobin affected by peripheral blood pulses [Kranjec et al. 2014; Tamura 2019].

In the course of this thesis, we compared the following types of recording of the human cardiac activity: (i) blood pressure (BP) monitoring, (ii) electrocardiography (ECG), (iii) photoplethysmography (PPG), compared in Fig.IV.4, as well as (iv) IR thermography.

Among BP monitoring techniques, the intra-arterial one is the most direct measurement of the state of the circulation, but it is also the most invasive method, requiring the placement of a catheter, with all the implied limitations.

The ECG is the most common clinical technique: once amplified, the electric potential between electrodes in contact with the skin on the chest and limbs provides a clear signal with sharp oscillating pulses, despite various possible artefacts (due to patient motions, changing skin-electrode contact, power-line interference, baseline wander...) [Bansal, Khan, and Salhan 2009; Cuiwei Li, Chongxun Zheng, and Changfeng Tai 1995; Saxena, Kumar, and Hamde 2002].

The PPG is another conventional clinical technique, based on a light emitter/receptor (visible to near infrared) on various part of the body. Particularly inexpensive and portable, the fingertip pulse oximeter gained recently a broad popularity for its ability to estimate the blood oxygen saturation level at home. This type of recording is also being integrated into

wearable devices (watch, smartphone, ...) for cardio-respiratory monitoring applications. Its repeating pattern is a simpler and smoother pulse, making it potentially more vulnerable to artefacts.

Infrared thermography is a non-contact technique collecting the thermal electromagnetic radiations in the far infrared wavelengths. From the theory of the blackbody radiation, it images the body surface temperature. In some regions of interest such as vascularized area, the cardiac pulse can be detected as periodic variations in the thermal map. However, it is more expensive and the temporal resolution is limited and more prone to noise. In addition to purely thermal characteristics, a thermogram is also influenced by the spatial motions of the body, whose analysis is left outside the scope of this thesis. Introduced for illustrative purpose in Fig.II.2, a temperature signal derived from an IR thermogram and containing a tenuous cardiac oscillation is also analysed in Appendix 2.

Both ECG and clinical PPG require contact, cables, which may not be adequate in every situation. This was the reason for the development of non-contact techniques such as infrared thermography. Alternative, such as RGB imaging (via a simple camera) or capacitively coupled (distant) electrodes, that especially inexpensive but subject to strong artefacts and require a more complex processing [Kranjec et al. 2014]. The possibility to handle the various forms of recorded cardiac oscillations would extend to new experimental situations, less and less invasive, more and more accessible.

1.2 Motivation: measuring the cardiac variability

Orchestrating the double circulation of the blood to oxygenate tissues, the contractile and rhythmic cardiac activity is adjusting constantly to the needs of organs (e.g. during physical activity), the breathing state (for an optimal gas exchange in lungs), the pressure in peripheral vessels (themselves adjusting by vasoconstriction or vasodilation), etc. In particular, the heart rate is finely regulated by the stimulatory (adrenergic / sympathetic) and inhibitory (cholinergic / vagal or parasympathetic) branches of the autonomic nervous systems, respectively increasing and decreasing it via a combination of electrical, chemical and mechanical signalling pathways [Opie 2004].

The variability of the heart rate has been linked to healthiness, and the lack of it to higher risk of diseases. Moreover, it provides precious insights into the activity of the autonomic nervous system, whose reciprocal action through the sympatho-vagal balance only constitutes a simplified description. In particular, a “fast” variability (at the breathing frequency or higher) is associated to vagal modulations, while a slower variability, possibly coexisting with the faster one, would rather be linked to sympathetic modulations (or both of them). The diversity of these slow and fast modulations of the heart rate are shown in Fig.IV.1 for several subjects. A long-standing interest into this variability produced various metric based on the estimation of the heart rate or period, and their correct quantitative interpretation is still under extensive investigation [de Geus et al. 2019].

While blood pressure (BP) or pulse oximetric (PPG) measurements are well-representative of the peripheral circulation, an electrical recording (ECG) close to the heart provides a direct monitoring of its activity. All these recording methods can serve for the estimation of the heart rate. An ECG is typically found among multiple cardiac recordings, and it provides details on the cyclic activity of the heart; for these reasons, we use it for illustration.

The cycle starts with the electrical depolarization of pacemaker cells of the sinoatrial node, stimulated and inhibited respectively by the sympathetic and parasympathetic nerves. The electrical impulse propagates, see Fig.IV.2 (A), and excites atrial myocytes, provoking the contraction of the two atria. Its front is delayed in the atrioventricular node while the two ventricles fill with blood (diastole). The depolarization wave then rapidly propagates across

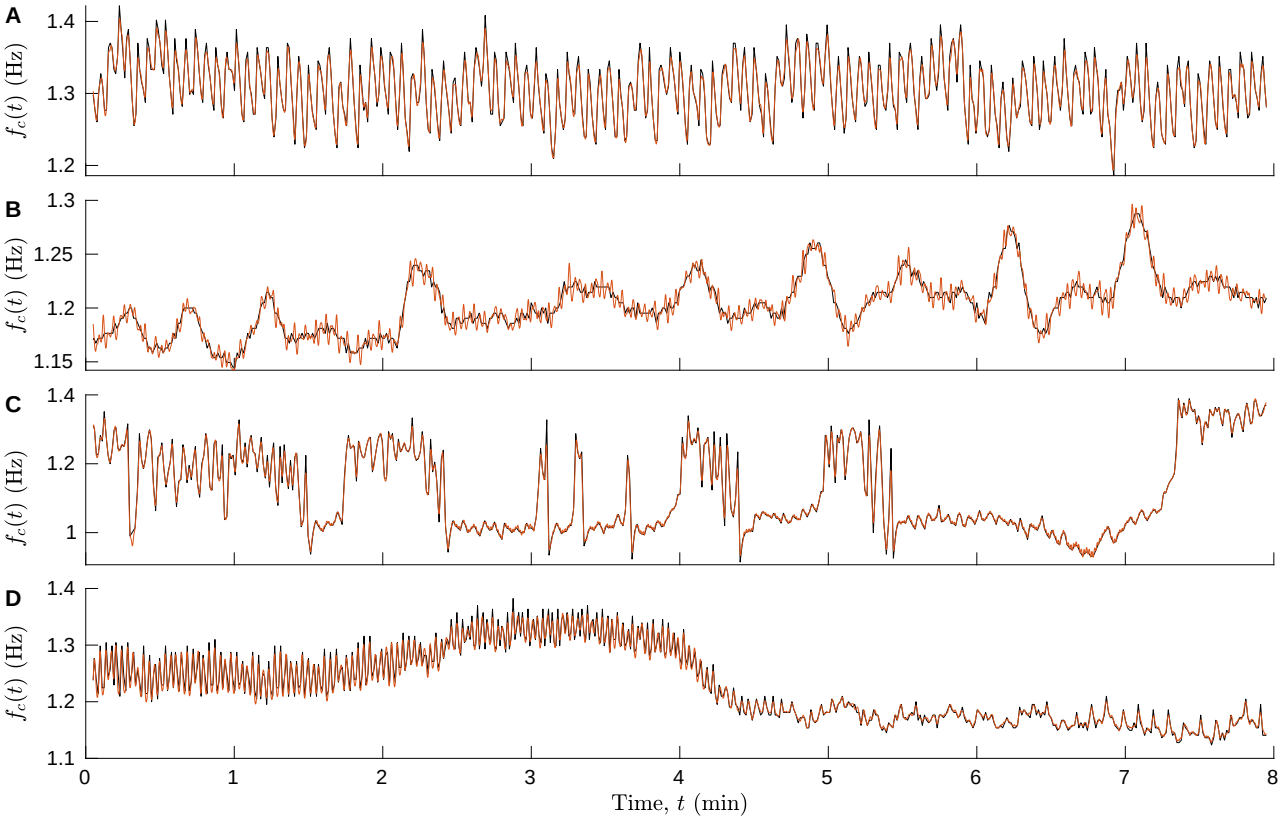


Figure IV.1: *Heart rate variability for 4 different subjects (A, B, C, D) for the capnobase dataset. Here, the cardiac frequency $f_c(t)$ is estimated from ECG recordings. A wavelet-estimator (red line) is compared to the reference estimator from the PhysioNet Cardiovascular Signal Toolbox based on the jqrs algorithm (black line), showing close agreement.*

the ventricles, whose myocytes contract and eject blood out of the heart (systole). The repolarization of cardiac muscle tissues complete the cycle [Opie 2004]. The resulting electromagnetic field can be effectively modelled as the result of a moving dipole, see Fig.IV.2 (B), undergoing a specific cycle both in strength and direction of the associated vector [Mohrman and Heller 2018]. It is measured at the surface of the skin as differences in the electric potential. The particular recorded waveform varies with the placement of the electrodes on the chest, but invariant features in the cardiac pulse allow to identify the phases of the heart cycle.

The extrema of these oscillating patterns are denoted by the letters P, Q, R, S and T, see Fig.IV.3, as alternating local maxima and minima. The P oscillation marks the atrial depolarization, whereas the strong Q-R-S pattern represents the ventricular depolarization, followed by the repolarization during the T oscillation. The detection of the Q-R-S pattern, which is very sharp, is the most common strategy to define the heart period $\tau_c(t)$ as the R-R time interval. As a reference conventional method, we use the *jqrs* algorithm from the PhysioNet Cardiovascular Signal Toolbox [Vest, Poian, et al. 2019], which first compute the Q-R-S intensity using a matched filter (band-pass) and a sequence of signal processing operations [Behar et al. 2014; Johnson et al. 2014]. Then, maxima of the intensity are detected in the resulting positive pulse train from a thresholding procedure. To validate the detection method or reject unreliable intervals, a signal quality index (SQI) can be computed, for instance from a comparison with another detector [Vest, Da Poian, et al. 2018]. An automated error correction, that interpolates missing beats and removes false ones based on physiological assumptions, can improve the estimation of the heart period (N-N intervals for "normal"), so that the adjustment of sufficient detection parameters (such as thresholds) can yield satisfactory results without the need for expert (but less reproducible) modifications.

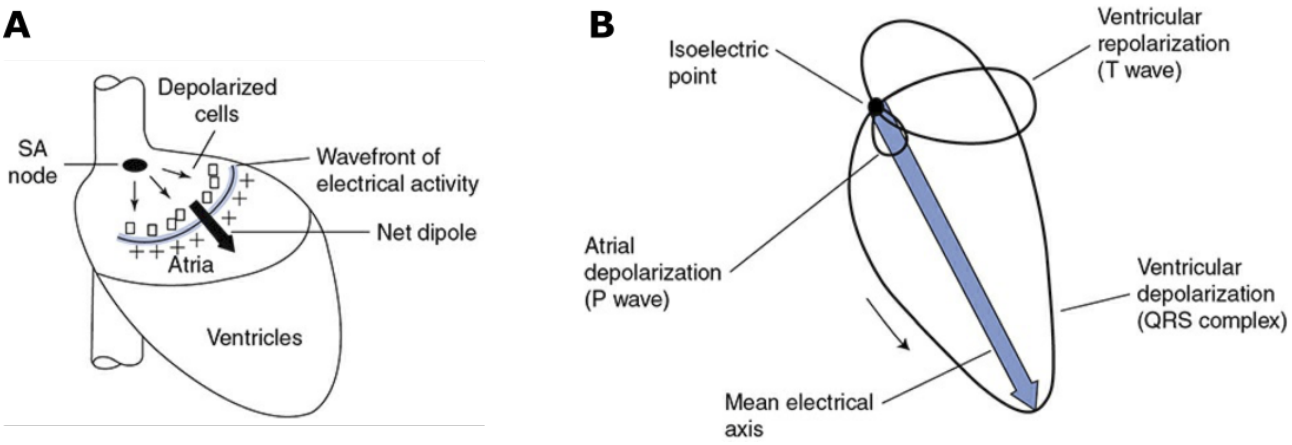
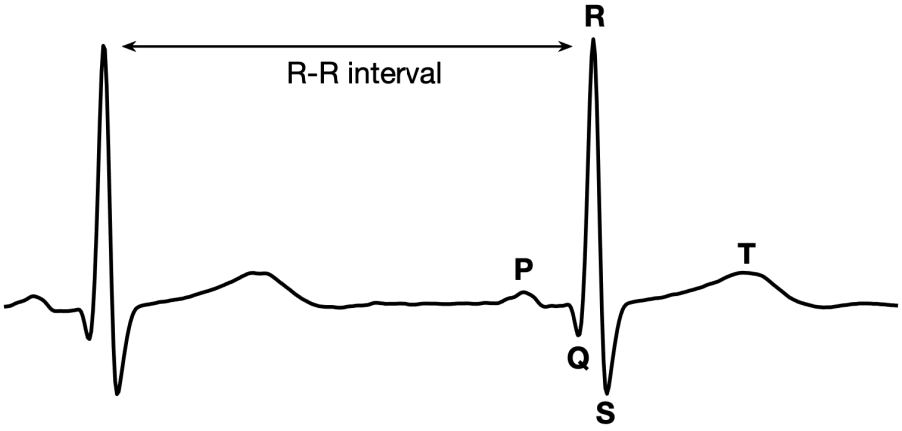


Figure IV.2: *Heart's electrical activity. (A) Pacemaker cells of the sinoatrial node (SA) induce the depolarization of cardiac cells, which initiates the contraction of the heart and generates a net cardiac dipole. (B) Typical vectorcadiogram: this dipole, characterized as a vector, undergoes a cyclic trajectory at each heart beat. Adapted from [Mohrman and Heller 2018].*

Figure IV.3: *Typical waveform of an ECG recorded on the chest: P, Q, R, S and T patterns, and R-R time interval. Depending on the subject and the position of electrodes, the S pattern can be more important than R, the T pattern can be more pronounced.*



1.3 Heart rate: but what is it?

These signals of repeating cardiac patterns have in common a varying periodicity $\tau_c(t)$, the heart period or inter-beat interval, equivalently expressed as its inverse the heart rate or cardiac frequency $f_c(t)$ (c for cardiac). Although intuitive in practice, this apparently simple characterization of the fluctuating cardiac activity can be quite paradoxical from a technical perspective. Is the cardiac rhythm a wave or a pulse train? Frequency or time-localized? We depict this most elementary biological quantity according to the dynamical approach introduced in chapter III, section 1: the heart rate is related to the phase $\phi_c(t)$ of a cyclic dynamics as $\dot{\phi}_c(t) = \Upsilon f_c(t)$. We recall how its physical interpretations, both temporal and spectral, are preserved and captured in the wavelet formalism. Different estimators for the heart rate signal are then compared, from the most generic version to the more sophisticated ones.

1.3.1 Cardiac harmony: beats or waves?

Looking at the BP, ECG and PPG recordings in Fig.IV.4, the signals (A, B, C) could be modelled either as a modulated wave Eq.(III.1), with a steadily growing cardiac phase $\phi_c(t)$ and waveform H, or as a pulse train whose beats $h(t - t_j)$ happen at each time of a sequence

t_j with amplitude a_j:

$$x(t) = A(t)H(e^{i\phi_c(t)}) + r(t) \tag{IV.1}$$

$$x'(t) = \sum_{j=-\infty}^{\infty} a_j h(t - t_j) + r'(t) \quad . \tag{IV.2}$$

Because of its sharp QRS pattern, the ECG signal is often treated as a pulse train. The smoother BP and PPG signals can be more easily associated to a modulated wave. Both descriptions would only coincide for a very regular rhythm, when the signal is close to periodic: $t_{j+1} - t_j \approx \tau_c(t) = \Upsilon/\dot{\phi}_c(t)$, $a_j \approx A(t)$, $h(t - t_j) \approx H(e^{i\phi_c(t)})$, $t_j < t < t_{j+1}$, which is unnatural for the cardiac rhythm (unless with a pacemaker).

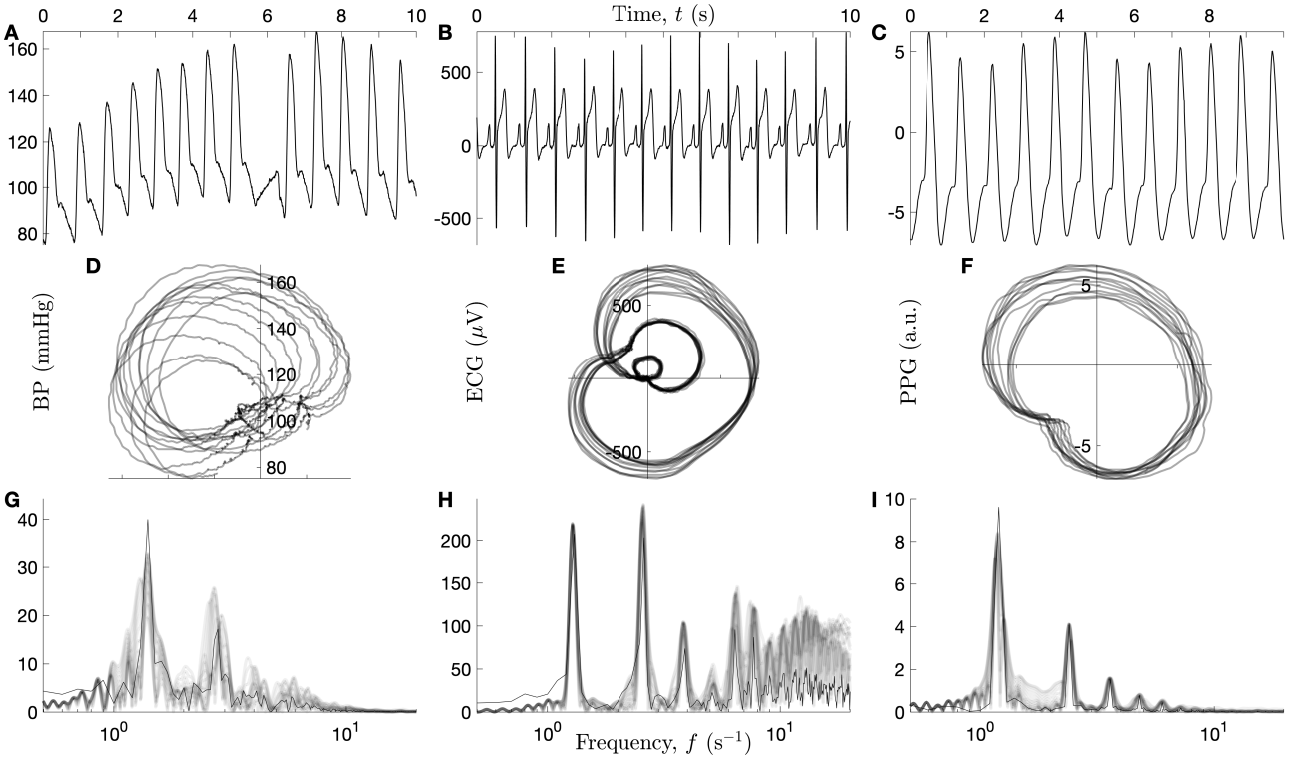


Figure IV.4: *10 seconds of cardiac activity recorded with different measuring techniques: (A, D, G) intra-arterial blood pressure (BP in mmHg), (B, E, H) electrocardiogram (ECG in μV) and (C, F, I) photoplethysmogram (PPG in arbitrary unit). Signals $x(t)$ on selected intervals for subjects 3 (A, B) and 2 (C). (D, E, F) Nearly cyclic trajectory of twice the analytic signals $2x^+(t)$: its real part ($x(t)$) is plotted against its imaginary part (the Hilbert transform of $x(t)$); positive rotation is here clockwise. (G, H, I) Spectral amplitude on the selected intervals of duration $T = 10$ s, computed in the signals' unit as twice the modulus of the Fourier transform $2|\hat{x}(f)|/T$ (thin black lines) and wavelet transform $2|X(t, f; Q)|$ with $Q = 30$ (thick light grey lines, stacked for each sample of t).*

In spite of these different interpretations, the cardiac dynamics undergoes a cyclic trajectory. The orbits are well visible, as represented in Fig.IV.4 (D, E, F) in the state space made of the real and imaginary parts of the analytic signal $x^+(t)$, Eq.(I.2). Hence, a dynamical cardiac phase $\phi_c(t)$ that parametrizes its cyclic orbit can also be associated to the ECG signal [Kralemann, Cimponeriu, et al. 2008], see also the discussion in chapter III, section 1.1. Sign of the inexact periodicity of the physiological rhythms, the orbits do not collapse on a single closed curve, due to amplitude modulations (clearly of respiratory origin for the ECG (B) and the PPG (C)), and from the presence of a varying trend, see BP in (A, D).

Note that a clean orbit can correspond to a rhythmic signal of slowly varying period but with constant waveform. When the amplitude also varies slowly, the angle of $x^+(t)$ still coincides with the genuine dynamical phase $\phi_c(t)$ in the case of a slowly modulated simple wave (sine or complex exponential model of chapter I, section 1.3). As can be observed in (C, F), the PPG signal is centred and approximately circular so that its angle effectively counts cycles. Even though it differs from $\phi_c(t)$, the angle $\Im\{\log x^+(t)\}$ estimates it in a simple way, so that it has been called a *protophase* [Kralemann, Cimponeriu, et al. 2008]. Similarly, a protophase may be constructed from the analytic BP signal; however, it requires a preliminary detrending (low-pass filtering) to centre the orbit.

Unluckily, the ECG signal, is too far from circular (E) / sinusoidal (B) to use its angle directly as a protophase: it winds several times per cycle. Yet, its specific PQRST local extrema each correspond to quite precise values of the dynamical phase, especially RS extrema which are the sharpest, whose detections yield a time sequence t_j such that $\phi_c(t_{j+1}) = \phi_c(t_j)$. Since the phase increases steadily within each cycle, it can be estimated by interpolation, and the same for the heart period $\tau_c(t)$ estimated from the discrete inter-beat intervals $t_{j+1} - t_j$. The difficulty to estimate the cardiac phase $\phi_c(t)$ from a wave model of the ECG is commonly called a "non-linearity" issue, abstractly referring to the waveform function $H(z)$ (in fact non-circular as $z = e^{i\phi_c(t)}$), and calls for the spectral distinction of circular components z^n of different orders n.

On a time interval containing few cycles, a cardiac signal is periodic enough for its spectrum to be harmonic, Eq.(III.3): it is decomposed into circular components, cycling with frequencies $nf_c(t)$. The typical heart rate in this interval is found straightforwardly at the frequency of the fundamental component, i.e. the first peak in the harmonic sequence. Each peak is associated to a Fourier coefficient, encoding the pulse or the waveform. For all recordings, the discrete harmonic structure with distinct orders n are easily recognizable in the amplitude spectra (G, H, I) of Fig.IV.4, despite an important perturbation in the case of the BP (G) (partly due to a missing beat (A) that compromises the periodicity). For the PPG (I) and the BP (G), the amplitude of the harmonics is decaying fast at increasing orders n, reflecting the weak non-circularity of these cardiac oscillations. In contrast, the second harmonic of the ECG (H) has a higher amplitude than the fundamental (first) one and numerous orders are visible, sign of the strong non-circularity of the ECG oscillation.

In the amplitude spectra (G, H, I), the amplitude of the wavelet transform matches the one of the Fourier transform because the wavelet lasts as long as the selected interval (with $\frac{Q}{\sqrt{\pi}} \approx 12$ oscillations). The factor 2 in their definition allows comparing directly their peak value to the amplitude of the corresponding harmonic wave in the signal. For longer durations, the Fourier amplitude gets more and more noisy, because of rhythmic fluctuations, so that the estimation of the power spectrum is preferred. The amplitude spectrum is estimated from the mean-squared modulus of the wavelet transform, similarly as the power spectral density $S_{xx}(f;Q)$, see Eq.(II.23). Their explicit relation is given here, for a signal $x(t)$ and with a quality factor Q:

$$2\langle|X(t,f;Q)|^2\rangle_t^{\frac{1}{2}} = 2\Big(\frac{\sqrt{\pi}}{Q} S_{xx}(f;Q)|f|\Big)^{\frac{1}{2}} \quad . \tag{IV.3}$$

In the following, amplitude spectra are preferred to power spectra for their oscillatory interpretation (instead of a stochastic one): the height of its frequency peaks directly provides a mean amplitude (radius) of the corresponding harmonic oscillation in the signal. The time-dependent amplitude is provided by twice the modulus of the (unaveraged) wavelet transform.

1.3.2 Wavelet strategy and selection of the quality factor

The ability to discriminate slow and fast oscillatory contributions in the frequency domain comes at the cost of losing the time dependence of the identified rate in the analysed portion of recording. This can be recovered from a time-frequency approach, as hinted in Fig.IV.4 (G, H, I), where the Fourier spectra restricted to a 10-second interval (thin black lines) superimposed to the time-varying wavelet spectra at a high quality factor ($Q = 30$), stacked for each time sample (thick light grey lines). The spread in frequency and amplitude of such a "persistence spectrum" directly represents the temporal variability of the cardiac rhythm. This is shown later with greater details on a longer 1-min ECG recording in Fig.IV.9 (C), together with the amplitude of the wavelet transform $X(t, f; Q)$ (B) with $Q = 8$. At this lower quality factor, less harmonic components are visible, whereas distinct sharp impulses are well represented. The choice of the quality factor is crucial in the time and frequency interpretations of the wavelet transform, through its relation to the time-frequency resolution $f\delta t = (\delta \log f)^{-1} = \frac{Q}{\sqrt{\pi}}$, Eq.(I.41).

Consider Fig.IV.5 where the same cardiac activity, simultaneously recorded from a PPG (A, B) and an ECG (C, D), is represented in the time-frequency domain using two different quality factors. As a first remark: all panels contain the cardiac rhythm as high amplitude (yellow) structures, that are mostly horizontal (waves) or vertical (pulses). An acceleration of the heart beat is readily observable as an increase in pulses density or in wave frequency. These slender regions of high amplitude repels zeros (dark blue singular points) that accumulates in interstices, adopting the same horizontal or vertical alignment.

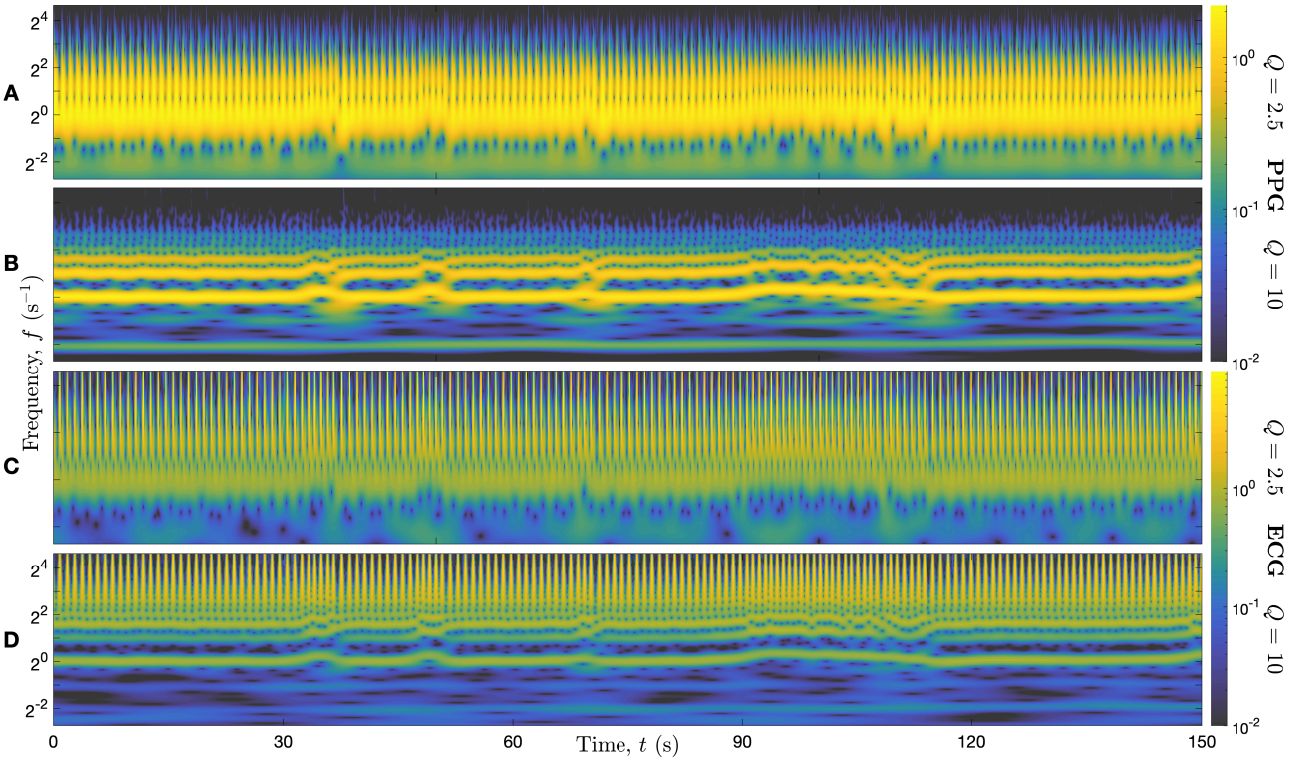


Figure IV.5: *Time-frequency representation of a PPG (A, B) and an ECG (C, D), simultaneously recorded on subject 1, computed with different quality factors: $Q = 2.5$ (A, C) or $Q = 10$ (B, D). Colours code for the amplitude, which is twice the modulus of the wavelet transform, of signals' unit, arbitrary here.*

To resolve the heart rate variability down to the scale of a single oscillation, the time resolution should be $f\delta t \sim 1$, so that the quality factor would be at most of the order of $Q \sim \sqrt{\pi} \approx 2.5$. This is illustrated in panels (A, C) where the cardiac rhythm is essentially composed of high amplitude time-localized pulses, joining at the cardiac frequency. Yet, it is

insufficient to exhibits the harmonic structure of the spectrum: the order n harmonics can be distinguished from the higher orders for at most $\delta \log f = \log \frac{n+1}{n}$, so that at least $Q \sim \frac{\sqrt{\Gamma}}{\log 2} \approx 3.6$ is required to separate the fundamental component from the next harmonics. In panels (B, D), we can verify that the quality factor $Q = 10$ is sufficient to distinguish clearly the 3 first harmonics from higher orders, in accordance with the spectral interpretation.

A similar reasoning from the time domain goes as follows: at least $f\delta t = n$ wavelet oscillations are required (fitted in a period of the rhythm) to resolve the order n harmonics, so that we start to resolve the second harmonics at about $Q \sim 2\sqrt{\Gamma} \approx 5$, and the fourth harmonics at $Q \approx 10$, see panels (B, D) of Fig.IV.5. Quick variations of the cardiac rhythm compared to the resolution $f\delta t \approx 4$ results in interferences in the harmonic lines. These spectral and temporal perspectives are consistently interlaced from the harmonic ordering: $n < (\log \frac{n+1}{n})^{-1} < n + 1$ Eq.(III.5).

Most of the rhythmic amplitude is found in the first harmonics for the PPG, whereas it is concentrated in pulses in the ECG. The advantage of the wavelet representation is the variety of approach that are available to study the cardiac rhythm: either temporal via the phase or pulse extraction, or spectral via frequency peak detection depending on the quality factor. Whether the cardiac rhythm is circular or pulsatile, a small quality factor ($Q \approx 5$) is beneficial compared to a higher one to resolve its fast temporal variability, see Fig.IV.6 (A, B). The spectral perspective (C) is not adequate in this context to estimate the heart rate because of the lack of frequency resolution, unless the frequency-reassignment Eq.(III.15) is used; based on the time-derivative of the phase, this is equivalent to a temporal approach.

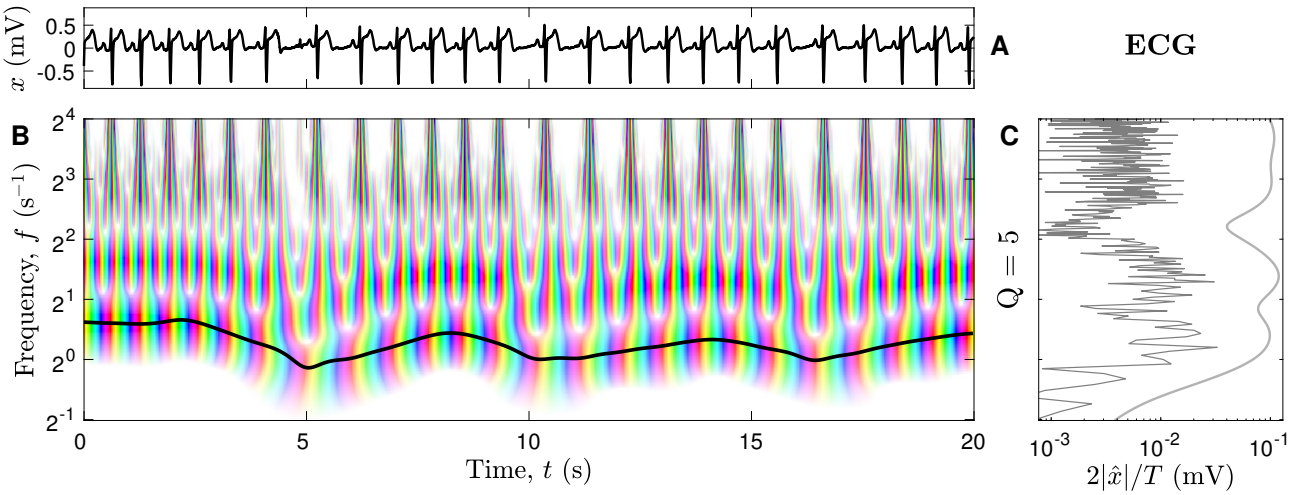


Figure IV.6: *ECG recording of subject 3 and cardiac frequency. (A) Signal on a selected interval of high variability of the cardiac frequency. (B) Time-frequency representation of ECG, computed with the quality factor $Q = 5$. Colours code for both the angle (hue) and the modulus (saturation) of the wavelet transform (see Fig.II.1 of chapter II). (C) Amplitude spectra obtained either from the Fourier transform (thin line) or from the wavelet transform (thick line) Eq.*(IV.3). *The cardiac frequency $f_c(t)$ is represented in (B) as a thick black line, superimposed to the fundamental cardiac component.*

However, the spectral distinction of the fundamental component from higher harmonic orders at each time opens the possibility to relate heart rate and cardiac recording in a very elementary way: it is the time-derivative of the angle of the fundamental component. Thus, the cardiac frequency $f_c(t)$ is identified with its stable ridge, defined in Eqs.(III.12, III.13) and illustrated for the ECG in Fig.IV.6 (B). Compared to the instantaneous frequency of the analytic signal Eq.(III.7), the spectral decomposition relieves the estimation of the cardiac phase from the activities at faster or slower scales. While the separation from higher harmonics reduces

Generic formula	*wide-band*	*narrow-band*	*approximation*
$\frac{\langle \dot{X} \rangle}{i\Upsilon\langle X \rangle}$	$F_x(t)$	$F_x(t,f;Q)$	$f_1(t;B)$
$\frac{\langle \dot{X}\overline{X} \rangle}{i\Upsilon\langle \lvert X\rvert^2 \rangle}$	$F_x(t;Q)$	$F_x(t,f;Q_+,Q_-)$	$f_2(t;B)$, $F_x(t,f;nQ,Q)$
Equation	III.7, III.28	III.9, III.29	III.25, III.27, IV.4

Table IV.1: *Summary of wavelet-estimators for the complex rate F_x of a real signal $x(t)$, as developed in chapter III. The real part $\Re\{F_x\}$ is interpreted as the frequency modulation and the imaginary part $\Im\{F_x\}$ is related to the amplitude modulation. The first row contains linear estimators, which can be numerically unstable (divergent), and their quadratic and stable counterpart are in the second row. Quadratic estimators reduce to linear ones in the limit cases $Q = 0$ (wide-band) or $Q_+ = Q_- = Q$ (narrow-band). Narrow-band estimators reduce to wide-band ones in the limit $Q, Q_- \to 0$, losing their frequency localization. In the generic formula, $\langle\rangle$ refers to a band integration (wide or narrow) of the wavelet transform $X(t,f;Q)$.*

non-circularity, slow perturbation of the cyclic orbit are also cleaned, such as the influence of the breathing rhythm, that can be noticed around 0.25 Hz in the PPG for $Q = 10$, Fig.IV.5 (B).

This quite universal approach, however, finds practical limitations when most of the rhythmic intensity settles out of the fundamental component, as is the case for some ECG recordings where a tenuous fundamental component competes with a stronger second harmonics. The extraction of the ridge for the fundamental component, can also be fragmented under the influence of an important noise level, complicating further signal analysis. We address these difficulties and their practical treatment in the next section.

1.4 Extracting the heart rate variability

There exists a huge number of processing method to estimate the heart rate, which are highly dependent on the recording technique [Bansal, Khan, and Salhan 2009; Elgendi et al. 2016; Kranjec et al. 2014], usually separated into temporal and spectral approaches. We have discussed above how these approaches meet in the time-frequency plane, and how the heart rate is captured by the wavelet transform, whatever the recording technique. It remains to extract from it a new signal that characterizes the fluctuations of the cardiac rhythm, essential to further study the heart rate variability in relation to other physiological activities.

Therefore, we put into practice the wavelet framework for rate estimation developed in chapter III section 3 to cardiac recordings. It contains a generic family of rate estimators, summarized in Table IV.1, that are not specific to the waveform, yet letting room for a fine-tuning through the selection of a quality factor and/or a band of interest. We apply them, from the most simple to the most sophisticated version, on cardiac signals of the *capnobase* dataset, for which both the ECG and PPG recording are available for each subject. We also extract from the ECG a reference cardiac frequency using the conventional *jqrs* algorithm from the PhysioNet Cardiovascular Signal Toolbox; it will serve to compare the outcome of the different estimators and assess this generic wavelet approach.

1.4.1 Paradox of the frequency localization

In order to extract the cardiac frequency, we apply the temporal strategy to the wavelet transform $X(t,f;Q)$ of a cardiac recording. It consists in computing the instantaneous frequency

$F_x(t, f; Q)$, see Table IV.1, which is detailed at each location in the time-frequency plane. Contrary to the purely temporal approach, $F_x(t)$, the frequency parameter allows separating fast or slow perturbations, but also to break strong non-circularities into harmonics and pulsatile contributions. That is beneficial for the estimation of the fundamental cardiac phase, but it introduces the frequency variable f. To examine the fundamental cardiac component only, it should be chosen equal to the cardiac frequency $f = f_c(t)$, that we try to estimate. Because of this paradoxical situation, it is in general difficult to extract the time-dependent frequency of the recorded rhythm without some assumptions about its frequency localization.

This is precisely what is done in the reference cardiac frequency estimation algorithm from the PhysioNet toolbox, which assumes the existence of the rhythm in a physiologically likely band (0.375 to 2 Hz by default), allowing to discard unlikely R-R intervals and interpolate holes. Although not generic, this assumption is relevant to the cardiac rhythm and related to the one of homeostasis.

Let us recall elementary time-frequency notions, and their limitations in estimating $f_c(t)$. When we only need to plot (but not to extract) the heart rate signal $f_c(t)$, it might be sufficient to construct the frequency-reassigned time-frequency power density Eq.(III.17), that concentrates the intensity near time-frequency lines $f(t)$ —a single one of which is the heart rate $f_c(t)$— facilitating its visualization. These lines satisfy the phase ridge conditions, in particular $\Re\{F_x(t, f(t); Q)\} = f(t)$. When the extraction of $f_c(t)$ is required for further processing, however, this self-consistent expression is hard to use for several reasons: the ridge of interest can be fragmented into a set of non-connected lines $\{f_k(t)\}_{k=1,2,...}$ (especially for a small quality factor), simultaneous overlapping lines coexist at different frequencies that need to be sorted, and off-ridge information is not exploited, which inefficiently requires a very precise f-sampling, or iterative methods [Delprat et al. 1992].

1.4.2 Practical workarounds

This inspired us related wavelet-estimators for the cardiac frequency, that rely on the previous knowledge of a frequency band B containing the cardiac frequency fluctuation around a typical value f_c, near 1 Hz for most subjects. From the rough frequency localization in this band, a greater precision is retrieved taking advantage of the squeezing effect against ridges. This effect can be observed in Fig.IV.7 (B): curves $\Re\{F_x(t, f; Q)\}$ are plotted for fixed values of f (regularly sampled) and concentrate around particular lines, such that the fundamental cardiac frequency $f_c(t)$ near 1 Hz. The band B of attraction around it, and its typical frequency f_c can be anticipated from small values of $\Im\{F_x(t, f; Q)\}$, as shown in Fig.IV.7 (C).

The instantaneous frequency of the analytic signal reconstructed in this band B is the real part of the linear estimator $f_1(t; B)$ (see Table IV.1), that can yield a very precise estimation of $f_c(t)$. However, the resulting estimator is not different from (and approximately related to) $F_x(t, f_c; Q_-)$, where $Q_- < Q$ corresponds to the increased bandwidth of B. On long recordings, this estimator suffers from occasional numerical divergences. An instability is caused by the passage of the fixed lines $f = f_c$ close to a phase vortex (singular point), supposedly repelled out of the rhythmic components, see Fig. IV.6.

Its quadratic counterpart $f_2(t; B)$, interpreted as an average of $F_x(t, f; Q)$ in the band B weighted by the intensity $|X(t, f; Q)|^2$ (vanishing at singular points) is found to circumvent this instability, due to its approximate relation to a canonical quadratic estimator $F_x(t, f; Q_+, Q_-)$. Applied to the ECG and PPG recordings of Fig.IV.8 (A, C), it follows closely and smoothly (red lines in B, D) the reference heart rate extracted from R-R intervals (black lines), up to a small delay in panel (B) of physiological origin: there is a transit time from the electrical impulse in the heart (ECG) to the oxygenation impulse in the fingertip (PPG).

Estimators in the wide-band limit are interesting when the band B of interest is unknown:

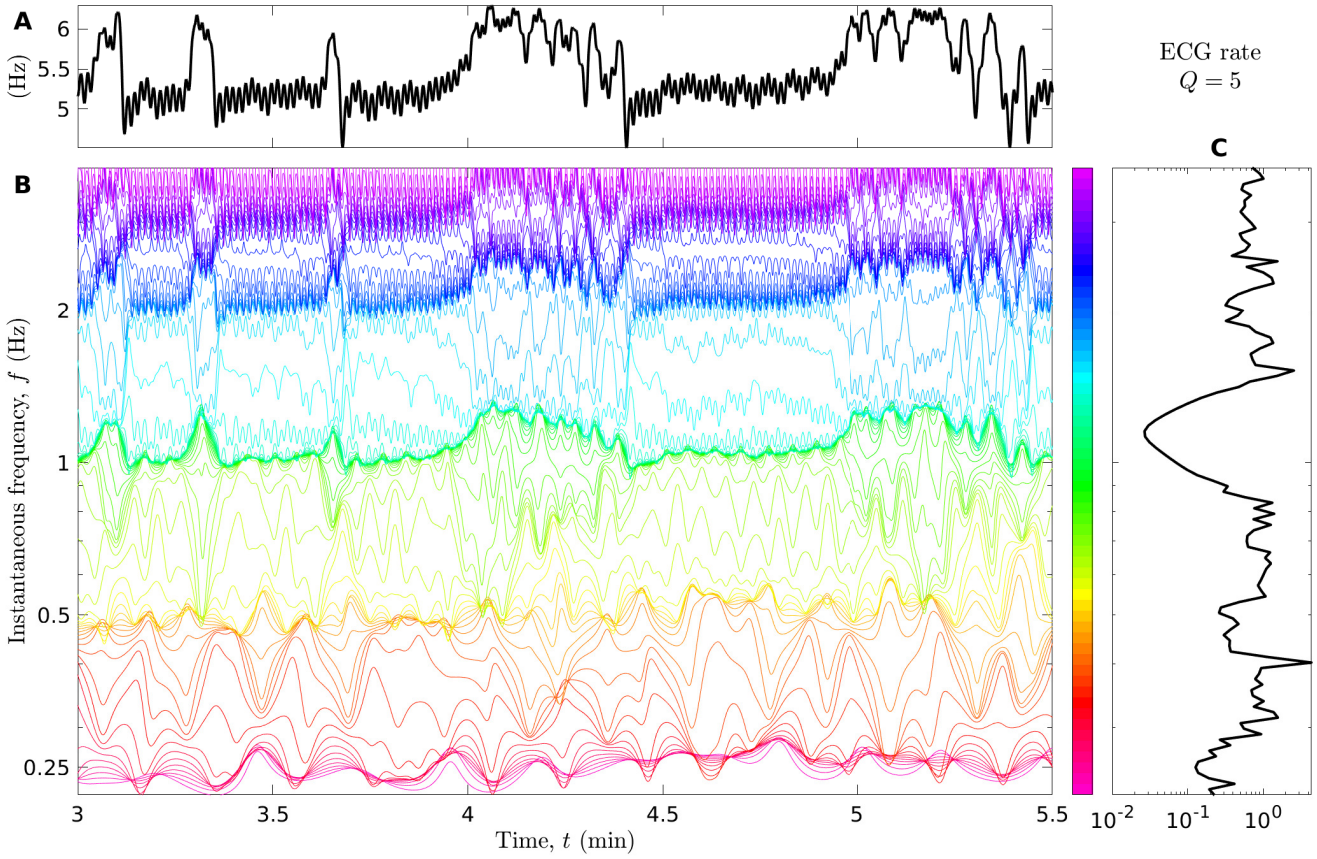


Figure IV.7: *Instantaneous frequency estimations from the ECG of subject 1, for $Q = 5$. (A) Wide-band estimator $\Re\{F_x(t;Q)\}$, which is the weighted average of (B) all curves $\Re\{F_x(t,f;Q)\}$ for fixed values of f, regularly sampled, each corresponding to a distinct colour. (C) Typical value (root-mean-square) of $\Im\{F_x(t,f;Q)\}/f$. For an easier visualization, oscillations faster than 1 Hz have been smoothed in (A, B).*

with fewer parameters, they are the most generic. In particular, the quadratic wide-band estimator $F_x(t;Q)$ (blue lines) reproduces very well the cardiac frequency modulations from both the PPG and ECG signals, up to an obvious bias. Caused by the weighted average of all components, from fundamental to harmonics, the estimated rate signal is necessarily an overestimation of the cardiac (fundamental) frequency. However, this bias is nothing but a coefficient, that is approximately constant because of the high regularity of the cardiac waveform (contrary to the voice vibrato in Fig.III.5). On a log-frequency scale, this proportionality is visible as shift from the reference heart rate in panels (B, D), small for the quite circular PPG and large for the strongly non-circular ECG.

For comparison, we also plot the even simpler (linear and wide-band) estimator $F_x(t)$ (green line): without any parameter, it can be directly computed from the analytic signal. Although the outcome could be acceptable when applied to the PPG, the signal extracted from the ECG are both biased and inaccurate, with clear instabilities. In comparison, $F_x(t;Q)$ is much more robust in estimating modulations of the cardiac frequency, in spite of a systematic bias. Note that such constant coefficient is ignored in a correlation or coherence analysis.

1.4.3 Non-circular modulations and time-smoothing

An important feature of these estimators has been hidden in Figs.IV.7, IV.8: they all oscillate at the cardiac frequency (~ 1 Hz), sometimes tremendously, so that we have attenuated it (with a smoothing) for an easier visualization and interpretation. Especially strong for small quality

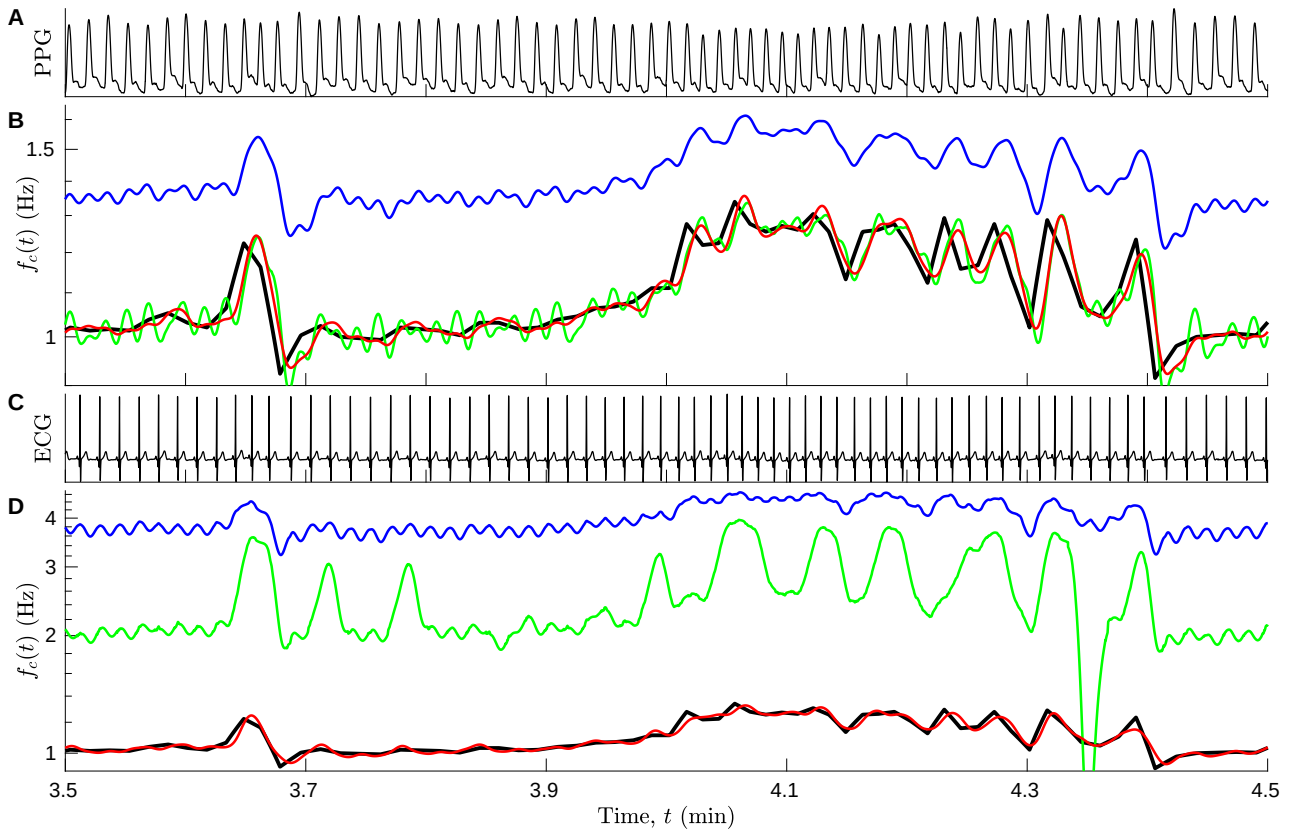


Figure IV.8: *Modulations of the heart rate of subject 1, obtained either from the PPG signal (A) or from the ECG signal (C), using different wavelet estimators for the cardiac frequency (B, D). The estimators are the real part of $F_x(t)$ (green lines), $F_x(t;Q)$ for $Q = 5$ (blue lines) and $F_x(t, f_c; Q_+, Q_-)$ for $Q_- = 2.5$ and $Q_+ \approx 3.5$ (red lines). Sign of non-circularity, strong oscillations shorter than 1 s are smoothed for a clearer comparison to the reference heart rate (black lines in both panels B and D), estimated from R-R intervals of the ECG from the jqrs algorithm. A short delay is visible between PPG and ECG estimations (B), this is the pulse transit time.*

factors or broad band, it is caused by the non-circularity of oscillations due to interfering and beating harmonics. The frequency modulation estimator translates this effect as a very fast modulation at the carrier cardiac frequency $f_c(t)$.

These very fast modulations induced by the non-circularity of the cardiac waveform (especially for the ECG and at low quality factor) are an important limitation when trying measuring the precise value of the heart rate $f_c(t)$ at a certain time. However, it is superimposed to the slower and legit cardiac modulations with an important separation of scales. From a time-frequency perspective this repetition of the carrier wave inside the estimator enriches it beyond the spectral range of the cardiac modulations.

The estimator that is the least affected by this effect is the quadratic and narrow-band estimator $F_x(t, f; Q_+, Q_-)$, approximated from a short time-smoothing (instead of band-integration):

$$F_x(t, f; nQ, Q) = \frac{S_{\dot{x}x}(t, f; nQ, Q)}{i\mathcal{T}S_{xx}(t, f; nQ, Q)} \quad , \tag{IV.4}$$

from the practical power density estimator defined in Eq.(II.71). For important values of the time-smoothing parameter n, the very fast modulations are perfectly removed, however, the fast but legit time modulations of the heart rate are smoothed. A compromise is found taking values as small as $1 < n \leq 2$, sufficient to avoid instabilities.

The precision of the resulting estimator of the heart rate as $f_c(t) = F_x(t, f; nQ, Q)$ is quite lenient on the choice of a fixed value $f = f_c$, as long as $f_c(t)$ fluctuates around it within a bandwidth. This observation helps to choose the quality factor: on one hand the estimator with a small values of Q is both more tolerant and better resolved in time. On the other hand, the distinction of the fundamental component from the next harmonics value for $Q > \frac{\sqrt{\top}}{\log 2} \approx 3.6$ is useful to reduce very fast modulations induced by non-circularity, allowing a heart rate fluctuations on about an octave, $\delta \log f = \log 2$ (even more when the signal-to-noise ratio is important).

In practice, we find that $Q = 5$ is a good compromise, and that a fixed value of $f = f_c \approx 1$ Hz is sufficient in most cases for the cardiac rhythm. We can adapt f_c for each subject from a preliminary spectral estimation as the fundamental peak of $S_{xx}(f)$, or from the minimum of $\langle |\Im\{F_x(t, f; Q)\}|^2 / f \rangle$, as observed in Fig.IV.7 (C). For polysomnography signals, we have not felt the need to adapt the frequency variable at different times as $f = f(t)$ for the case of important heart rate variation (e.g. from simple to triple). This can be achieved extracting more specific information from the recording but also introducing additional sophistications.

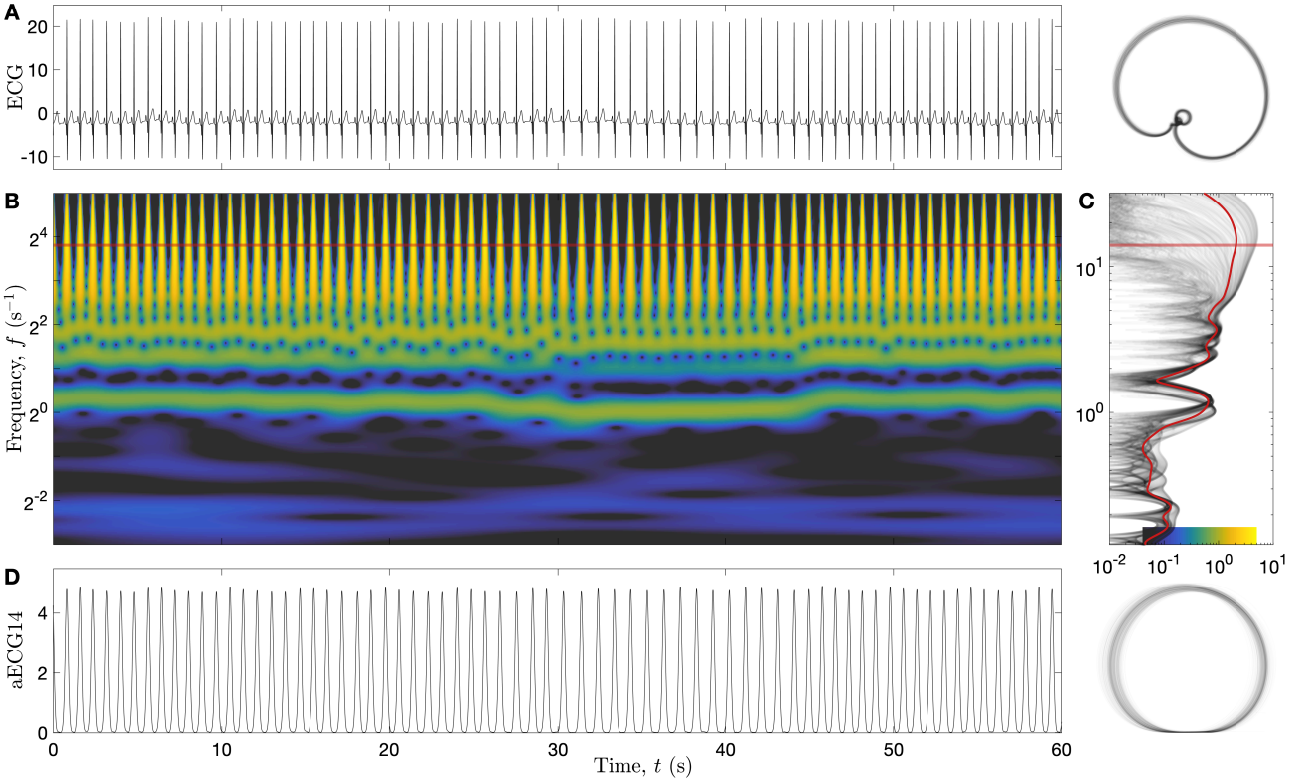


Figure IV.9: *1 minute of ECG, recorded on subject 1, and its narrow-band amplitude at 14 Hz, aECG14. (A) ECG signal $x(t)$ (arbitrary unit) and its cyclic trajectory. (B) Wavelet transform for $Q = 8$, the colour codes for the amplitude $2|X(t, f; Q)|$. (C) Amplitude profiles stacked for each time sample in the selected interval (light grey lines) into a persistence spectrum, and amplitude spectrum (thick red line) Eq.*(IV.3)*. The colour scale is aligned with the amplitude of the spectra, which are all in the signal unit. (D) aECG14 signal $x'(t) = 2|X(t, f; Q)|$ at $f = 14$ Hz (red line in panel B) and its cyclic trajectory.*

1.4.4 Change of input signal: strengthening the fundamental

Another type of sophistication has proved useful to the extraction of the heart rate from the ECG and more generally the rate of a strongly non-circular rhythmic oscillation. It aims at avoiding cases for which the fundamental component in the ECG recording has a weak or even

vanishing amplitude, while harmonics contains most of the rhythmic intensity. This typically happens when the R and S peaks, positive and negative, are comparable. The workaround consists in analysing another signal $x'(t)$ derived from the ECG recording $x(t)$ for which the fundamental component is important (i.e. more circular than the ECG), with the same method otherwise. We considered several quantities, such as the power of the ECG, $P_{xx}(t)$, or its magnitude $\frac{1}{2}\log P_{xx}(t)$.

Our best results have been achieved by reverse-engineering of the reference heart rate estimation from the PhysioNet toolbox: the *jqrs* algorithm is indeed harvesting the cardiac pulse intensity into a new (positive) signal, from which peak locations t_n can be more easily extracted [Behar et al. 2014; Vest, Poian, et al. 2019]. This preprocessing operation is made of a band-pass filtering of the ECG, then turned into a positive signal that represents the envelope of the Q-R-S complex. The filter is best matched by a log-normal wavelet centred at $f = 14$ Hz and with quality factor $Q = 8$. This corresponds to considering the amplitude of the ECG at $f = 14$ Hz, $x'(t) = 2|X(t, f; Q)|$. Denoted *aECG14*, this signal is illustrated in Fig.IV.9. The usefulness of the wavelet transform has been recognized in this particular task [Cuiwei Li, Chongxun Zheng, and Changfeng Tai 1995; Saxena, Kumar, and Hamde 2002].

Compared to the power signal $P_{xx}(t)$, that is wide-band, the narrow-band amplitude signal $x'(t) = 2|X(t, f; Q)|$ focuses specifically on the fast Q-R-S oscillations, filtering out contributions from other frequency bands. The fundamental component of the cardiac rhythm in this new signal is indeed more reliable than the one of the raw ECG, as confirmed by its better circularity (see panel C). This is verified statistically in Fig.IV.10, where we compare the error (in %), relative to the reference heart rate, of the (narrow-band, quadratic) estimator Eq.(IV.4) applied to the ECG, $F_x(t, f; nQ, Q)$, to the same estimator applied to the aECG14, $F_{x'}(t, f; nQ, Q)$. Applied to all subjects in the *capnobase* dataset and all time-intervals in which the signal quality index (hence the quality of the reference estimation) is maximal (SQI = 1), the error is smaller than 0.3% half of the time (the interquartile range is 0.057%) with aECG14, while it is twice higher and with a slight bias (the median is 0.08%) directly from the ECG.

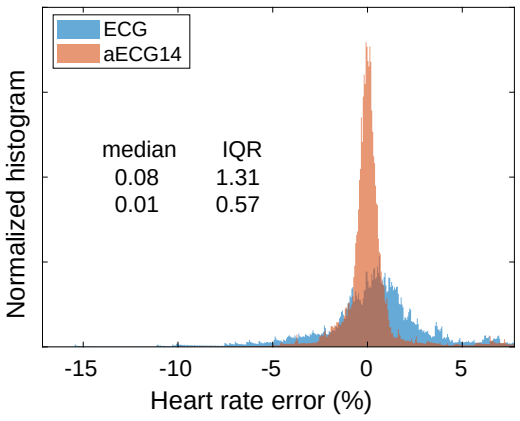


Figure IV.10: *Comparison of the precision of the heart rate, when estimated directly from the ECG or from its amplitude signal at 14 Hz. Histogram of the error, $f_{est}/f_{ref} - 1$ (in %), between the narrow-band quadratic estimator $f_{est}(t)$ based on either the ECG (blue) or the aECG14 signal (red), and the jqrs reference $f_{ref}(t)$, over all subjects in the ppg dataset and all times of maximal signal quality index. Median and interquartile range (in %).*

In the following, we simply refer to this sophistication of the heart rate estimation as the *precise* estimator, that allows to read the precise cardiac frequency at each time. This precise estimator has been illustrated at the beginning of this chapter in Fig.IV.1 (red) against the reference heart rate (black). We can observe that the slight remaining error consists in insufficiently sharp extreme variations, related to the time-smoothing in Eq.(IV.4). The worst estimation in Fig.IV.1 (B) is still closely following the reference, in spite of an important oscillation at the cardiac frequency. This perturbation is caused by an especially prominent T wave in the ECG, that induces a strong second harmonic, in competition with the fundamental component, hence increasing non-circularity. In contrast, we observed subjects for which this precise estimator is able to follow very sharp variations of the heart rate, while the estimation from PhysioNet toolbox incorrectly judges it artefactual and interpolates through it.

The change of input signal allows taking advantage of the pulsatile (vertical) nature of the ECG. We have also applied this method in the low signal-to-noise context of a temperature signal derived from IR thermography in Appendix 2, for comparison with a converse high quality factor perspective called the cepstral approach. This alternative approach is based on the detection of spectral periodicity between the (horizontal) harmonic lines, omnipresent in the Gabor paradigm, and used in state-of-the-art heart rate extraction methods [Cicone and Wu 2017; Li et al. 2019; Lin, Su, and Wu 2018; Su and Wu 2017].

1.4.5 Comparison of the estimators against the reference as a coherence

The comparison of the other generic estimators cannot be compared directly to the reference heart rate signal as done in Fig.IV.10, either because of the delay between the PPG and the ECG for PPG-based estimations, or because of their important non-circularity-induced fast oscillations and/or important bias, see Fig.IV.8. For this reason, we do not refer to them as heart rate estimators, but rather as heart rate modulation estimators. Indeed, they are able to capture well its variability, as verified at each modulation frequency from a coherence analysis.

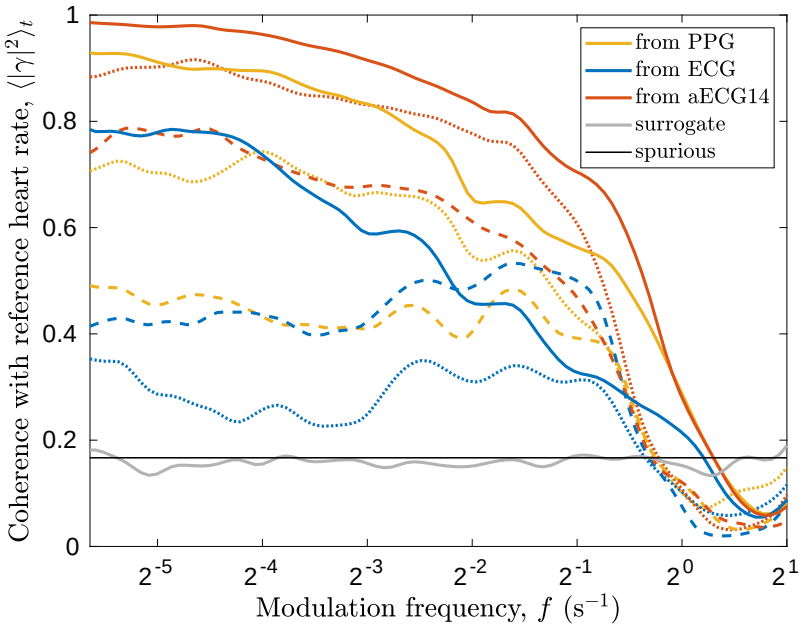


Figure IV.11: *Comparison of wavelet estimators $x(t) = f_{est}(t)$ with the reference cardiac frequency $y(t) = f_{ref}(t)$, in terms of the coherence $\gamma_{xy}(t, f; Q_+, Q_-)$, $Q_+ \approx 5Q_-$ and $Q_- = 5$. For all subjects in the PPG database for which the ECG recording is of good quality ($SQI > \frac{1}{2}$ at all times), the typical squared coherence value is obtained from the average over all cumulated times: $\gamma^2_{xy}(f) = \langle|\gamma_{xy}(t, f; Q_+, Q_-)|^2\rangle_t$. The line styles refer to the wavelet estimators $f_{est}(t)$: wide-band for $F_x(t)$ (dotted line) and $F_x(t; Q)$ (dashed line), and narrow-band for $F_x(t, f_c; Q_+, Q_-)$ (plain line). The colours refer to the input signals x: PPG (yellow), ECG (blue), aECG14 (red), white Gaussian noise surrogate (grey). The thin black line is the expected level of spurious coherence.*

Therefore, we evaluate the time-frequency coherence $\gamma_{xy}(t, f; Q_+, Q_-)$ of different wavelet-estimators x and input signal (PPG, ECG, aECG14), with the reference heart rate signal y. The phase of the coherence is very close to zero, except for PPG estimations (because of a small delay) so that we consider the squared coherence only. We average it over all times and selected subjects in the *capnobase* dataset to obtain a coherence profile $\gamma^2_{xy}(f) = \langle|\gamma_{xy}(t, f; Q_+, Q_-)|^2\rangle_t$ for each cardiac frequency modulation signal. Subjects are selected so that the quality of the

reference estimation is sufficient (SQI $> \frac{1}{2}$) during the entire recording. The significance of the coherence is controlled against the expected spurious level and verified from a surrogate signal.

The result in Fig.IV.11 is the following: all estimators of the cardiac have a significant coherence with the reference heart rate at modulation frequencies lower than the cardiac frequency. The narrow-band estimator is the most coherent one for all input signal, and the most coherent among them is based on the aECG14. The instantaneous frequency $F_x(t)$ yields the worst estimation when applied to the ECG, because of its strong non-circularity, while it is satisfying for the other cardiac signals. For the wide-band estimator $F_x(t;Q)$, its squared coherence reach about 0.5 for both the ECG and PPG recordings, which is quite good considering it has a single control parameter. In the following, we refer to the wide-band estimator based on the ECG as the *generic* estimator of the heart rate modulations, in contrast to the *precise* estimator, narrow-band and based on aECG14.

Summary of section 1

The cardiac rhythm has been introduced, together with some of its numerous recording techniques and the interest in extracting its key observable: the heart rate. Its variability (HRV) provides an elementary insight into the sympathetic-vagal balance, implied in the regulation of physiological functions.
Conceived as the modulated instantaneous frequency of the fundamental rhythmic component, a generic approach of the heart rate is obtained by combining temporal and spectral perspectives in the wavelet transform of the cardiac recording. In spite of its various waveforms or pulse shapes depending on the measurement technique, the regularity of the cardiac rhythm allows to isolate its fundamental component with a precise time resolution and mostly within a wavelet bandwidth at a small quality factor $Q = 5$. A generic extraction strategy is formulated based on a family of time-frequency estimators, applied on PPG and ECG signals.
These estimators are either wide-band or narrow-band, and either linear or quadratic. Quadratic estimators require an additional parameter but prove more robust, in relation to their statistical interpretation. The narrow-band estimator is especially precise, but the wide-band version only requires a single parameter. The latter is sufficient to extract modulations of the cardiac frequency, even though it is biased and contains fast oscillations for non-circular waveforms. The ECG is particularly non-circular, so that most of its intensity settles into impulses at higher frequencies. The precision of the rate estimation is increased when taking into account its pulsatile nature, by defining a new aECG14 signal derived from the wavelet transform at 14 Hz.
These estimations are compared with a reference heart rate, estimated from R-R intervals in the ECG. The precision of the frequency modulations is evaluated in terms of coherence at each frequency, retained for its insensitivity to various effects, such as the presence of a bias, a small delay between ECG and PPG (the pulse transit time), or fast oscillations at the cardiac (fundamental) frequency. Their coherence with the reference is significant, and increases with the circularity of the input cardiac signal and with the specificity of the band selection.

2 Cardio-respiratory interaction as a coherence

The heart and breathing activity are finely intertwined rhythms, forming the cardio-respiratory system that balances ventilation and perfusion to ensure suitable gaseous transport between the metabolism and the environment. We apply here the wavelet framework to the observation of its rich living dynamics, from the comparison of a non-invasive respiratory recording, the

airflow signal, to the heart rate signal estimated in the previous section.

We begin with the phenomenon of respiratory sinus arrhythmia (RSA), which actually is a very rhythmic modulation of the heart rate at the breathing frequency, as can be observed in Fig.IV.1 (A, D). Applying the coherence analysis, the RSA coupling can be characterized in time at the breathing frequency in terms of fluctuating phase and strength of coherence between the heart rate and the airflow recording. This provides a direct insight into the dynamics of the parasympathetic activity, that conveys this interaction involved in the regulation of ventilation. We also show that a significant but also intermittent coherence in the cardiac and respiratory variability exists at other scales, in particular at slow and less precise frequencies. The sympathetic system is an actor of this low frequency activity, that reflects the regulation of perfusion [Moser, Fruhwirth, and Kenner 2008; Moser, Frühwirth, et al. 2006], and we suggest it can be approached via bursts of slow coherence.

Then, we illustrate how the choice of the pair of cardio-respiratory signals, either direct recording or extracted modulation signals, affects the observed cardio-respiratory coherence and its interpretation. In particular, the use of breathing modulation rate signals, describing its instantaneous frequency and its amplitude modulation, are shown to provide a complementary perspective on the cardio-respiratory interaction. While they provide a less direct perspective on the fast RSA component, they allow a more reliable interpretation of the coherence and phase shift between slow cardio-respiratory modulations.

In short, the coherence analysis provides precise characteristics of the cardo-respiratory inter-regulation. Limited to finding linear coevolution in standardized physiological fluctuations, important information is left out: the actual values of these regulated biophysical quantities. One of these physiologically relevant quantity is the ratio between the cardiac and the respiratory frequencies, also called pulse-respiratory quotient (PRQ) [Scholkmann and Wolf 2019]. Sometimes, cardio-respiratory rhythms synchronize: the heart and breathing frequencies are commensurable [Schäfer et al. 1998]. To determine this frequency ratio at each time, we construct a ratio distribution from power densities of the cardiac and respiratory recordings, advantageously concentrated from frequency reassignment. We illustrate 3 situations, of well-defined ratio without or with synchronization, or with unstable ratio related to sleep apnea.

2.1 Respiratory recording

The breathing activity can be recorded by various means, from the most accurate but also invasive such as spirometry (air volume and flow) and capnography (CO_2 partial pressure), to the less invasive measurements methods such as nasal-oral airflow sensors (thermistor, thermocouple or pressure transducer), thoracic and abdominal effort belt (inductance plethysmography or piezoelectric) or chest motion tracking (accelerometer, video). Respiratory signals can also be derived from pulse oximetry (PPG). Less invasive methods are largely preferred for long recordings because of their simplicity. Several techniques are usually combined in a polysomnography, each with their specific artefacts: the resulting respiratory signals are only semi-quantitative [Farre 2004].

In this section, we focus on the airflow signals, measured with a thermocouple in the *shhs2* database and with a thermistor for subject 3 from another database, see descriptions in Appendix 1. The sum of the thoracic and abdominal effort signals from *shhs2* will also be used (as a volume signal from inductance plethysmography) to bring additional confirmation. We have selected the airflow signal because we have observed fewer artefacts than the effort respiratory signals in *shhs2*, occasionally occurring due to sleep position changes, affecting the amplitude of the oscillations. In the following, the sum of the effort signals will actually prove to be more accurate in tracking very slow variations of the lung's volume.

Exposed to temperature change of the nasal-oral airflow during inspiration (room temperature) and expiration (37 °C), the electric property of these sensors (tension and resistance respectively) varies. A correct calibration is assumed so that the signal's unit is in litre per second, and we invert the sign of the signal so that a positive value of the breathing oscillation corresponds to an inspiratory flow, and a negative one to an expiratory flow. Its resolution is about $10^{-3}\text{l}\cdot\text{s}^{-1}$ and sampled at 250 Hz in *slpdb*, $8\cdot10^{-3}\text{l}\cdot\text{s}^{-1}$ in *shhs2*. However, these sensors are known to provide an inaccurate measurement of airflow [Farré et al. 1998]: the waveform is smoothed and delayed (of up to one second) and the amplitude variation is non-linear, but these techniques are sufficient to capture the rate and modulations of breathing. In *shhs2*, an instrumental high pass filter at 0.05 Hz is mentioned, and the sampling frequency is 8 or 10 Hz.

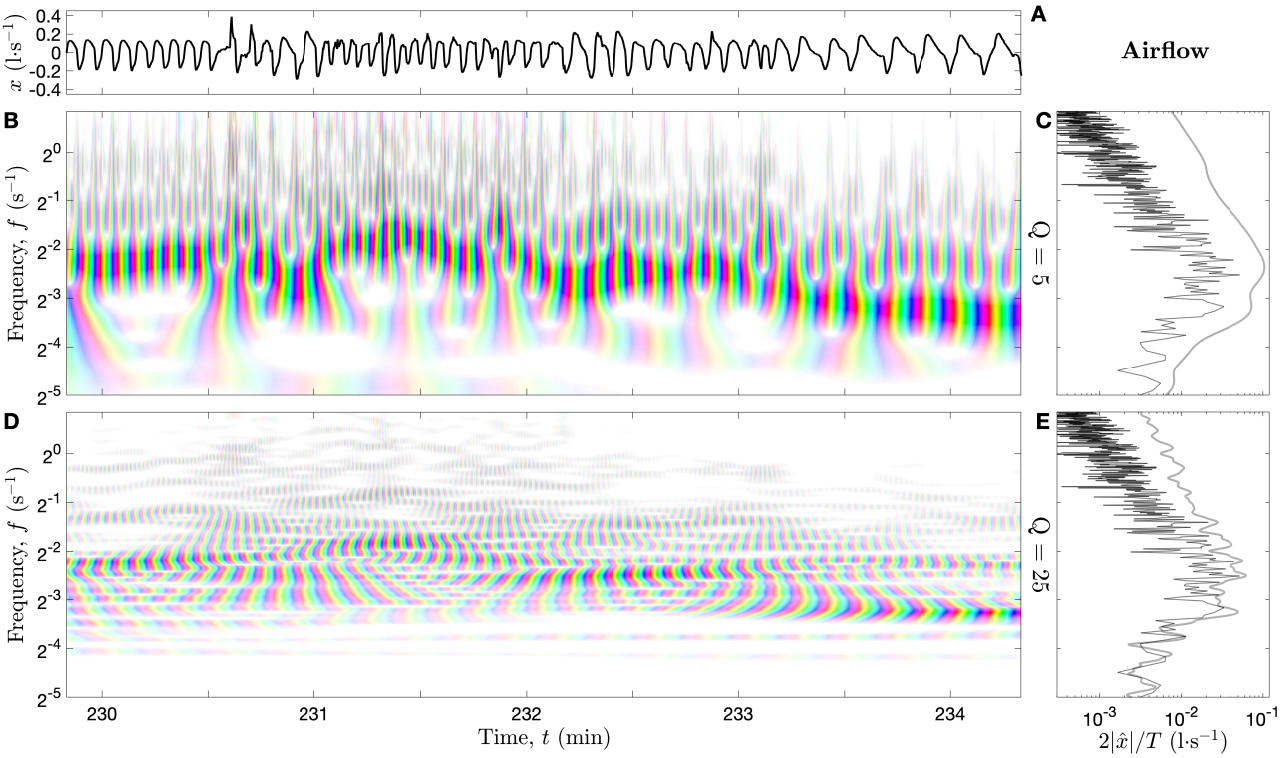


Figure IV.12: *Respiratory recording of subject 3. (A) airflow signal (in* $\text{l}\cdot\text{s}^{-1}$*) on a selected interval of high variability of the breathing rhythm. (B, D) Time-frequency representations, computed with the quality factor* $Q = 5$ *in (B), and* $Q = 25$ *in (D). Colours code for both the angle (hue) and the modulus (saturation) of the wavelet transform (see Fig.II.1 of chapter II). (C, E) Amplitude spectra obtained either from the Fourier transform (thin black line) or from the wavelet transform Eq.*(IV.3) *(thick grey line).*

An airflow signal is illustrated in Fig.IV.12 (A) for a selected time interval, that starts with a few steady breathing oscillations, followed by strongly modulated ones. The complete representation (amplitude and phase) of its wavelet transform is given in panels (B) and (D) for $Q = 5$ and $Q = 25$ respectively. Although they both contain the same information, the fast modulations of the breathing rate are much more easily recognized at the smaller quality factor $Q = 5$ (B), for which the fundamental breathing component is clearly present.

The effect of the quality factor in the frequency domain is summarized in the amplitude spectra $2\langle X(t,f;Q)|^2\rangle_t^{\frac{1}{2}}$ (root-mean-squared amplitude) obtained from the wavelet transform, thick grey line in Fig.IV.12 (C, E). Although more resolved at $Q = 25$ (E), the amplitude of the spectra is reduced compared to the one observed on the signal (A). This is due to breathing rate variations that are both wider than the wavelet bandwidth and shorter than its duration, causing a destructive interference between the signal and the wavelet. The amplitude is thus

more accurate at $Q = 5$ than at $Q = 25$. The wavelet amplitude spectrum approaches the Fourier estimation $2|\hat{x}(f)|/T$ (thin black lines, T duration of the selection) when the quality factor increases, see Eq.(II.8); an accurate amplitude is only expected for a pure wave.

2.2 Respiratory sinus arrhythmia as a coherence between heart rate and airflow

During inspiration, the intrathoracic pressure rises and affects the cardiac and circulatory activity. In a reflex compensatory mechanism enabled by baroreceptors, the parasympathetic activity is inhibited, causing an increase of the heart rate [Mohrman and Heller 2018]. The expiration has the converse effect: the heart rate decreases, saving its energy when the oxygen intake is low. This cardio-respiratory interaction is commonly observed as a fast variability of the heart rate at the breathing frequency, called respiratory sinus arrhythmia (RSA), see for instance panel (A, D) of Fig.IV.1.

The RSA is often characterized from the amplitude of the heart rate modulation, amplitude that tends to decreases with the breathing rate and increases with the tidal volume [Hirsch and Bishop 1981]. It is also strong for young or athletic subjects and decreases with aging, depend on the position of the body, etc. Even in the rest state, the occurrence of RSA is subject to an important temporal variability, sign of the complexity of the regulation mechanisms in the cardio-respiratory interaction.

This cardio-respiratory interaction is measured as a coherence between the heart rate and the breathing oscillation, in time and/or in frequency. Following the development of section 4 in chapter II, the significance of the RSA coherence [Faes, Pinna, et al. 2004] is also evaluated individually.

2.2.1 Temporal and spectral perspectives on the interaction

The most simple way to characterize the strength of the interaction between the heart rate $x(t) = f_c(t)$ and a respiratory signal $y(t)$ is certainly to compute their correlation coefficient ρ_{xy}. This can be done in a time-dependent manner thanks to its wavelet estimator $\Gamma_{xy}(t,t;Q) = \Re\{\Gamma_{x^+y^+}(t,t;Q)\}$ (without delay), circumventing the stationarity hypothesis controlled by the quality factor Q (see section 4.1 in chapter II). The complex version $\Gamma_{x^+y^+}(t,t;Q)$ is further able to enrich its interpretation with a continuous phase difference (the angular argument) instead of a discrete sign.

This is presented in Fig.IV.13 (A, B, C): the heart rate (A, red line, best wavelet estimator) is clearly oscillating with the airflow recording (B, blue line) during most of the time selection, causing a quite high correlation modulus, with close to no phase difference (colour-coded in green). This obvious manifestation of RSA is cluttered by sudden events (7 of them are visible), where the heart rate transiently rises above a base value, but also the breathing rate seen as faster oscillations in panel (B). Looking more closely to the airflow recording, these events are associated to more frequent and ample inspirations, called hyperpnea, directly followed by a reduced breathing activity, called hypopnea. These cardio-respiratory events affect the correlation, with a phase difference that is dragged towards negative values (lag of the heart rate) and a modulus either reduced or enhanced. Note that at phase quadrature, the complex correlation coefficient is essentially imaginary so that its real part (the usual correlation coefficient) vanishes.

This situation is clarified by examining the signals and their correlation in the frequency domain, as shown in Fig.IV.13 (D, E). Amplitude spectra for the full 8 hour-signals are shown in panel (D). As expected, they both exhibits a peak at the breathing frequency (about 0.22 Hz), sign of the RSA phenomenon in the variability of the heart rate, also composed of a slower

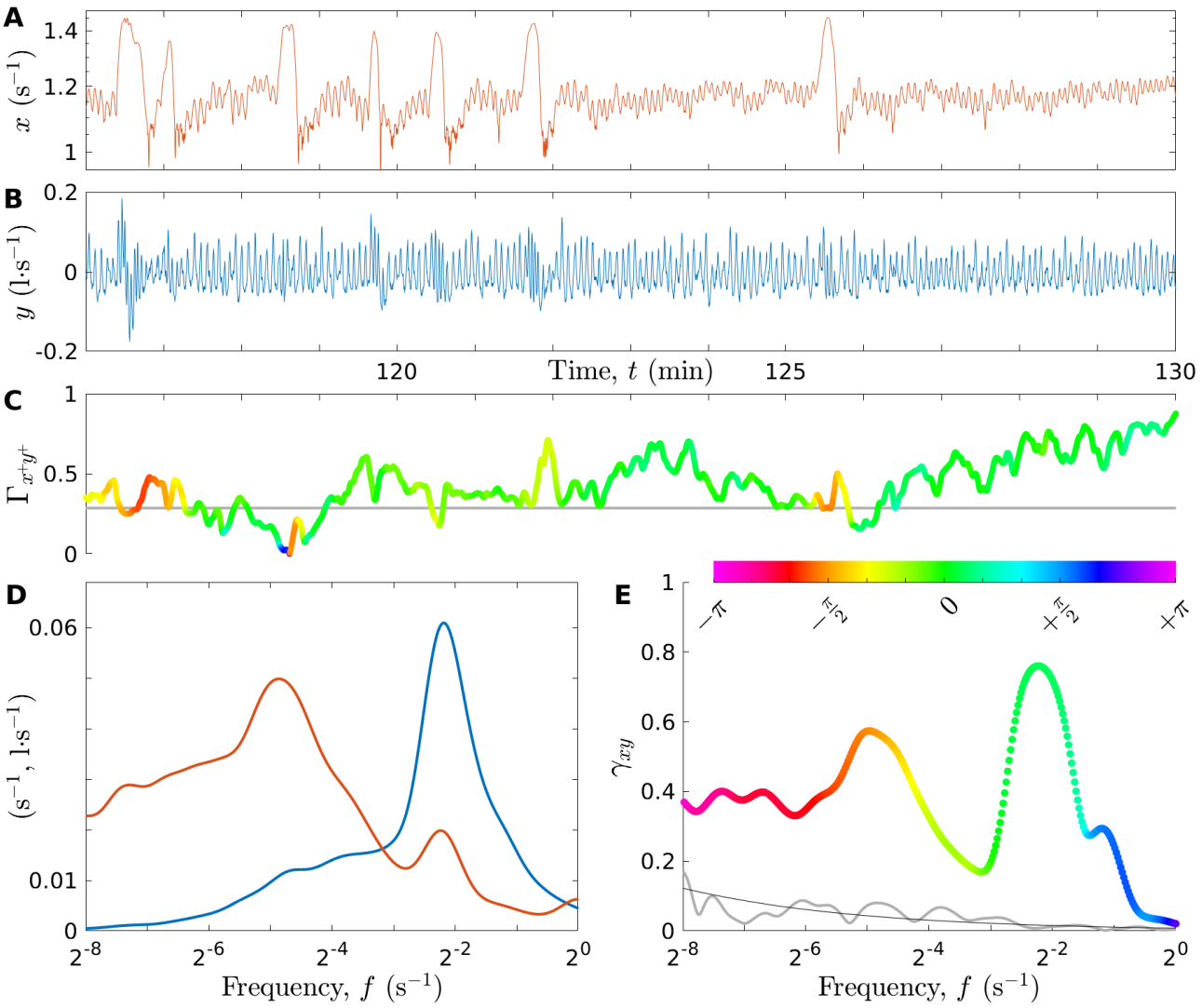


Figure IV.13: *Correlation between heart rate (red lines) and airflow (blue lines), either in time or in frequency, from the polysomnography of subject 4. (A) Heart rate signal $x(t) = f_c(t)$ (in s^{-1}), narrow-band wavelet estimation from the aECG14 signal. (B) airflow signal $y(t)$ (in $l{\cdot}s^{-1}$, positive for inspiratory flow). (C) Time-dependent complex correlation coefficient (temporal coherence) $\Gamma_{x^+y^+}(t;Q)$, and estimated spurious level. Only a selection time interval is shown. (D) Amplitude spectra Eq.*(IV.3) *for the entire signals (8 h long) with the same unit. (E) Spectral coherence $\gamma_{xy}(f;Q)$, with expected spurious level (thin black line) and simulated one (thick grey line). Every wavelet estimator is obtained for $Q = 5$. In both the temporal (C) and the spectral (E) cases, the coherence modulus is plotted, and its phase is colour coded with the chromatic circle, see the colour wheel (bar).*

wide-band component (peaked at about 0.033 Hz), precisely reflecting the strong variations of the heart rate. The RSA rate amplitude is as low as ± 0.02 s^{-1}, i.e. ± 1.2 beats per minute here. Non-harmonic waves give some amplitude at frequency of higher harmonic orders, explaining the right-skewed asymmetry of the peaks. Also notice the damping of the airflow amplitude below $2^{-4} \approx 0.06$ Hz, caused by its instrumental high-pass filtering.

The panel (E) is the spectral coherence $\gamma_{xy}(f;Q)$ (its modulus with a colour-coded phase), normalized version of the cross-spectrum $S_{xy}(f;Q)$ able to distinguishes correlations at different scales. The two cardio-respiratory modulations are thus well separated, and their associated phase is clarified: slightly above 0 for the fast RSA component, and about $-\frac{\pi}{2}$ for the slow one. We can hypothesize that the intermediary phase values observed in time is a superposition

of both types of coherent modulations. As measured by the squared modulus of coherence $|\gamma_{xy}|^2$, the proportion of one spectral density linearly explained by the other one is at most $0.77^2 = 0.6$ for the respiratory component. The remaining 0.4 proportion is a sign of the temporal variability of the fast RSA modulations, even higher for the slower modulations. A very weak amplitude can be sufficient to yield a significant coherence, as is the case for the second order harmonics of breathing, clearly appearing with a phase shift $+\frac{\pi}{2}$, as well as for the slow (even very slow) airflow modulations that has been partially filtered. This can be verified by comparing the modulus to the (much lower) level of spurious coherence, expected from theory (thin black line, obtained from table II.1) or simulated (thick grey line, from a white Gaussian noise surrogate).

2.2.2 Time-frequency view on the cardio-respiratory interaction

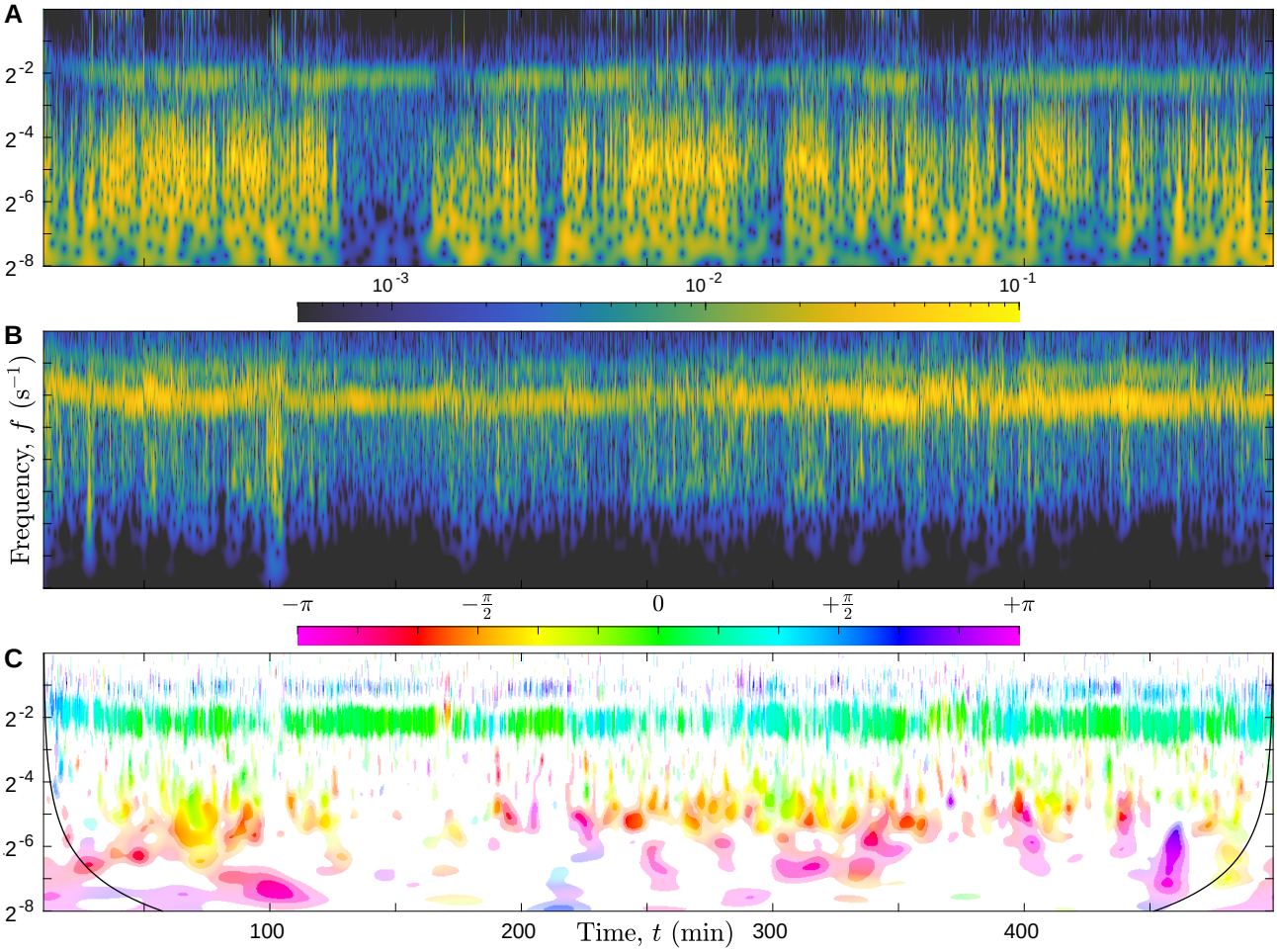


Figure IV.14: *Correlation between the heart rate and airflow of subject 4 in the time-frequency plane. (A) Amplitude of the heart rate modulations, $2|X(t,f;Q)|$ (in s^{-1}), and (B) amplitude of the airflow modulations, $2|Y(t,f;Q)|$ (in $l{\cdot}s^{-1}$), with shared logarithmic colour-scale. (C) Time-frequency coherence $\gamma_{xy}(t,f;nQ,Q)$ estimated for $Q=5$ and $n=5$, and cone of influence (black line) delimiting regions with possible border effects. The hue-saturation colour coding for phase and modulus is here discretized into 3 ranges of significance, with p-value above 10^{-3}, 10^{-2} and 10^{-1} with decreasing colour saturation (white below), that corresponds to threshold moduli $|\gamma_{xy}| = 0.87, 0.78, 0.61$ respectively. The time and frequency axis are identical for all images.*

These time and frequency perspectives are combined in Fig.IV.14: that represents the

wavelet transform (the amplitude) of the heart rate (A), the airflow (B), and their time-frequency coherence $\gamma_{xy}(t, f; nQ, Q)$ (C). These images disclose the extent of the temporal variability and the spectral separation for these modulations during sleep, in terms of amplitude, phase difference and coherence modulus. This level of details in the dynamics of the cardio-respiratory interaction is obtained with $Q = 5$ and from a low number of statistical degrees of freedom in the coherence analysis, close to $n = 5$ here: it corresponds to resolutions of $f\delta t \approx 10$ oscillations and relative bandwidth $\delta \log f \approx 0.5$ (n is their product). Computed from a time-smoothing over the 10 oscillations, the ensemble average estimator favours the representation of the quite harmonic (horizontal) coherent component at the breathing frequency, to the detriment of a precise time localization of the very pulsatile (vertical) slow components. That few degrees of freedom come at another cost: it constrains the analysis of strong coherence only, the spurious coherence level being as high as $\gamma_{\text{sp}} \approx (1+n)^{\frac{1}{2}} = 0.4$ so that a true but weak coherence cannot be distinguished from spurious coherence. To improve the readability of the coherence image, its significance is directly encoded into the saturation of the colour, increasing zero (white) to one (colourful) when the p-value gets lower than 10^{-1}, 10^{-2} and 10^{-3}. These different coherence levels are reproduced in Fig.IV.15 (B).

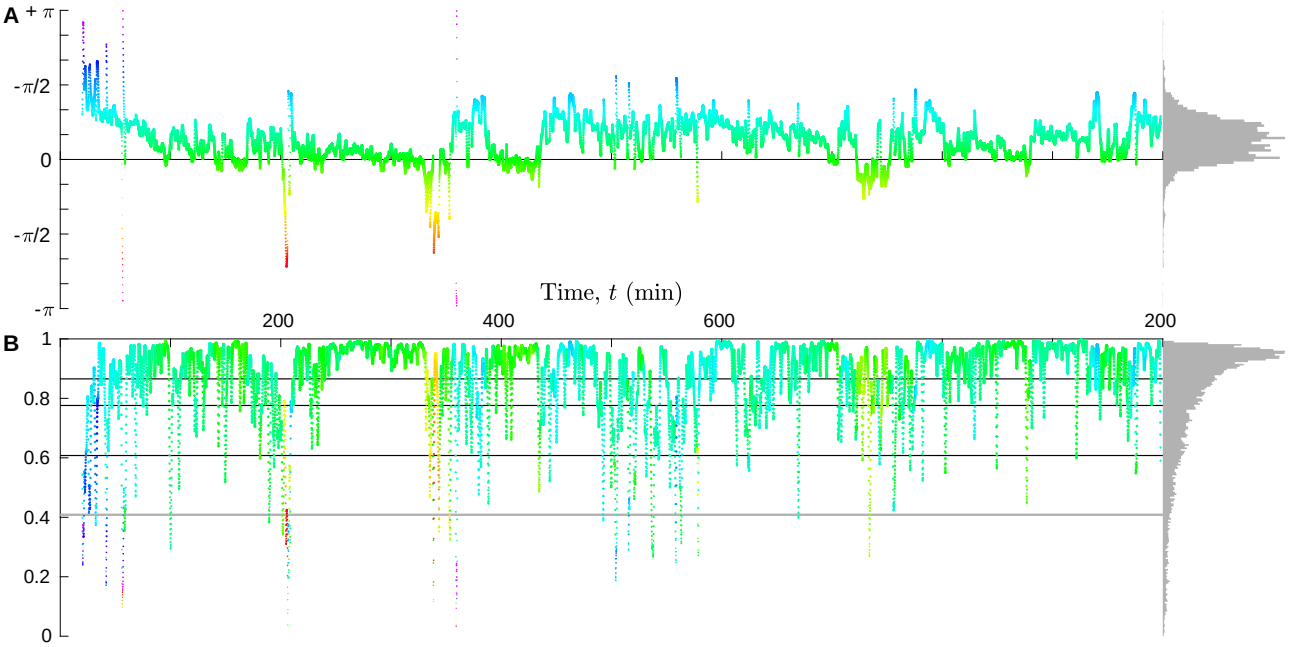


Figure IV.15: *Temporal variability of the respiratory sinus arrhythmia of subject 4, measured as $\gamma_{RSA}(t) = \gamma_{xy}(t, f; nQ, Q)$ from Fig.IV.14 for a fixed respiratory frequency $f = 0.22$ Hz (maximizing $\gamma_{xy}(f; Q)$). The phase and modulus of this complex signal are plotted in (A) and (B) respectively, and repeated in the colour and size of the markers. Different significance levels are represented in (B): the expected spurious coherence (thick grey line) and the increasing significance thresholds (thin black lines) with p-values 10^{-1}, 10^{-2} and 10^{-3}.*

To focus on the temporal variability of the fast RSA component, free from the influence of the slower ones, we can extract the dynamics of the amplitude and coherence into new signals. For simplicity, we exploit the fact that the breathing frequency here fluctuates within one wavelet bandwidth (that should be increased if necessary), around the frequency $f = 0.22$ Hz. The phase and modulus of a coherence signal $\gamma_{\text{RSA}}(t) = \gamma_{xy}(t, f; nQ, Q)$ is presented in Fig.IV.15, thus constructed by slicing the coherence image at this fixed frequency. The resulting signal is similar to the complex correlation coefficient, except that it is narrow-band, hence free from the influence of the strong but slower modulations.

It is clear from panel (A) of Fig.IV.15 that the phase of the respiratory modulation in

the heart rate is most of the time early compared to the airflow phase, with a fluctuating difference ranging from almost $+\frac{\pi}{2}$ to slightly below 0. Important deviations from these phase shift are associated in panel (B) to a drop in coherence modulus (hence in its significance), that otherwise frequently exceeds $\gamma_{xy} = 0.95$.

2.3 Statistics over *shhs2* of the phase between heart rate and breathing recordings

To know whether these observations are specific to the selected subject or a widespread characteristic of RSA, we compute the time-frequency coherence between the heart rate signal and respiratory recordings for all 2650 subjects in the *shhs2* database, as in Fig.IV.14 (C). We are interested in the phase ϕ_{xy} of the coherence $\gamma_{xy} = |\gamma_{xy}|e^{i\phi_{xy}}$, more specifically in its typical values and temporal variability during sleep, at different scales.

We start with the airflow recording, and we reproduce the computation for a lung volume signal. This last signal was use in [Saul, Berger, Albrecht, et al. 1991] together with the heart rate signal to obtain a phase response relation from 7 subjects and 13 min recordings. We expect to extend their measure towards lower frequencies thanks to the long duration of the polysomnography and the numerous subjects.

2.3.1 Phase distribution as an angular density of coherence

We propose to construct an angular density of coherence (at each frequency) as follows:

$$G_{xy}(\phi\,;f) = \langle \gamma_{xy}\delta(\phi - \phi_{xy})\rangle_t \quad , \tag{IV.5}$$

where the coherence in the estimator $\gamma_{xy}(t, f; nQ, Q)$ and the time averages run over the 2.3 years of cumulated sleep time in the database. This resemble the normalized histogram (estimated probability density function) of ϕ_{xy}, where $\delta(\phi - \phi_{xy})$ is in practice a bin of finite width around the phase value ϕ. The only difference with a histogram is the presence of a weight, the coherence γ_{xy}, that favours significant phase values. This definition of the angular density of coherence is related to the time-averaged coherence by an integral:

$$\langle \gamma_{xy}(t, f)\rangle_t = \int_0^{\Upsilon} G_{xy}(\phi\,;f)\mathrm{d}\phi \quad , \tag{IV.6}$$

whose phase is the mean value of ϕ_{xy} in the sense of circular (or directional) statistics [Mardia and Jupp 2000].

The advantage of using such time average is two-fold. First, different subjects contribute to the statistics proportionally to their sleep duration. Then, it will allow in the next chapter to extract non-stationary characteristics by conditioning the time average to the observation of a certain event, such as sleep stages or apnea occurrences annotated in the polysomnography. It has the following limitations: it does not distinguish the intra-individual and inter-individual variabilities, and it does not keep track of the distribution of the coherence modulus.

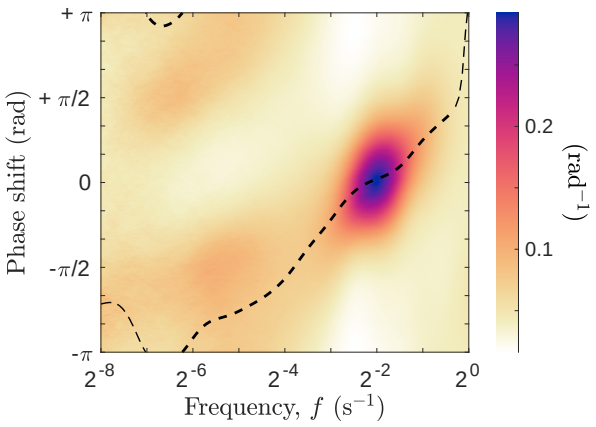
Note that the mean coherence uses two averages, one is local in time before normalization, and the other after normalization. The interest of keeping the local time-averaging is not clear. Without any time averaging before normalization, we obtain the so-called *phase synchrony* measure [Aviyente, Bernat, et al. 2011], whose meaning is similar to the coherence. Conversely, an infinite time-averaging before normalization yields the spectral coherence. On one hand, the spectral coherence weights the phase proportionally to the power density, which vanishes on disturbances causes by phase vortices. On the other hand the spectral coherence is much more sensitive to intense artefacts that phase synchrony. The extent of the time-averaging n limits the proportional averaging in time and interpolates between these two behaviours. It is

observed to slightly affects the value of the angular density of coherence, without changing it qualitatively.

The angular density $|G_{xy}(\phi\,;f)|$ of this cardio-respiratory coherence is shown has a heat map in Fig.IV.16. It is most dense at the breathing frequency at phase 0 with a typical spread of $\pm\frac{\pi}{3}$. This constitutes a measure of the phase variability of RSA across time and subjects.

Although the coherence is weaker at lower frequency, its angular density is more spread (sign of an important variability) but still polarized around a negative mean value. This mean phase is materialized by a thick dashed line, made thinner when less defined, at scales where the angular density is close to uniform. This is the case at very low and very high frequencies, the latter being explained by the loss of significance of the modulations in the rate signal $f_c(t)$ close to 1 Hz (see Fig.IV.11).

Figure IV.16: *Angular density $|G_{xy}(\phi\,;f)|$ of coherence γ_{xy} between heart rate (x) and airflow (y) over the 2.3 years of cumulated sleep in the shhs2 database. The intense coherent mode corresponds to the distribution of the phase between RSA modulations and airflow oscillations, see Fig.IV.15. This un-normalized density integrates to the time-averaged coherence $|\langle\gamma_{xy}\rangle_t|$ of Fig.IV.17, whose phase (the circular mean) is plotted against frequency (dashed line).*



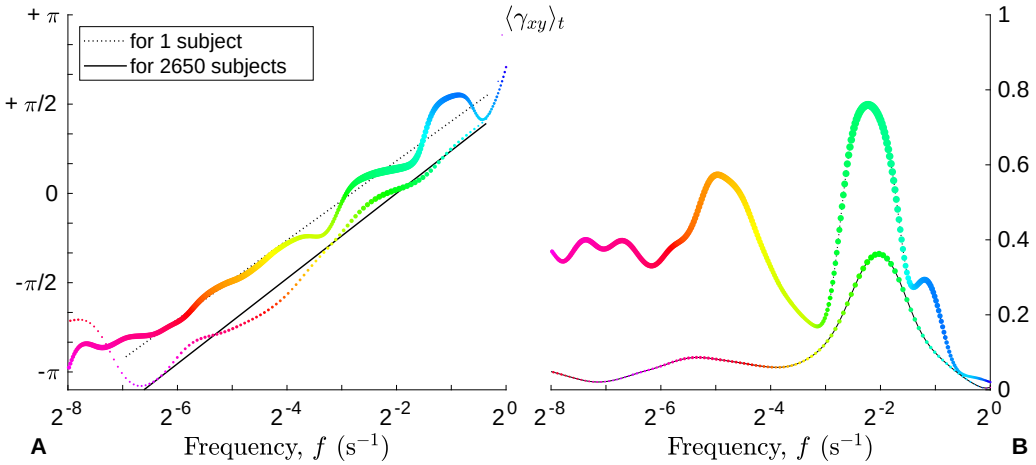


Figure IV.17: *Phase (A) and modulus (B) of the mean coherence $\langle\gamma_{xy}\rangle_t$ between the modulations of the heart rate and the airflow. The phase and modulus are repeated in the colour and size of the markers. Two coherence profiles, obtained from a time average $\langle\gamma_{xy}\rangle_t$, are compared: one over the 2.3 years of cumulated sleep recording for all 2650 subjects in the shhs2 database (lower curves), the other over the 8 h sleep of the subject 4 (upper curves). A linear relation is observed between coherence phase and log-frequency, and adjusted for the individual profile (black dotted line) and the collective profile (black plain line) in (A), see text.*

2.3.2 Trend of the mean phase across frequency

The phase of this collective mean coherence is reported in Fig.IV.17 (A) together with its modulus (B), and compared to the individual mean coherence computed for the single recording

of Fig.IV.14. From the panel (B), we observe that the collective cardio-respiratory coherence profile (lower curves) has a smaller modulus than the individual one. In agreement with the angular distribution in Fig.IV.16, this is either the sign of the important inter-individual phase variability for the slow component (0.015 to 0.06 Hz), that produces a destructive interference when averaged, or the sign of a weaker expected coherence for the fast component (0.25 Hz), associated with a lower phase variability.

An intriguing observation is the sensibly linear relation between the mean coherence phase and the log-frequency:

$$\phi_{xy}(f)/\Upsilon = \alpha \log \frac{f}{f_0} \quad , \tag{IV.7}$$

on both the collective and individual profiles up to a small phase shift. The slope α and the reference frequency f_0 were adjusted by linear regression over the 2 decades (0.009 to 0.9 Hz) where the phase distribution is clearly non-uniform. For a greater precision, we weight the data by the inverse of the circular variance, estimated as $1 - |\langle \gamma_{xy} \rangle_t|$. The numerous data points are very correlated, but we expect them to cover $\frac{\log 100}{\delta \log f} \approx 9$ distinct wavelet bandwidths. The result is as follows for the collective (respectively the individual) cardio-respiratory coherence profile: the goodness of the fit, expressed as a coefficient of determination, is $R^2 = 0.990$ (0.980), and the model parameters are $\alpha = 0.166 \pm 0.004$ (0.16 ± 0.03) and $f_0 = 0.24 \pm 0.01$ Hz (0.142 ± 0.003 Hz). Notice how the phase shift between the individual and collective profiles gets absorbed into their different value of f_0.

What is the physical interpretation of this model? Considering a signal $x(t) = y(t - \tau)$ lagged of a time τ compared to another signal $y(t)$, their coherence is $\gamma_{xy}(f) = e^{-i\Upsilon f \tau}$. Therefore, we expect a phase proportional to frequency (not log-frequency) in case of a pure delay. However, we may have two distinct delays associated to each coherent components, which is difficult to distinguish with only 9 effective data points. We can reconcile this interpretation with Eq.(IV.7) considering that it interpolates different delays at different scales. This idea can be expressed as a group delay (of x compared to y):

$$\tau_{xy}(f) = -\frac{\mathrm{d}\phi_{xy}(f)}{\Upsilon \, \mathrm{d}f} \tag{IV.8}$$

$$= -\frac{\alpha}{f} \quad , \tag{IV.9}$$

that is inversely proportional to frequency. This has the following scale-invariant interpretation: the group delay $\tau_{xy}(f)$ is a constant portion $\alpha = 0.17$ of an oscillatory cycle of frequency f.

2.3.3 Verification from thoracic and abdominal respiratory effort signals

We now try to reproduce this observation when replacing the airflow recording with other available respiratory signals, in particular the lung volume for comparison with the phase responses in [Saul, Berger, Albrecht, et al. 1991].

The variations of the lung volume are estimated by summing the thoracic and abdominal effort belt recordings, measured by inductance plethysmography. We use the convention of a positive oscillation when lungs are filled. The airflow being the time-derivative of the volume, we expect the slow volume oscillations to be enhanced and delayed by a constant phase shift $-\frac{\pi}{2}$ (that does not affect the group delay): $x(t) = \dot{V}(t) \quad \Rightarrow \quad \hat{V}(f) = \hat{x}(f)/(i\Upsilon f)$. Surprisingly, we did not observe such phase difference between the "airflow" and the "volume" signals (that is close to zero). We hypothesize that the recording from the thermocouple is associated with a delay of the order of 1 s compared to the actual airflow [Farré et al. 1998], that would explain this observation. This imposes to be careful on the interpretation of these respiratory

signals, which are indirect and semi-quantitative measurements of the corresponding physical quantities. The associated imprecisions and uncertainties are is discussed in the next section.

We reproduce the computation over the complete *shhs2* database of the coherence between cardiac frequency modulations and respiratory oscillations. As in Fig.IV.16, their phase shift is represented as a coherence-based angular distribution at each frequency in Fig.IV.18.

The previous result from the airflow signal is recalled in Fig.IV.18 (A), whereas the result from the lung volume signal is shown is panel (B). First, the fast RSA component is slightly more coherent with a similar phase variability but a slightly lower mean phase due to the change of respiratory signal.

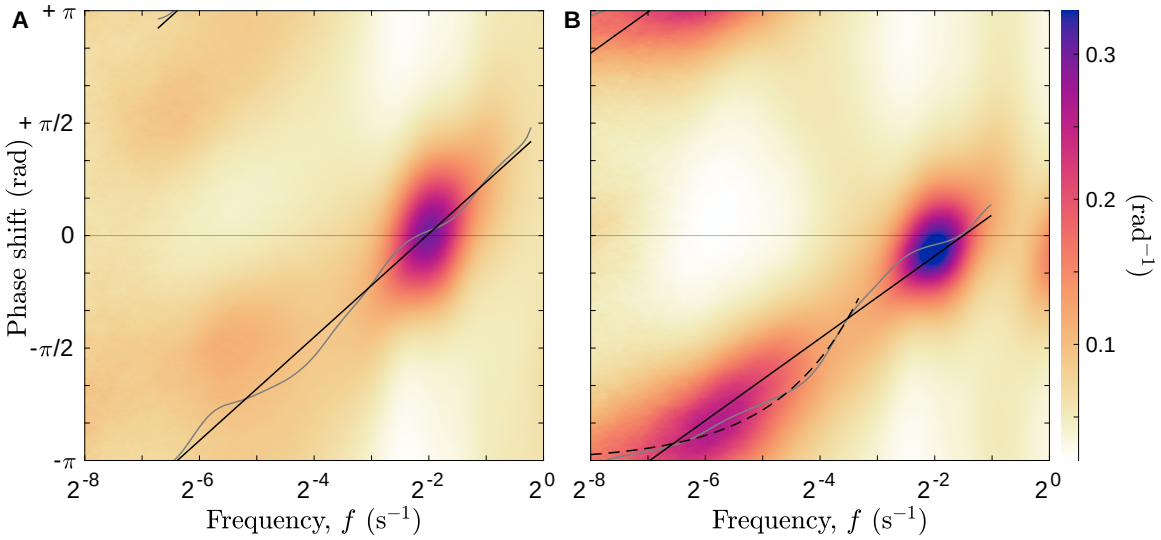


Figure IV.18: *Comparison of the phase distributions $|G_{xy}(\phi\,;f)|$ in shhs2 between the heart rate (x) and two different respiratory signals (y): either (A) the airflow measured by the thermocouple (same as Fig.IV.16) or (B) the lung volume measured by the effort belts. The mean phase (plain grey line) is plotted with its closest linear adjustment with log-frequency (plain black line) in (A, B). The linear adjustment with frequency (dashed black line) of the slow coherent component in plotted in (B).*

Then, we observe a similar trend of the phase across log-frequency is observed in both cases. This trend has a better angular concentration and a higher coherence at low frequency for the volume signal. This suggests a more precise measurement of breathing from the effort belts rather than from the thermistor, whose lower precision results in angular dispersion and incoherence. We fit the same model Eq.(IV.7) as previously (plain line) to the mean phase $\phi_{xy}(f)$ between cardiac frequency and breathing, from the volume signal. The slope is significantly smaller than before, $\alpha = 0.132 \pm 0.006$ ($f_0 = 0.35 \pm 0.04$ Hz), and the model appears less suitable ($R^2 = 0.955$) to explain the trend of the mean phase in spite of the increased coherence compared the panel (A). We conclude that the description of both the slow and fast components with the model Eq.(IV.7) may not be suitable.

Therefore, we take advantage of the lower angular dispersion in panel (B) of Fig.IV.18 to adjust the phase of the slow coherent mode only ($f < 0.05$ Hz). An affine relation to the frequency, i.e. the model:

$$\phi_{xy}(f)/\Upsilon = f\tau + \beta \quad , \tag{IV.10}$$

dashed line, better explains the trend ($R^2 = 0.983$): we obtain a phase opposition $\beta = -0.501 \pm 0.004$ and a delay $\tau = 3.6 \pm 0.1$ s (less than a breathing cycle) between slow cardiac frequency modulations and slow lung volume variations.

In [Saul, Berger, Albrecht, et al. 1991], the phase response of the cardiac frequency to the variations of the lung volume (with a forced random breathing rate to increase its bandwidth) were measured in controlled situations with either a dominant vagal or sympathetic regulation. Our result from polysomnography is consistent with the phase response in the vagal situation, expected during the sleeping state. They mainly observe the RSA fast component (with close to zero phase shift), but the slow component is harder to observe due to their limited data (7 recordings of 13 min): only the negative sign of its phase is visible on the measure but neglected in the modelling. We have precisely characterized it as a phase opposition and a delay: the slow modulations of the heart rate precede the slow variations of the lung volume by about $\tau = 3.6 \pm 0.1$ s while they evolve in opposite ways ($\beta\Upsilon = \pm\pi$). Even though it appears counterintuitive, this is not incompatible with the fact that the phase of the slow heart rate modulations is smaller (late) compared to the slow trend of the volume ($\phi_{xy}(f) \in [-\pi, 0]$ for $f < 0.05$ Hz). The existence of the delay, visible in the phase-frequency relation, suggests a causality directed from the heart rate to the slow trend of the lung volume.

2.3.4 Summary and instrumental limitations

From the individual example alone, we cannot doubt of the multi-scale and non-stationary nature of the cardio-respiratory interaction, depicted as a coherence between the heart rate and the airflow. We showed how to precisely locate the time and scale of very coherent oscillations in both signals. In particular, the respiratory sinus arrhyhtmia is described as a fast and coherent component at the breathing frequency, subject to some temporal fluctuations of its strength and phase difference. We also observed that a second coherent component occur significantly at the low frequencies supposed to be filtered out of the respiratory recordings. The modulus of the coherence in indeed independent on the actual amplitude of the modulations, which provides some robustness to instrumental non-linearity and filtering.

The proposed analysis of the coherence across the *shhs2* database further provides a statistically robust description of the coherence, in particular of its phase between the heart rate and airflow signals at different scales. This methodology provides the typical strength of the RSA coherence and the extent of its phase variability for different persons. The temporal variability of the slower coherent component appears especially important with the airflow signal. A part of this variability is of instrumental origin, explained by the progressive loss of the physiological signal into the noise below the high-pass frequency 0.05 Hz. The slow coherence is indeed stronger for the lung volume signal (hence even stronger for the true biophysical quantities), also subject to the same instrumental filtering, but with stronger slow oscillations due to its integral relation to the airflow.

A typical phase-frequency profile emerges, with a close to vanishing phase for the fast RSA component and a negative phase decreasing with the frequency, interpreted as a delay of the slow variations of the lung volume (i.e. the ones of its trend or running average) compared to the cardiac frequency slow modulations. Already visible with the airflow signal, the phase of coherence between these slow modulations has been characterized precisely as the result of a delay $\tau = 3.6 \pm 0.1$ s of the slow lung volume variations compared to the slow heart rate variations, evolving in phase opposition. These are physiological characteristics of the mainly vagal inter-regulation of the cardio-respiratory rhythms during sleep.

However, these observations can be subject to several instrumental biases, especially the value of the phase, which is very sensitive to the instrumental characteristics. The airflow recording with a thermoelectric sensor is indeed problematic since it is delayed of about one second compared to the actual airflow [Farré et al. 1998]. This can imply a bias of about $-\frac{\pi}{2}$ on the airflow phase at the breathing frequency, that makes the thermistor oscillation close to the lung volume oscillation, so that the interpretation of the RSA phase difference extracted

in Fig.IV.15 would flip from a phase-leading heart rate to a phase-leading airflow. The latter is consistent with the close to zero (slightly negative) RSA phase between heart rate and lung volume, assumed without a delay.

More concerning, the phase response of the instrumental high-pass filter at 0.05 Hz, mentioned for all respiratory signals, is not known. Apart from the damping of slow oscillations, responsible for a loss of coherence at very low frequency, the filter can also introduce a phase bias. The simplest such filter introduces a phase shift of the oscillations ramping from 0 at high frequencies to $\frac{\pi}{2}$ at the filtered low frequencies. The true phase difference between the coherent slow heart rate and respiratory modulations could then be smaller than measured, as well as their delay.

The extraction of the heart rate signal from a recording (here the ECG) makes it less bounded to the instrumental specificities of this recording. Although a delay can still exist between different cardiac recording methods (for instance between the ECG and the PPG), the resulting slow cardiac modulations are not subject to the instrumental high-pass filtering since they are carried by faster oscillations.

Without a proper calibration, the airflow measurement from a thermoelectric sensor is a semi-quantitative description of the actual airflow into the lungs. Heavily influenced by instrumental characteristics that are not provided, it appears as a poor choice to describe accurately the physiological regulation mechanisms from a phase-frequency relation. Its is nevertheless rich in physiological information, that can be better exploited through its extracted modulation rate signals.

2.4 Coherence: from recordings to rate signals

The previous discussion highlights the ambivalence of the coherence formalism: one hand, it provides a very detailed analysis of correlated oscillations, at different times and scales, between two correlated dynamics; on the other hand its precise interpretation (in particular the one of the phase) highly depends on the selected fluctuating quantities, that can inextricably mix experimental characteristics to physiologically relevant information.

We propose a comparison of the similarities and differences in the interpretation of the cardio-respiratory coherence when we vary the nature of the input signals. We distinguish 3 of them: the recorded oscillations, their frequency modulations (FM) and their amplitude modulations (AM).

Respiratory FM and AM can be estimated similarly as for the cardiac rhythm for the respiratory recording x, here the airflow (AF). Due to its conveniently low non-circularity, but its important variability, the wide-band estimator $F_x(t;Q)$ is especially suited for the breathing rhythm. Its real part, the mean instantaneous frequency, is superimposed to wavelet transform in Fig.IV.19 (B). As for the cardiac signals in Fig.IV.8, the slight non-circularity yields an overestimation of the breathing fundamental frequency when using a quadratic estimator. Nevertheless, the coherence analysis is insensitive to the actual value of the rate signal (from standardization), and the breathing rate modulations are effectively captured by this generic wide-band estimator. Its imaginary part, related to the amplitude in Eqs.(III.8, III.30), is not straightforward to interpret. Instead, we show the amplitude estimated by inverting their relation (exponential of the integral, with an integration constant set by hand) in thick red lines in Fig.IV.19 (A), so that it directly compare to the envelope of the airflow recording (thin black line).

In the following, we compare cardio-respiratory pairs of signals, among recordings and FM or AM signals, in a time-frequency analysis. Notice that the "cardiac AM" is not considered, because its does not contain cardiac information but respiratory one: it captures the periodic changes in position of the ECG electrodes on the chest relative to the heart and impedance

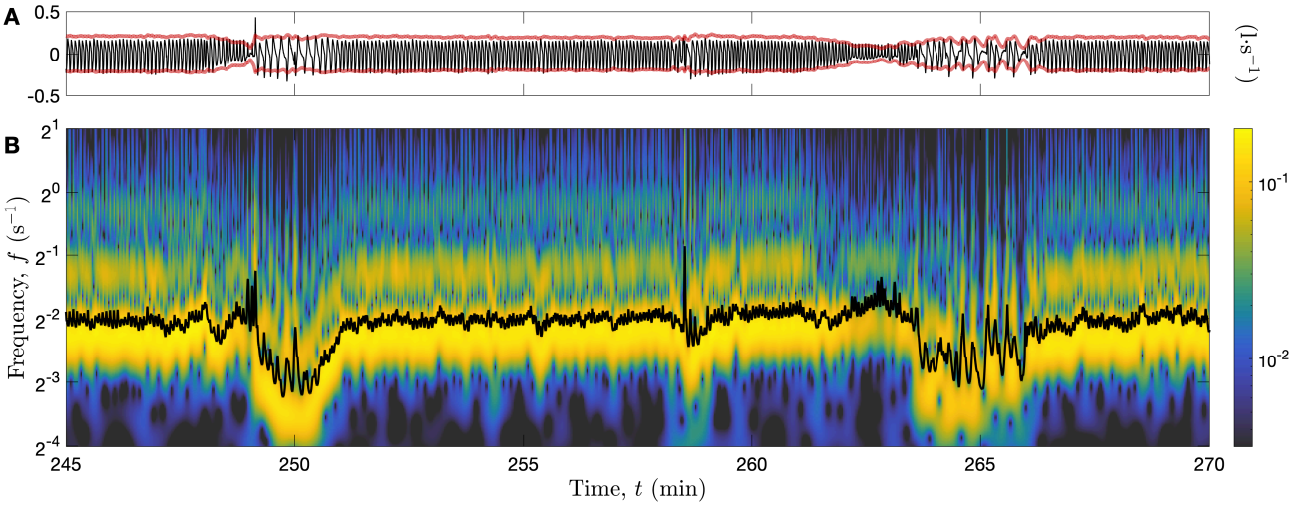


Figure IV.19: *Respiratory recording of subject 3 and estimation of the breathing rate modulations. (A) Airflow signal (in $l{\cdot}s^{-1}$, thin black line) and (B) amplitude (same unit) from the wavelet transform of quality factor $Q = 5$. (A) Breathing amplitude (thick red lines) and (B) mean instantaneous frequency (thick black line), derived from the wide-band generic rate estimator $F_x(t;Q)$.*

variations as the lungs fill and empty [Moody et al. 1986]. This ECG-derived respiration signal could have been an interesting candidate to replace the airflow recording, as it is shielded from its instrumental high-pass filter limitations.

2.4.1 Recording vs. recording

In Fig.IV.20, the first panel (A) is the direct comparison of the ECG and the airflow raw recordings. Their coherence bears a heavily instrumental information: the cross-talk between the sensors. A band of strongly significant coherence is found at high frequencies, in particular at the breathing frequency and at its second harmonic order (one octave higher). It means that a respiratory oscillation is present in the ECG, as a weak variation of the electric potential baseline. The coherence in this band is intermittent with a varying phase. Compared to the airflow, the reference respiratory signal, this ECG breathing oscillation is a mediocre representation of the respiratory oscillation, that can be considered an "echo" in the ECG electric baseline of the respiratory impedance modulation (clearer in the amplitude modulations of the ECG). The change of coherence phase is certainly due to positional changes during sleep. No significantly coherent event are observed at lower frequency, apart from isolated singular events associated to shared motion artefacts. Other examples of cross-talk between recording are shown in the Appendix 5.4, Fig.II.7.

2.4.2 Frequency modulation vs. recording

The next panels (B, C, D) of Fig.IV.20 involve at least one modulation rate signal. We showed in chapter III how the frequency and amplitude modulations could be both captured into a complex rate signal derived from the recording.

In the previous section, we used as a heart rate the sophisticated cardiac frequency narrow-band estimator $f_c(t) = \Re\{F_x(t, f; nQ, Q)\}$ where $x(t)$ is itself derived from the ECG (aECG14 signal, the amplitude at 14 Hz), that yields a very good estimate. We replace it by its most generic and wide-band alternative, the mean instantaneous frequency $\Re\{F_{ECG}(t;Q)\}$, compared once again to the airflow (AF) recording in panel (B). The previously described coherent and

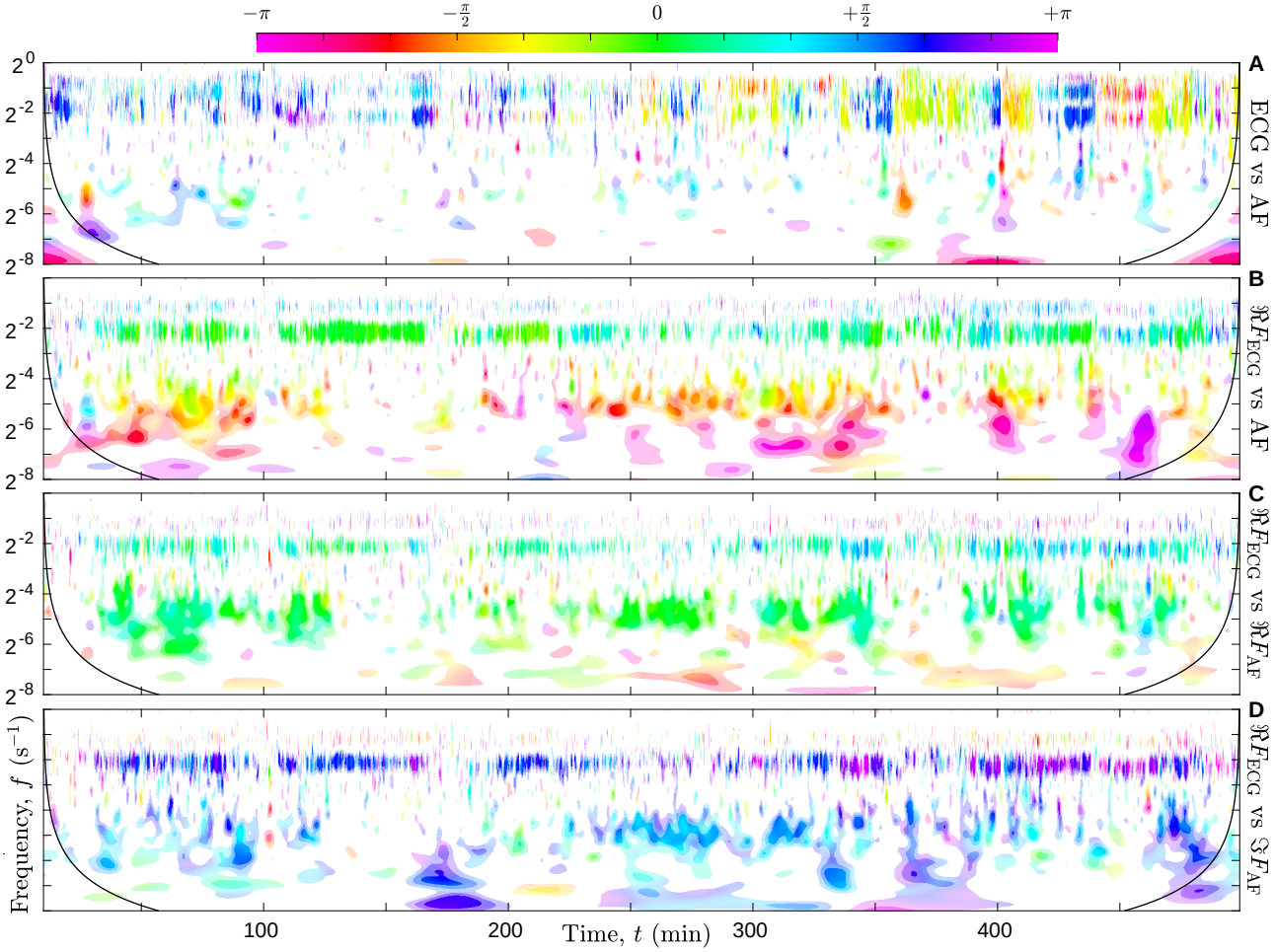


Figure IV.20: *Comparison of the time-frequency coherence $\gamma_{xy}(t, f; nQ, Q)$ associated to different pairs of cardio-respiratory signals (x vs. y) for subject 4 ($Q = 5$, $n = 5$). (A) ECG vs. airflow (AF) bare recordings. (B) Cardiac frequency vs. AF recording. (C) Cardiac frequency vs. respiratory frequency. (D) Cardiac frequency vs. respiratory magnitude rate. Mean instantaneous frequency and magnitude rate of a recording x (ECG or AF) refer here to the real and imaginary parts of the wide-band rate signal estimator $F_x(t; Q)$, for $Q = 5$.*

fast RSA component is well reproduced, as well as the slow one. In fact, the resulting coherence is hardly distinguishable from its counterpart of Fig.IV.14 (C) that uses the more precise heart rate. The only difference between these input rate signals is a proportionality constant (due to higher order harmonics), that does not affect the coherence analysis, and a repetition of the cardiac oscillation at about 1 Hz, limited to the upper limit of the investigated frequency range. See the different heart rate signals in Fig.IV.8 (D). This constitutes a strong methodological simplification: the generic and the sophisticated heart rate estimators contain the same cardiac FM information.

2.4.3 Frequency modulation vs. Frequency modulation

We carry on this approach for the breathing rhythm, by comparing in Fig.IV.20 (C) the heart rate to the breathing rate obtained with the same generic wide-band estimator $\Re\{F_{AF}(t; Q)\}$ (plotted in Fig.IV.19 (B)). The horizontal line of the RSA component is also present with a similar phase but weaker coherence than in panel (B). This is a beneficial effect of the repetition of carrier (breathing) wave in the rate estimator. The study of the RSA coherence through this indirect effect is certainly less reliable than using the airflow recording, but a useful echo. The main advantage of this cardio-respiratory coherence is the possibility to interpret easily the

phase between of the slow rates modulations. Free from the instrumental high-pass filtering, it unambiguously shows that there is close to no phase difference between a variation of the heart rate and a variation of the breathing rate. The last concern is the global delay of the airflow recording compared to the ground truth, and transmitted to the breathing rate signal. Of the order of 1 second, such delay has a limited effect on the phase of the slow breathing modulations since it represents no more than 1/16 of a cycle below $2^{-4} \approx 0.06$ Hz.

2.4.4 Frequency modulation vs. amplitude modulation

In the last panel (D) of Fig.IV.20 the heart rate is compared to $\Im\{F_{AF}(t;Q)\}$, the imaginary part of the generic wide-band estimator. Its relation to the signal's amplitude is not straightforward, so we recall it:

$$\Im\{\Upsilon F_y(t;Q)\} = -\partial_t \log \sqrt{P_y(t;Q)} \quad , \tag{IV.11}$$

where $\sqrt{P_y(t;Q)}$, estimator of the time-dependent standard deviation, represents the (wide-band) amplitude of the airflow ($y = AF$). Slow coherent patterns are present in panel (D) but distinct from the ones of panel (C): this allows to distinguish heart FM related to breathing AM rather than FM (or to both of them). It remains to interpret observed phase, close to $+\frac{\pi}{2}$, in relation to the actual amplitude modulation: the minus sign brings $\pm\pi$, the time derivative adds $+\frac{\pi}{2}$ (and a negligible log-frequency shift $-\frac{1}{Q^2}$), and the logarithmic function is a non-linear transformation of the waveform that does not affect its phase. Therefore, oscillations in $\Im\{F_{AF}(t;Q)\}$ are shifted of $-\frac{\pi}{2}$ compared to the breathing amplitude, which means that the heart FM has close to no phase difference with breathing AM. The horizontal line at the breathing frequency is also present, which only reflects the carrier wave shared between $\Im\{F_{AF}(t;Q)\}$ and $\Re\{F_{AF}(t;Q)\}$ (hence identical to panel (C) up to the phase quadrature).

2.4.5 Statistics over *shhs2* of the phase between heart rate and breathing rates

Once again, we compute the coherence-based distribution of the cardio-respiratory phase as a function of the modulation frequency over the 2.3 years of sleep recording in the *shhs2* database. We focus on the phase between heart rate and breathing rate in Fig.IV.21 (A) and the one between heart rate and breathing amplitude (B). The individual cardio-respiratory coherence are the one from Fig.IV.20 (C, D) based on the generic (wide-band) rate signals of the ECG, $x = \Re\{F_{ECG}(t;Q)\}$, and of the airflow recordings, $F_{AF}(t;Q)$ (real and imaginary parts). In order to represent the correct phase of the breathing AM in panel (B), we shift the associated rate signal by $\frac{\pi}{2}$ using $y = i\Im\{F_{AF}(t;Q)\}$.

The use of rates signals improves the observation of the slow cardio-respiratory coherence. In Fig.IV.21, the slow mode (2^{-6} to 2^{-4} Hz) appears more clearly, with a higher coherence and smaller angular spread, when the modulations of the frequency or amplitude of the airflow oscillation is used instead of the bare recording (Fig.IV.16). In this frequency range, the respiratory frequency and amplitude modulations evolve (nearly) in phase synchrony with the cardiac frequency.

We can further expand the observation to very slow frequencies ($< 2^{-6}$ Hz), for which respiratory frequency and amplitude modulations differ. While the phase difference between very slow modulations of the heart rate and the breathing amplitude shifts only slightly from zero, the very slow breathing rate modulations are in phase opposition with the corresponding ones in the heart rate.

Finally, the fast coherent mode associated to RSA (at the breathing frequency 2^{-2} Hz) is still visible. As discussed previously, the origin of this fast modulation in the breathing rate signals is the non-circularity of the recorded breathing waveform. Here, we push the limit of

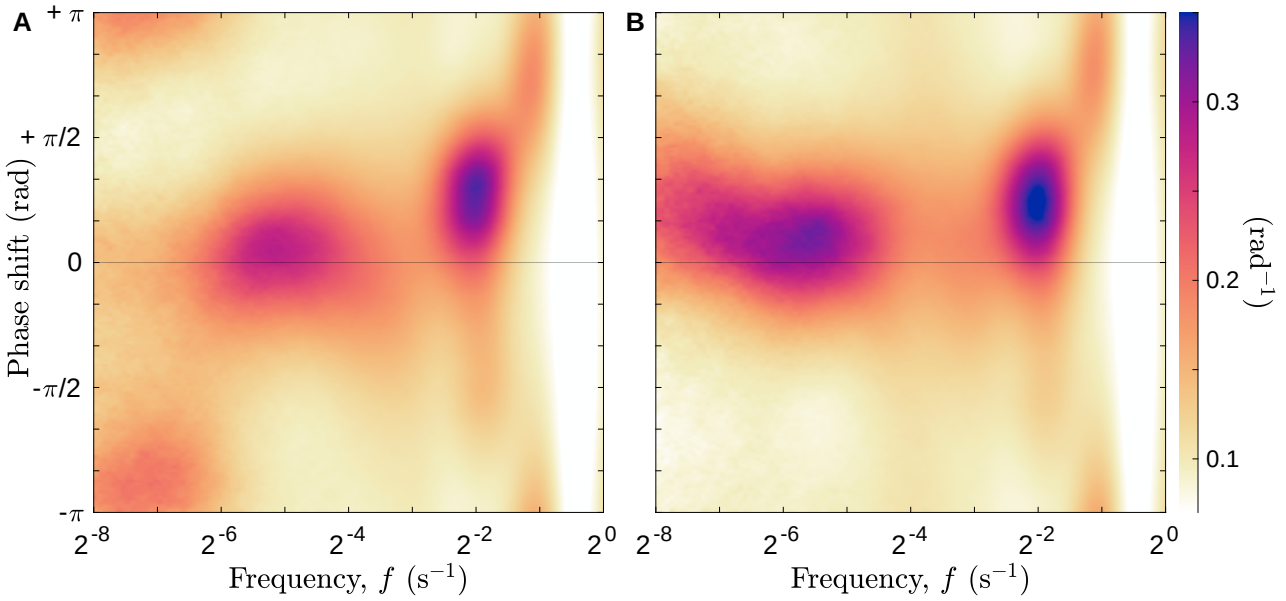


Figure IV.21: *Comparison of the phase distributions $G_{xy}(\phi\,;f)$ in shhs2 between heart rate and breathing rates modulations. (A) Cardiac frequency ($\Re\{F_{ECG}(t;Q)\}$) vs. respiratory frequency ($F_{AF}(t;Q)$). (B) Cardiac frequency vs. respiratory amplitude ($i\Im\{F_{AF}(t;Q)\}$).*

its interpretation as a rate modulation by considering this non-circularity as a rate modulation at the carrier (breathing) frequency. This allows to assess the presence of RSA. The behaviour at higher frequency ($> 2^{-2}$ Hz) cannot be interpreted easily.

Even though these results have been obtained from the generic estimators, the same computation with more precise estimators only improves marginally the angular distributions (the coherence is slightly increased). We show it in Fig.B.24 (Appendix 3.1), for which the precise heart rate estimator is used instead of the generic one. Because of the suppression of the fast cardiac oscillation in the precise estimator, the main difference concerns very high frequencies ($> 2^{-2}$ Hz), which are not relevant to breathing modulations.

A description of these two slow and very slow scales of cardio-respiratory interaction can be found in [Thomas, Mietus, Peng, and Goldberger 2005], termed high-frequency and low-frequency cardio-pulmonary coupling. It is quantified as a positive index, intermediary between a cardio-respiratory coherence and their cross-spectral density (it is half-normalized), where the respiratory signal has the particularity of being derived from the ECG amplitude modulations. Our coherence approach allows putting into evidence the phase of between modulations, in particular a positive and a negative sign of the slow and very slow FM-FM cardio-respiratory coherence respectively. Their antagonist mechanism allows the regulation of the ratio of the cardiac and respiratory frequencies, that is investigated in time in the next section.

3 Cardio-respiratory coordination as a frequency ratio

This section aims at approaching the topic of synchronization between the cardiac and respiratory rhythms, that is not in the reach of the coherence analysis. This phenomenon consists in the settlement of the cardiac and respiratory frequencies close to a simple rational number $q = \frac{a}{b}$: it occurs when the heart beats a times while b breathing cycles are completed, for a significant time duration. A strong cardiac FM versus breathing FM coherence with no phase difference, as in Fig.IV.20 (C), is necessary for the observation of synchronization, because $f_c(t) = qf_r(t)$. But this is not a sufficient condition, since the cardiac and respiratory frequencies can vary together non-proportionally.

Usually described as a phase locking phenomenon [Pikovsky, Rosenblum, and Kurths 2001], the synchronization process is described as the result of a dynamical interaction between phases [Kralemann, Cimponeriu, et al. 2008; Kralemann, Frühwirth, et al. 2013; Rosenblum et al. 2001]. Thus, the cardio-respiratory synchronization is commonly investigated from their phase, with a stroboscopic representation (of the breathing phase at successive ECG pulse) called synchrogram [Schäfer et al. 1998]; it facilitates the determination of the number $q = \frac{f_c}{f_r} = \frac{a}{b}$ when it is rational. This dimensionless number, called the pulse-respiration quotient (PRQ) [Hildebrandt 1976; Moser, Lehofer, et al. 1995; Scholkmann and Wolf 2019], is a crucial parameter for the cardio-respiratory system, biophysically more relevant than the breathing or heart rates alone.

3.1 Distribution of frequency ratio

Ideally, the determination of the dynamics of this ratio results from the equally ideal determination of the breathing and heart rates. Although the time-frequency estimation of their modulations is quite straightforward and sufficient for the coherence analysis, a quantitative and stable estimation is more difficult to obtain, as illustrated for the heart rate. We propose to use instead a distribution of this ratio, that concentrates on the desired line $q(t)\frac{f_c(t)}{f_r(t)}$ in the ideal case, and remains spread when the rhythmic oscillations lose their regularity. This happens occasionally for the breathing rhythm, more rarely for the heart rate (atrial fibrillation for some subjects), and can have an instrumental origin when the signal is transiently lost during a motion artefact for instance.

Given a cardiac recording x and a respiratory one y, their power densities appear as appropriate distributions for the frequency of their rhythm in the stationary case. The distribution of the pulse-respiration ratio q can be defined from the following integral:

$$R(q) = \int S_{xx}(f) S_{yy}(qf) |qf| \mathrm{d}f \quad , \tag{IV.12}$$

that is a multiplicative convolution (or cross-correlation), corresponding to the regular (additive) one when considering the variable $\log f$ (for $f > 0$). Notice that we are in fact comparing power log-frequency density here, naturally associated to the scale-invariant (Haar) measure $\frac{\mathrm{d}f}{|f|}$, so that it integrates to the power. Whenever the power of each recording is normalized, $P_x = 1 = P_y$, the ratio distribution is also normalized $\int R(q)\frac{\mathrm{d}q}{|q|} = 1$, and dimensionless.

This formalism extends to time-dependent distributions in the wavelet framework, the simplest of which is obtained directly from the wavelet transforms as:

$$R_{xy}(t, q; Q) = \int |X(t, f; Q) Y(t, qf; Q)|^2 \frac{\mathrm{d}f}{|f|} \frac{Q^2}{\pi} \quad . \tag{IV.13}$$

The introduction of the quality factor Q explicitly fixes the time-ratio uncertainty trade-off of this expression, where the ratio resolution is related to the frequency resolution as $\delta \log q = \sqrt{2}\delta \log f$.

An illustration of Eq.(IV.13) is given in Fig.IV.22 (C), where X and Y originate from the same signal. Based on a single PPG $z(t)$ (A) and its wavelet transform $Z(t, f; Q)$ (B), we notice that this recording contains both the cardiac and breathing rhythms. We thus split the frequency domain above and below $f_0 = 1$ Hz (red dashed line) to define the cardiac and breathing frequency distributions from $X(t, f; Q) = Z(t, f; Q)\Theta(\log f/f_0)$ and $Y(t, f; Q) = Z(t, f; Q)\Theta(\log f_0/f)$ (with Θ the Heaviside step function). Their ratio distribution Eq.(IV.12) is then obtained in (C) from a cross-correlation with respect to $\log f$ (we only keep positive frequencies). We used the quality factor $Q = 15$ to get a correct ratio concentration, but we cannot increase it too much not to lose the time resolution. This results into numerous

harmonics that are also participating in the ratio distribution as secondary lines of the form $\frac{m}{n}q(t)$ for small integers m, n. These ratio lines are cross-terms that are also present when comparing the entire PPG to itself ($X = Y = Z$), however, they are mixed with the always dominant auto-terms with fundamental line $q = 1$, justifying the splitting of the frequency range.

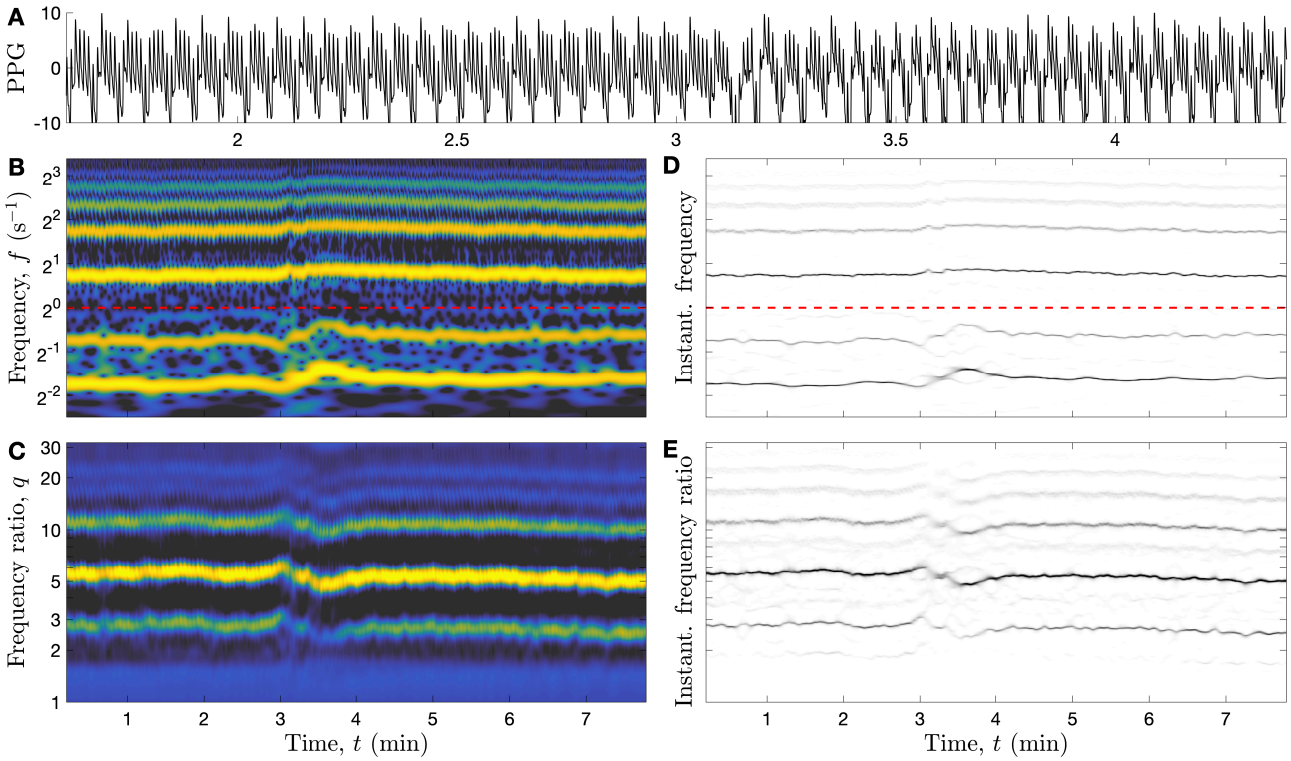


Figure IV.22: *Distribution of the ratio between the heart and breathing rates of subject 2 from a single PPG recording. (A) PPG signal. (B) Wavelet transform of the PPG with the quality factor $Q = 15$ (squared modulus on a logarithmic colour scale). The horizontal red dashed line at 1 Hz separates the respiratory rhythm form the cardiac one. (C) Ratio distribution computed from the power densities obtained above and below 1 Hz in (B). (D) Frequency-reassigned power density of the PPG. (E) Ratio distribution based on the frequency-reassigned power densities (D) above and below 1 Hz.*

3.2 Distribution of instantaneous frequency ratio

As discussed in the chapter III, the frequency uncertainty is a constraint on the smallest distinguishable interval between simultaneously oscillating components of the signal, rather than a maximal achievable precision for the determination of their frequency. Indeed, the latter one can be greatly improved from the frequency reassignment procedure Eq.(III.17), that exploits the time derivative of the phase or instantaneous frequency $\Re\{F_x(t, f; Q)\}$ Eq.(III.9), see Fig.IV.22 (D). We use this ultimate sophistication to achieve the goal of concentrating the ratio distribution near the ideal line $q(t)$, as represented in Fig.IV.22 (E). This operation is explicitly expressed as:

$$\tilde{R}_{xy}(t, q; Q) = \iint |X(t, f; Q)Y(t, f'; Q)|^2 \delta\left(q - \frac{\Re\{F_x(t, f; Q)\}}{\Re\{F_y(t, f'; Q)\}}\right) |q| \frac{\mathrm{d}f\,\mathrm{d}f'}{|f f'|} \frac{Q^2}{\pi} \quad , \qquad \text{(IV.14)}$$

where the tilde refers to the frequency ratio reassignment procedure. The concentration of the resulting cardio-respiratory rate distribution is greatly improved, so that a much smaller quality

factor can be used to retrieve a precise time resolution. For rhythmic signals, we do not observe any noticeable difference using $\delta\left(\log \frac{q|F_y|}{|F_x|}\right)$, and we restrict frequency integrals to the positive side. This could be considered as a marginal distribution for the more general distribution that also considers the phase of F_y/F_x, interpreted from Eq.(III.23) as a delay related to vertical features or strong modulations.

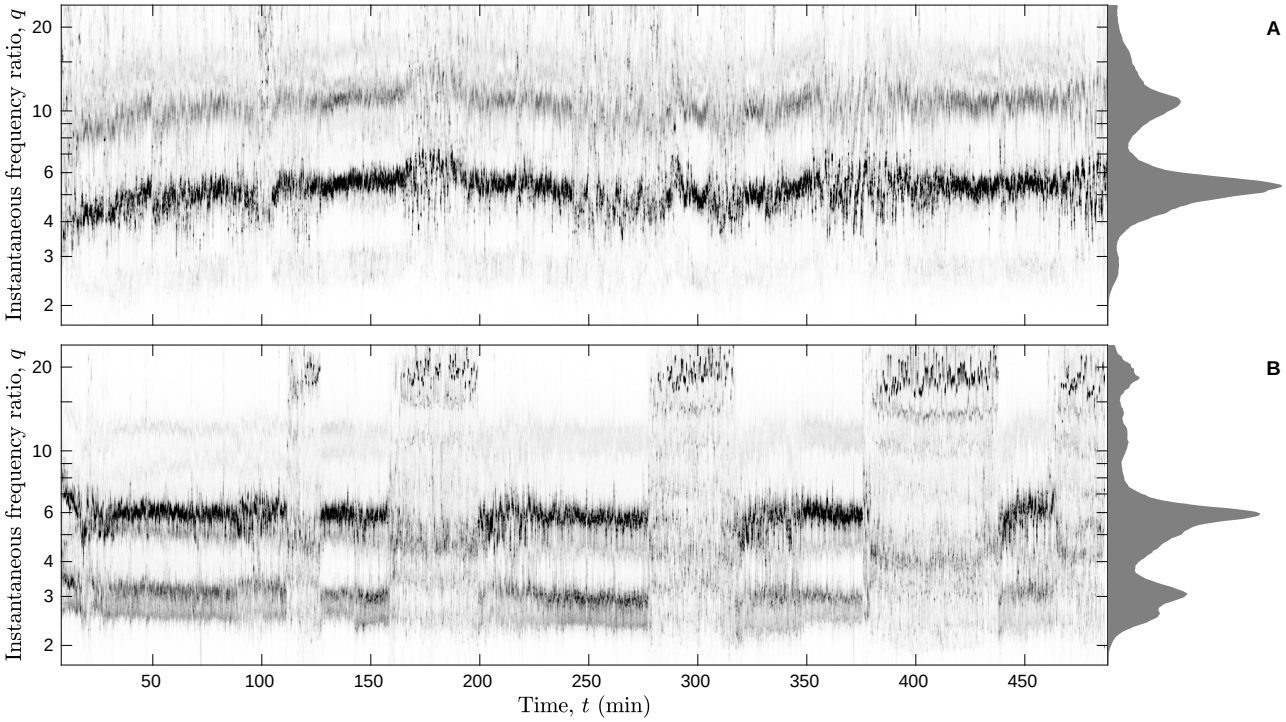


Figure IV.23: *Distribution of the ratio between the cardiac and respiratory instantaneous frequencies, $R_{xy}(t, q; Q)$, during the full 8 h night of sleep for subjects 4 (A) and 7 (B). The recordings are respectively $x = aECG14$ and $y = AF$. The quality factor is $Q = 4$ and the ratio distribution normalized at each time for clarity. The time-averaged ratio distribution is given on the right.*

A second example is given in Fig.IV.23 for 8 h of sleep of two different subjects (A, B). Instead of the two-in-one PPG, we use two distinct recordings: the airflow signal for the breathing rhythm, and the aECG14 signal for the cardiac one. We prefer it to the raw ECG because its harmonics are more damped, thus clarifying the distribution. It is indeed convenient to have a dominant fundamental component, so that the line of interest $q(t)$ is the most weighted one. The harmonic contributions are further limited by choosing a quality factor as low as $Q = 4$, so that the fundamental component is still well-separated from harmonics, but harmonic components are not distinct. The influence of second harmonics can still be noticed as echoes an octave above or below $q(t)$.The ratio distribution is also normalized at each time for a clearer representation.

For the subject 4 of panel (A), which is the one of the coherence analysis, the PRQ wanders between 4 and 6 during the night, without settling around a particular value, except perhaps rare stops near $\frac{11}{2}$. We can further observe important differences in the fluctuations of the PRQ: sometimes it remains concentrated with short fluctuations around a well-defined value, whereas at other times, it fluctuates a lot more. These stable and unstable episodes are in direct correspondence with episodes of respectively strong and weak RSA coherence in Figs.IV.14 and IV.20. The observation of the slow coherence in Fig.IV.20 (C, D) tends to evolve in the opposite manner, although this would require more confirmations.

This characterization of steady and unsteady episodes from the fluctuation of the PRQ is

reminiscent of the time-delay stability analysis [Bashan et al. 2012]. Applied on the cardiac and respiratory rates [Bartsch, Liu, Ma, et al. 2014], it bears a close relation with the phase methods developed here (both the instantaneous frequency ratio and the temporal or time-frequency coherence).

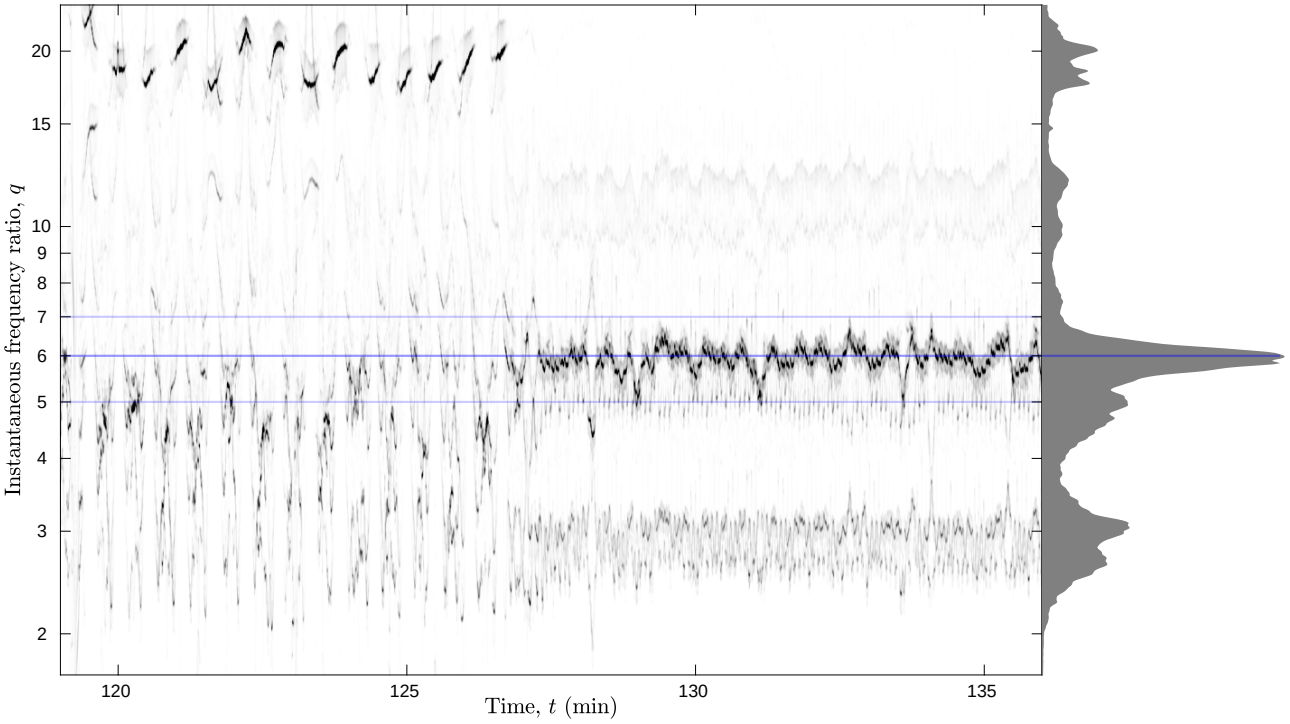


Figure IV.24: *Distribution of the ratio between the cardiac and respiratory instantaneous frequencies for subject 7. Zoom of Fig.IV.23 (B) on a selected time interval that contains a transition from a state of repetitive obstructive sleep apnea to a state of 6:1 cardio-respiratory synchronization.*

In addition to the stable and unstable episodes, the distribution of the PRQ for the subject of panel (B) exhibits a third type of behaviour, that consists in a fast oscillation between extreme values of the PRQ, about 3 to 20. This is the repetitive occurrence of obstructive sleep apnea: the subject periodically stops breathing during about 20 heart beats, before recovering from apnea and accelerating breath up to once every 3-5 heart beats. This oscillation is better distinguished on Fig.IV.24, zoom of Fig.IV.23 (B) on a 15 min time selection that contains a transition from obstructive sleep apnea to a steady episode. At this timescale, we can better appreciate the precision of the time-ratio resolution for $Q = 4$. Contrary to the previous subject, the stable episode here clearly fluctuates close to a stationary value that turns out to be an integer number, this is a 6:1 cardio-respiratory synchronization. This corresponds to the observation that episodes of cardio-respiratory coupling are more frequent for apneic subjects [Riedl et al. 2014]. The rich phenomenology of the cardio-respiratory interaction [Bartsch, Schumann, et al. 2012] is reflected into the important or weak fluctuations, locking to a simple ratio or oscillating behaviour of the PRQ. These episodes are put in correspondence with REM and non-REM sleep stages at the end of chapter V.

3.3 *Sonance* in a nutshell: towards a multi-frequency synchronization index

The PRQ distribution is not perfectly concentrated on a line, even with a careful choice of the quality factor and selection of the signals. This is an effect of the non-circularity of the

oscillation, and the associated phenomenon of harmonics and beats. We notice that the inclusion of these harmonics, as integer multiples of the fundamental frequency, does not affect the rationality of the frequency ratio distribution in case of synchronization.

Can we define a synchronization index for the distribution of rational numbers? Although we can not provide a simple and definitive answer on this question, we can propose a side-trip around this question, that conveys an important musical meaning. Indeed, in the high quality factor context of Music ($Q > \frac{12}{\log 2}\sqrt{\top} \approx 43$), the non-circularity is a desirable timbre of sounds, and their synchronization is called consonance. Harmonics participate in the sensation of consonance or dissonance [Sethares 2005], and was already studied in details by pioneers of biophysics [Helmholtz 1870]. The synchronization of the heart and breath resemble two voices that sing at unison, except that these phenomena occurs with respectively very low and very high intrinsic quality factors.

The extensive and slightly off track exploration of this question is compiled in an independent "bonus" part C, that is self-consistent. We summarize here briefly our findings. A singular measure of rationality or irrationality can be defined from the definition of a statistical ensemble over (positive) rational numbers $q = \frac{a}{b} \in \mathbb{Q}^+$. This measure is based on an energy (or "Hamiltonian") of the form $\log ab$, where a, b are relatively prime. Introducing an "inverse temperature" exponent, that we have denoted σ, we can introduce "Boltzmann weights" $e^{-\sigma \log ab} = (ab)^{-\sigma}$ and a partition function:

$$\mathcal{Z}_\sigma = \sum_{\frac{a}{b}\in\mathbb{Q}^+} (ab)^{-\sigma} \quad , \tag{IV.15}$$

associated to the *sonance* measure of a ratio distribution $R_{xy}(q)$:

$$\mathscr{C}_\sigma[R_{xy}] = \mathcal{Z}_\sigma^{-1} \sum_{\frac{a}{b}\in\mathbb{Q}^+} R_{xy}\left(\frac{a}{b}\right)(ab)^{-\sigma} \tag{IV.16}$$

$$= \int_0^\infty R_{xy}(q)\mathrm{d}\mathscr{C}_\sigma(q) \quad . \tag{IV.17}$$

Its definition is unambiguous for $\sigma > 1$, but $\sigma \leq 1$ requires to specify an enumeration of rational numbers, which can be obtained from successive rational approximations in binary trees [Calkin and Wilf 2000]. In practice, we stop the series at a certain height in the tree.

The resulting measures are fractals, see Fig.2 in the bonus part C, and we can identify special cases: $\sigma \to +\infty$ is concentrated on $q = 1$, the most rational number, $\sigma \to -\infty$ is concentrated on $\phi = \frac{\sqrt{5}+1}{2}$ and $\phi^{-1} = \frac{\sqrt{5}-1}{2}$ the Golden —most irrational— numbers, $\sigma \to 1$ seems to converge very slowly to an infinitely flat measure, and $\sigma \to 0$ is an interestingly peculiar measure called Minkowski's question mark function [Alkauskas 2008; Minkowski 1910]. This "infinite temperature" limit, i.e. with an equipartition among rational numbers, is in fact a measure of irrationality (since rationals are dense in reals).

This measure can be implemented efficiently in the Fourier spectral domain conjugated to $\log q$, called its Mellin transform. The same strategy yields an efficient numerical implementation of the multiplicative convolution Eq.(IV.12). Although we obtained the expected behaviour of a dissonant measure in the idealized situation of discrete distributions (for which synchronization yields a vanishing value), its application to concrete situations is difficult to interpret. We have just started to unfold the thermodynamic formalism, by considering the first derivative with respect to σ, and another exponent (the norm p not necessarily equal to 2 in Eq.(IV.13)).

Summary of sections 2 and 3

Computed from the ECG and airflow (AF) recordings and their rate signals, the cardio-respiratory coherence and its significance is evaluated individually in time and frequency, and the distribution of its phase is obtained from collective statistics over all times and subjects in the *shhs2* database. The comparison of different respiratory signals, either the bare recordings or amplitude and frequency modulations (AM and FM), to the heart rate variability (HRV, or cardiac FM) highlights different aspects of the cardio-respiratory interaction. The generic wide-band estimator proves to be sufficient to estimate modulations of the cardio-respiratory rates, not only from the airflow recording but also from the ECG.

Overall, three different scales of coupling of cardio-respiratory rhythms as been observed. The fast mode (2^{-2} Hz) is commonly known as respiratory sinus arrhythmia (RSA): the heart rate increases during inspiration and decreases during expiration. Straightforwardly observed by comparing HRV to a raw respiratory recording, RSA coherence is also visible using respiratory AM and FM signals, thanks to an echo of the breathing carrier wave that enriches the rate signals obtained from generic estimators.

Respiratory rate signals are crucial to investigate slow dynamics, circumventing the limitations of the instrumental high-pass filtering in the airflow recording that damps slow oscillations, dilutes the apparent slow coherence and potentially biases its phase. A slow mode (2^{-6} to 2^{-4} Hz) indicates that both the breathing frequency and amplitude evolve in-phase with the heart rate at this scale. For the breathing AM, this slow coherent component extends to very slow frequencies ($< 2^{-6}$ Hz) with a very slight phase shift, while a distinct very slow mode in phase opposition appears between the heart and breathing FM.

We may understand the difference of sign between these slow and very slow modes as a regulatory mechanism made of a positive and a negative feedback loop between the cardio-respiratory rates: at the scale of about 30 s, heart and breathing rhythms agrees in their variations to adapt gaseous exchanges to the need of the metabolism; in contrast, to the slow deceleration of one rhythm at the scale of a few minutes, the other rhythm respond with an acceleration (and vice versa). These antagonist mechanisms are essential to regulate the ratio between the heart and breathing rates.

This ratio, also called pulse-respiration quotient, can be estimated directly from the pair of cardio-respiratory signals as a time-dependent distribution of their frequency ratio. The use of frequency-reassignment is especially suited to concentrate this ratio on a horizontal line, whose fluctuations can be precisely resolved in time using a small quality factor. Distinct episodes of weak or important variability, locking to a simple rational number or important oscillations can be identified in relation to sleep phenomena such as stages and the occurrence of apnea.

Chapter V

Sleep phenomena: wavelet perspectives on polysomnography

In this last chapter, we focus on three types of recordings from polysomnography databases: the electroencephalogram (EEG), the electrocardiogram (ECG), and the airflow (AF). Among these, the EEG remains the most complex, because its spectral signature is a mixture of rhythms of different natures: some of them have been recognized with a physiological origin, others which are more volatile (unsteady) can be interpreted falsely from spectral decomposition [Jones 2016]. The correlations between these EEG "rhythms" and other physiological signals (such as the heart and breathing rates) can help decipher this complexity; we propose a methodology to assist this clarification.

The cardiovascular system is vital for feeding and clearing the whole body organs: its failure in the brain or other neural tissues leads rapidly to irreversible issues, it must therefore be finely regulated to keep a correct flux and filtration of blood. The cerebral blood flow has been reported to increase during sleep, both in slow wave (NREM) sleep (4% to 25%) and in REM sleep (25% to 80%) [Kryger, Roth, and Dement 2005]. Recently, it was also shown that the brain rhythms can be placed in resonance with the HRV and respiration when modulating the respiration frequency to lower bands [Hinterberger et al. 2019]. Continuous time-frequency representation is reported in [Prerau et al. 2017] as a key to simplify the reading of EEG, both globally and locally, and in [Aviyente, Bernat, et al. 2011; Aviyente and Mutlu 2011; Aviyente, Tootell, and Bernat 2017] as a way to quantify functional integration in the brain as a phase synchrony in EEG.

In Fig.V.1 (B, C, D) we illustrate simultaneous EEG, ECG and AF recordings from subject 3 (described in Appendix 1). As summarized by the *hypnogram* (panel A, from a clinician annotation), a transition from NREM sleep to wake phase can be noticed on the three signals as drastic changes in their behaviour near 51.5 min. The amplitude of these three recordings is also represented in the time-frequency plane; these images are identical to scalograms (wavelet-based spectrograms obtained as the squared modulus) because of their logarithmic colour scales. We recognize the fundamental modes of ECG ($\sim 2^0 = 1$ Hz) and AF signals ($\sim 2^{-2} = 0.25$ Hz) in Fig.V.1 (C, D), and some of their harmonics (first three, only visible before 51 min for the AF).

The EEG in Fig.V.1 (B) is completely different, there is no clear fundamental mode: it is a mixture of complex dynamics spread over a large frequency range (at least up to 125 Hz and down to the instrumental cut-off visible near 2^{-4} Hz, illustrated in this example), but also concentrated at specific frequencies. During the wake stage, α waves, which are typical of the phase of wakefulness with closed eyes [Prerau et al. 2017], are visible in a very thin frequency band near 10 Hz. A similar band, much more intermittent and less intense, can be noticed near 12 Hz during the NREM sleep stage: it is the σ band composed of bursts of sleep

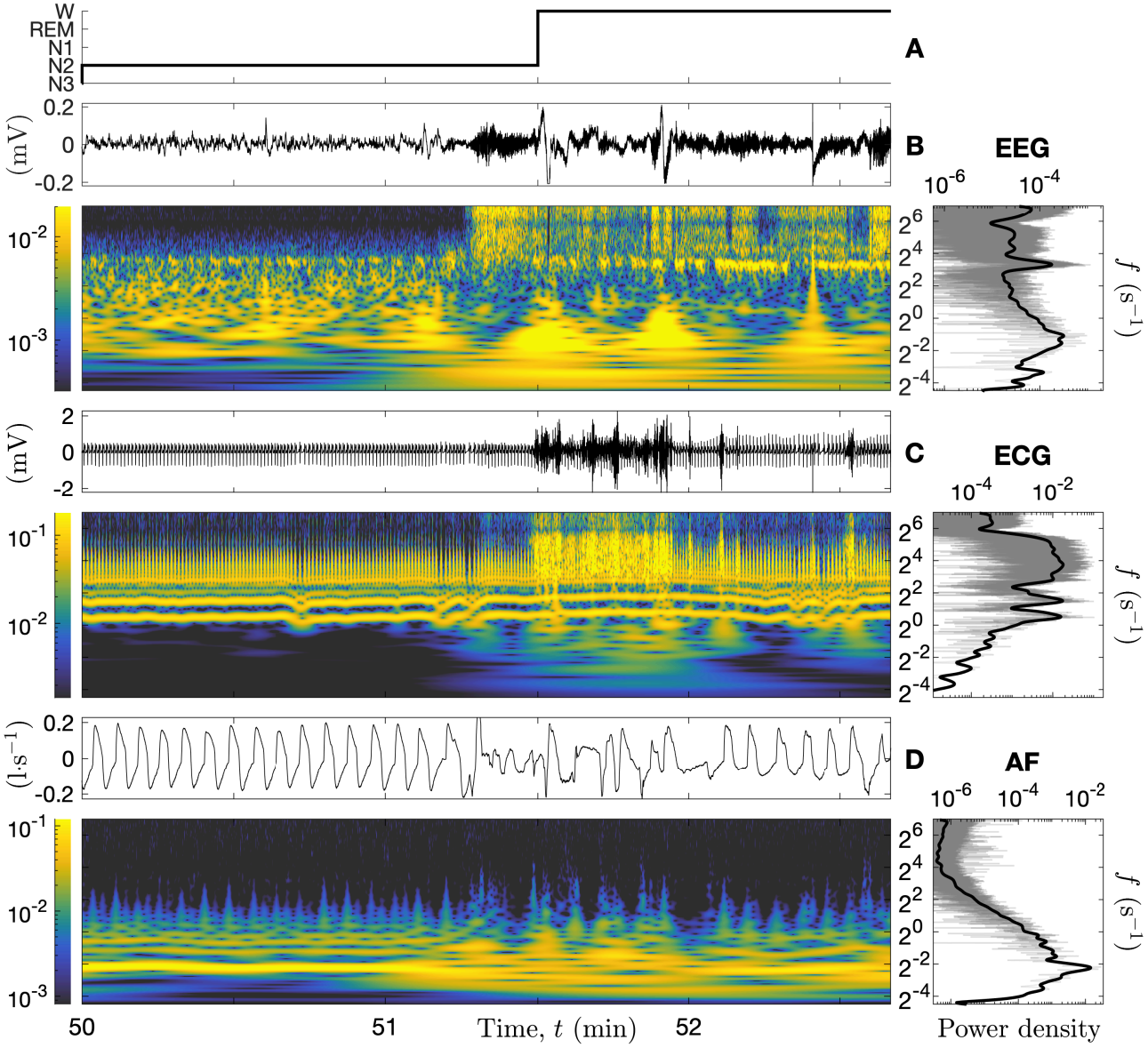


Figure V.1: *Comparison of three polysomnographic recordings of subject 3. (A) Hypnogram of the person (black line), who is shifting from NREM sleep stage 2 (N2) to wake phase (W). (B) EEG (C3-O1) in millivolt, (C) ECG in millivolt, and (D) nasal-oral airflow (AF) in litre per second. For each recording, the signal (top) and the power density (right) are aligned with the corresponding wavelet representation with quality factors $Q = 10$. Only its amplitude (twice the modulus) is colour-coded, with logarithmic scales in signals' unit. The power log-frequency density $S_{xx}(f)|f|$ (in signal's unit squared) is estimated from either the squared Fourier (thin grey line) or wavelet (thick black line) transforms on the selected time interval.*

spindles. Below 4 Hz we observe localized bursts (with vertical cone rather than horizontal band shape) corresponding to sharp and sudden events in the signals. This time-frequency representation is very helpful to recognize different components; singular events are expressed as vertical structures, whereas periodic components translate into horizontal bands. The latter are summarized as a peak in the power log-frequency density $S_{xx}(f)|f|$, estimated either from the wavelet or Fourier transforms as in Eqs.(II.23, II.25).

In this chapter, we further develop our discussion of the nervous system, previously limited to the sympatho-vagal regulation of the cardio-respiratory activity, to the brain activity that can be recorded from the scalp. In every polysomnography, it is recorded as a multi-channel electroencephalography (EEG); the scale-free property of the wavelet transform appears especially well-suited to its multi-scale and non-stationary dynamics, as shown in Fig.V.1 (B). We describe the global architecture of its sleep-wake patterns, structured in frequency-bands and time-stages, both from the perspectives of power density and coherence. Then, we introduce EEG band modulation signals, that we compare to other physiological modulations signals in terms of coherence. In particular, we highlight the influence of sleep apnea, both through individual examples and collective statistics. In the last section we approach the local phenomenon of sleep spindles, made of short wave trains.

1 Neural activity from EEG

Shared among most animals, the nervous tissue is composed of a network of nerve cells, or neurons, that convey chemical and electrical influxes across different parts of the body. The central nervous system, in particular the brain, is a hub for this information flow made of a dense connected network of neurons. Short electrical impulses, called action potentials, travel along individual neurons as wavefronts of de- and re-polarization of their membrane, lasting about 1 or 2 ms and followed by an equivalent or longer refractory period. Collective neural behaviours effectively produce oscillating electrical dipoles, whose resulting electric field can be recorded on the head.

1.1 Electroencephalogram

The brain activity can be measured with electrodes recording the fluctuating electric potential (usually in microvolts, μV) between different standardized locations on the scalp. The resulting electroencephalogram (EEG) exhibits both erratic (noisy and/or scale-free) and rhythmic behaviours in a very wide range of frequencies [Buzsáki and Draguhn 2004]. It represents the integrated activity of a network of multiple neurons throughout the brain, in particular the one of cortical neurons that is the least attenuated by conduction through layers of tissues and the skull due to the proximity with the electrode.

1.1.1 Spatiotemporal characteristics

The duration of the action potential sets the fastest temporal scale (a maximum frequency around up to 500 Hz) for the resulting neural activity, that spans multiple scales down to the lifetime. The human brain signals found in a polysomnography (EEG) can only record oscillations in a bandwidth of three decades (from about 0.1 to 100 Hz), during a full night of sleep. These oscillations can be modulated at much slower scales, reflecting the activity of the body (digestion, sleep-wake, etc), that are only limited by the duration of the recording.

The multiple channels of the EEG also constitute a discrete spatial measurement of the electric field, positioned according to the 10-20 system as illustrated in Fig.V.2. The spatial

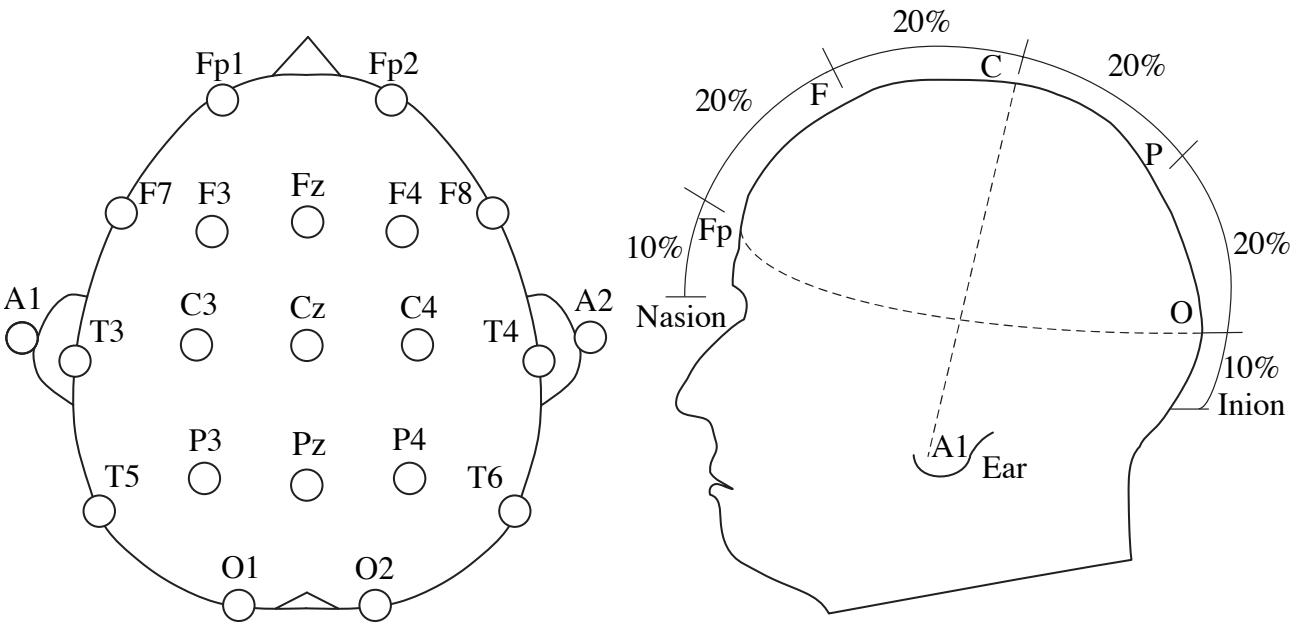


Figure V.2: *Standardized electrode placement for EEG, called the 10-20 system in reference to the percentage of the distance on the skull that separates reference points (nasion to anion and between ears). Cortical zones are labelled as follows: frontal pole (Fp), frontal (F), central (C), parietal (P), temporal (T), occipital (O) and auricular (A). Odd (respectively even) numbers refers to the left (right) side of the head, whereas z refers to the sagittal plane. Adapted from the thesis of [Hoffmann 2007].*

resolution increases with the density of electrode, up to a certain limit due to volume conduction. However, only a few electrodes were recorded in the polysomnography databases we use.

In the polysomnography of subject 3, a single bipolar EEG is available, measured between different pairs of locations depending on the subject (C4-A1, O2-A1 or C3-O1). Its resolution is about 0.1 μV, sampled at 250 Hz, without specification of the instrumental filter. In *shhs2* (subjects 4, 5, 6), the two bipolar EEG signals are recorded to the central left and right zones, C3-A2 and C4-A1, where A refers to the mastoids, bones assumed to be electrically "quiet" and commonly taken as contralateral reference electrodes. The resolution is 1 μV, sampled at 125 or 128 Hz with an instrumental high-pass filter at 0.15 Hz. The polysomnography of subject 7 contains a unipolar multi-channel EEG, from at least 11 electrodes (F3, F4, Fz, C3, C4, Cz, Pz, O1, O2, A1, A1) with a reference electrode on the forehead (the sagittal plane is also assumed "quiet" since above the *corpus callosum*). Its resolution is about 0.2 μV, sampled at 256 Hz, without specification of the instrumental filter. A unipolar recording gives the possibility to re-reference the EEG offline, for instance by taking the difference between two chosen electrodes, $x_i(t) - x_j(t)$.

1.1.2 The reference problem

Beyond pairs (bipolar EEG), several referencing methods can be applied to a unipolar EEG for a greater spatial resolution and a better control of artefacts. We refer to this arbitrariness as the reference problem. A first approach aims at approximating an ideal reference, the electric potential at infinity, inexistent in practice because of electromagnetic interferences in our environment. The simplest method consists in averaging all channels into an average reference [Osselton 1969], based on the assumption of a globally neutral field. Weights can be introduced in the construction of a more precise neutral reference, that can include corrections from a realistic model of volume conduction in the human head [Hu et al. 2019; Yao et al. 2019]. A common reference has the drawback of introducing residual oscillations that are firmly non-

local. The second approach consists in constructing local references for each electrode from neighbours. This guaranties the spatial locality of the recording, which does not represent the electric potential any more, but a discrete approximation of its surface Laplacian [Hjorth 1975], proportional to the current source density (from Poisson equation and Ohm's law [Gulrajani Sept.-Oct. 1998]). The spherical geometry is sufficient to increase spatial resolution at a low computational cost [Kayser and Tenke 2006b, 2015; Perrin et al. 1989], (further improved with a realistic model of the skull [Babiloni et al. 2001]). The temporal distinction of successive events at different locations is also improved [Burle et al. 2015].

In addition to bipolar EEG recording, we experimented the average reference, most straightforward to apply, and the spherical surface Laplacian, for which a simple implementation from conversion matrices is available [Kayser and Tenke 2006b]. Although the latter method should be preferred, its physical interpretation is rather rough from only 11 electrodes (30 is considered a minimum [Kayser and Tenke 2006a]). Eventually, our use of re-referencing essentially aimed at the reduction of artefacts, rather than a better spatial resolution, that was not fully exploited.

1.1.3 A note on grounding

Invisible in the EEG data, an electrode is used as a ground for the differential amplifier and the shielding. This ground is isolated from the one of the power supply, for electrical safety reasons. It is placed on the head, on Cz in *shhs2*, in order to reduce the capacitive coupling (antenna effect) with external electromagnetic sources [Ferree et al. 2001], resulting in a reduction of power line artefacts.

No grounding of the body itself is applied, even though a connection to earth can yield physiological differences that are visible on the EEG [Chevalier, Mori, and Oschman 2006]. The degree of earthing of the subject may be a significant but hidden parameter that could be controlled.

1.1.4 EEG artefacts

Different types of artefacts are commonly found in EEG signals, either of external or physiological sources. The electromagnetic interference of the power supply produces an important but precisely localized horizontal line at 50 Hz or 60 Hz (as well as less intense harmonic lines at rationally related frequencies), that is removed by a notch (band-stop) filter. The rejection of common modes in a bipolar or well-referenced EEG recording, is indeed incomplete, partly due to unavoidable impedance mismatch between electrodes [Ferree et al. 2001]. Other narrow-band spectral lines can be found at various frequencies (with a smaller amplitude) depending on the surrounding electrical apparatus. A low-amplitude broad-band noise is also associated to the EEG amplifier.

Physiological artefacts are broadly caused by any kind of motions of the body. The EEG can directly record the electrical activity of eyes movements, facial muscles and heart beat (EOG, EMG, ECG effects), characterized by time-localized events of strong amplitude. The impedance of the skin-electrode junction (about or higher than 10 kΩ) can also be affected, for instance by a mechanical tension on the cables and electrodes, sudden (time-local) during postural changes, or subtle but regular (frequency-local) for breathing. A global drift in the recording conditions can also be caused by transpiration or drying of the conductive gel.

When separate recordings are available, part of them can be identified as a significant inter-recording coherence, or cross-talk, as showed in Fig.II.7 between EEG and ECG (A, B) and between EEG and airflow (C, D), in which case this coherence can be used to construct a corrected signal (incoherent with a chosen signal by construction). Otherwise, the removal or the attenuation of these physiological artefacts is a difficult task that can be attempted

with different algorithmic methods, see reviews [Jiang, Bian, and Tian 2019; Urigüen and Garcia-Zapirain 2015], to produce "clean" EEG signals. We have tried to attenuate high-amplitude time-events in a simple but imperfect way by standardizing the recording with its time-dependent power. However, we decided not to apply any cleaning preprocessing to the EEG signals for the applications shown in this chapter, that can either be considered as (i) keeping this additional complex physiological information as part of the measurement, (ii) a way to test the robustness of time-frequency methods, (iii) some room for improvement. The true rationale behind this choice is the will to conserve a sound wavelet framework, applied to the raw recordings, into which any processing operation can be expressed in a simple mathematical form. We did not reach a clear expression for this cleaning operation, beyond the above-mentioned partial attempts, and we found that the scale-adapted time-resolution of the wavelet is able to confine these intense events to very short durations in high frequency regions. Their possible influence on the results will thus be discussed case by case using the knowledge of their localization in time or frequency.

1.2 Structure of EEG: bands and sleep stages

As shown in Fig.V.1 (B), the complexity of the EEG can be represented in the time-frequency plane to distinguish both its non-stationary and multi-scale aspects. The traditional way of dealing with this complexity consists in discretizing time into stages and frequency into bands. Wake-sleep stages have been classified into the wake (W), rapid eye movement (REM) and non-REM (or NREM) phases, and the NREM phase is further decomposed into the stages N1 (lightest), N2 and N3 (deepest sleep). This classification is based on a set of criteria about the relative EEG power in each frequency band (possibly complemented by the EOG and EMG), inside 30 seconds time epochs (R & K system and improvements [Grigg-Damberger 2012; Rechtschaffen 1968]). The succession of these stages, available in most polysomnographic databases as a clinician annotation every 30 s, is a simplified representation of sleep called a hypnogram, see Fig.V.1 (A). Aiming at the empirical distinction of neural waves, the frequency bands are roughly defined in slices of 4 Hz and denoted with a Greek letter: δ up to 4 Hz, θ from 4 to 8 Hz, α from 8 to 12 Hz, σ from 12 to 16 Hz, β from 16 to 20 Hz, and γ above 20 Hz. For instance, the γ waves are intense in the wake state, while they vanish during NREM sleep, characterized instead by intense δ waves (especially during N2 and N3 stages). In contrast, the EEG power is low in all bands during REM sleep. We refer to Table 1 of the review [Prerau et al. 2017] for the proposed mechanism generating these neural waves. This practical discretization of the time-frequency plane remains an approximation that cannot account for the continuous dynamics and the micro-structures of sleep (such as sleep spindles and K-complexes).

1.2.1 Full night EEG in time and frequency

The time-frequency representation of an EEG can simplify considerably the scoring of the sleep stages but also their reading at the global scale of sleep, while conserving the information about micro-structures [Olbrich, Claussen, and Achermann 2011; Prerau et al. 2017]. In the same spirit, we use the wavelet transform, that appears especially suited to represent the many frequency decades that are spanned by neural oscillations [Buzsáki and Draguhn 2004]. The EEG is presented in Figs.V.3 and V.4 as the amplitude of its wavelet transform (panels A) for the entire night of sleep of two subjects, together with the corresponding hypnogram (panels B) and amplitude spectrum (panels C). We can clearly observe the alternation in time of different stages, characterized by specific frequency content in distinct bands.

For instance, the time interval from 17 to 51 min in Fig.V.3 corresponds to the person falling in deep sleep and waking up again. Before 17 min, the α band (near 10 Hz) and the presence of higher frequency, β-γ, means that the person is awake (with eyes closed), then

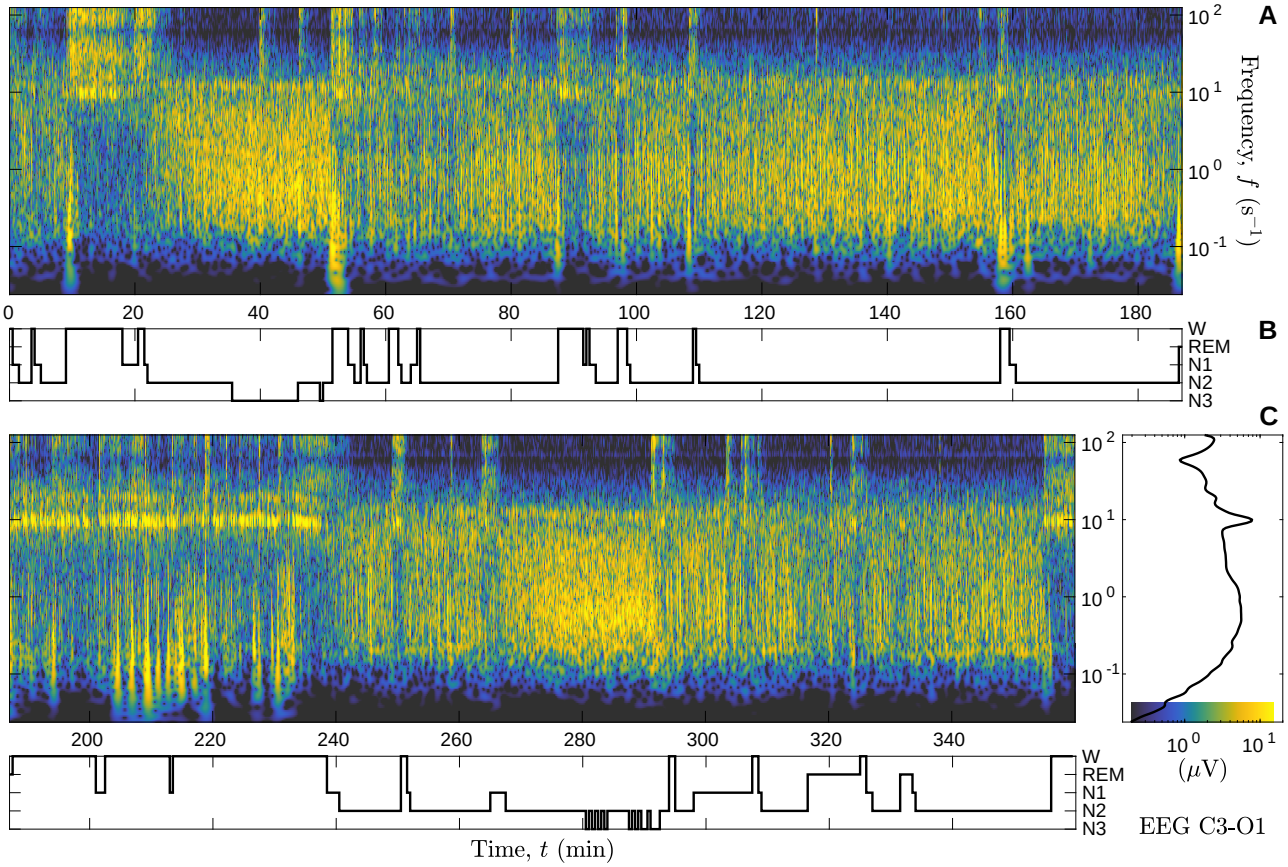


Figure V.3: *EEG (C3-O1) recorded during sleep on subject 3. (A) Amplitude of the wavelet transform, $2|X(t, f; Q)|$ with $Q = 10$ (divided in two panels). (B) Hypnogram. (C) Amplitude spectrum (root-mean-square) aligned with the colour scale (in μV).*

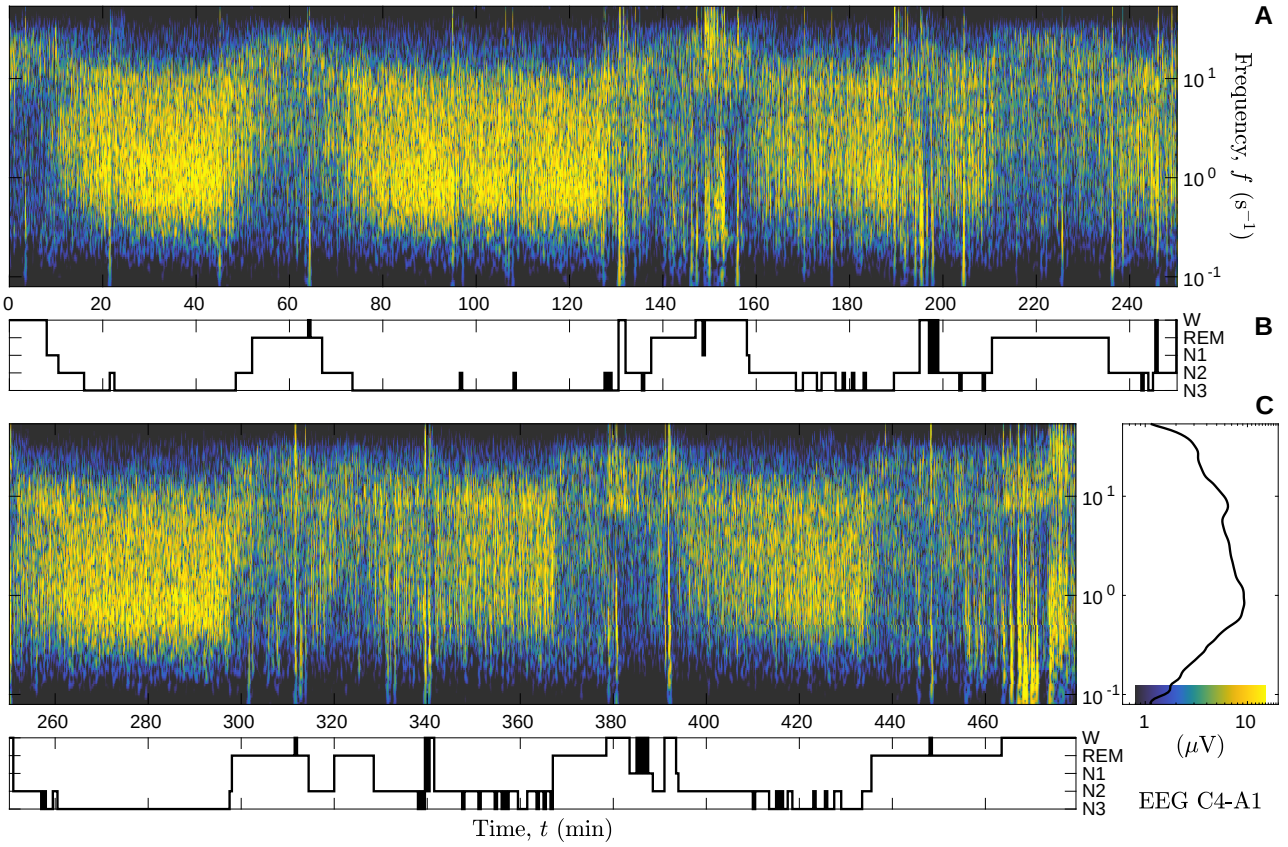


Figure V.4: *EEG (C4-A1) recorded during sleep on subject 5. Same as in Fig.V.3. Sleep stages are here structured into regular cycles of 80 min (ultradian rhythm).*

disappearing completely after some last bursts at 20 min, corresponding to the N1 stage (non-REM 1 or light sleep). Afterwards, we observe a continuous dynamics up to time 51 min: the amplitude at low frequencies (θ and δ "bands") increases, and discrete bursts appear in the σ band (around 12 Hz, the sleep spindles), occasionally interrupted by short arousal events characterized by the return of high frequencies. This has been scored as the stages N2 and N3 (deep sleep), with no obvious event at their transition. At 51 min, the succession of the burst of higher frequencies and motion artefacts (vertical line rooted in low frequencies) followed by the return of the α band marks a short wake stage. This sleep cycle is later described (in Fig.V.6) as distinct band magnitude signals, and Fig.V.5 gives a closer look at the α band during wake state.

For the rest of the night, the subject of Fig.V.3 is mainly in the N2 stage, with short waking or light sleep, and interrupted by high frequency and strong α waves (awake with eyes closed) from 187 to 237 min (motion artefacts can be observed in this interval). The deep sleep stage (N3) is only scored again intermittently between 280 and 293 min, with a short REM stage (rapid eye movement sleep) between 316 and 324 min, of equally low amplitude at all frequencies. These observations are not representative of a typical night sleep: the subject is indeed affected by a sever obstructive sleep apnea, responsible for the persistence of the intermediate N2 stage. Sleep apnea is discussed later in a distinct section.

The EEG recording of the second subject, Fig.V.4, is representative of an "ideal" sleep architecture: the night is here divided into 6 distinct and quite regular cycles. This repetition, about every 80 min for this subject, is called the ultradian rhythm. It is rarely observed with such regularity in practice, so that stochastic models are proposed as more realistic descriptions of sleep stage transitions [Kemp and Kamphuisen 1986]. The probability for each first order transition is estimated in the Appendix 4 and represented in Fig.B.26; although insufficient to describe this complex dynamics [Kishi et al. 2018], it provides an illustration of the transition network.

The idealized cycle is composed of a non-REM phase (strong slow δ waves, no amplitude at high frequencies), followed by the REM phase (low amplitude at all frequencies). In a non-REM phase, N1 is transient (very short), and alternating annotations of stages N2 and N3 (in the hypnogram) are not associated to obvious changes in the EEG: the non-REM stages appear as different degrees on a continuum, questioning their discreteness. Non-REM tends to be more important (longer and deeper in N3) early in the night, while REM tends to increase at the end of the night. Often concomitant with the wake state (W), motion artefacts are again visible as vertical singularities.

Many instrumental specificities or limitations are noticeable by comparing the first EEG in Fig.V.3 to the second one in Fig.V.4. Starting at high frequencies, they are damped (low pass filtered) near the power line frequency 60 Hz for the second recording, whereas the frequency range extends to 125 Hz in the first one. Instead, a notch filter at 60 Hz removes this major source of electromagnetic pollution, letting a hole in the γ band. We can also notice a line of constant frequency and low amplitude at 70 Hz caused by another unfiltered artificial source. Then, α waves and the σ band are much more intense and distinct for the first subject. This can be caused by difference in the montage, in particular the electrode locations, C3-O1 for the first EEG and C4-A1 for the second: the occipital area favours their observation [Prerau et al. 2017].

At lower frequencies, the breathing rhythm is slightly but surely present at 0.2 Hz in several time intervals of the first EEG, especially in the last part of the signal (after 250 min). Commonly observed in EEG of the *shhs2* database, the heart beat seems absent from this EEG; in fact it is also present with a small amplitude, as put into evidence by its significant coherence with the simultaneous ECG recording (used as an illustration of the time-frequency coherence in Fig.II.7 (A, B)). No such incursions of cardio-respiratory rhythms are visible in

the second EEG. Finally, both recordings are high pass filtered, at a lower frequency for the first EEG (near 0.1 Hz) than for the second (near 0.3 Hz, see amplitude spectra in panels (C)). Below this cut-off, only intense events such as motion artefacts are still visible.

Compared to the EEG spectrogram (i.e. the Gabor transform) used in [Olbrich, Claussen, and Achermann 2011; Prerau et al. 2017], the most striking difference of the EEG scalogram (i.e. the wavelet transform) is its logarithmic frequency axis (and measure $\mathrm{d}\log f$), consistent with the constant quality factor (and relative frequency resolution) of the wavelet transform, instead of a constant frequency resolution and a linear frequency axis (and measure $\mathrm{d}f$). The log-frequency axis emphasizes multi-scale aspects and details much more the low frequencies than the linear axis. This reveals that the δ band contains occasionally, in some recordings, the cardio-respiratory rhythms. Moreover, the use of the log-frequency measure is suited to represent the approximately constant EEG power per decade (see the root-mean-square amplitude in panels C): the trend of the EEG spectrum is much closer to the one of the pink noise than the one of the white noise, see discussion in section 3.3.2 of chapter II. Hence, the readability (the contrast) of the representation is improved at high and low frequencies with the scalogram. Eventually, the use of a constant quality factor means that the time resolution of an oscillating event is adjusted to its frequency, i.e. to the duration of a minimum number of oscillations $f\delta t = \frac{Q}{\sqrt{\pi}}$ (the one of the wavelet). These arguments are clearly in favour of the wavelet paradigm for analysing the EEG.

Contrary to EEG oscillations in the low frequency bands (δ and θ) or the high ones (β, γ), which are more time-localized (vertical lines) or irregular and wide-band, the intermediary bands α and σ (between 8 and 16 Hz) contain quite frequency-localized structures, i.e. of many regular oscillations. These narrow-band activity are illustrated as intermittent horizontal lines in Fig.V.1, σ waves (or sleep spindles) in the non-REM phase and α waves in the wake phase. The waves in these intermediary bands benefit from the spectrogram representation, that typically yields a more precise frequency resolution than the scalogram at high frequencies.

We argue that a well-chosen quality factor permits to distinguish these different types of event conveniently. The distinction of the central frequencies of the α and σ bands, say 10 and 13 Hz, requires for the log-frequency resolution $\delta\log f = \frac{\sqrt{\pi}}{Q} < \log\frac{13}{10}$, i.e. a quality factor $Q > 9.5$. Moreover, the optimal time localization of a micro-structure such as sleep spindles, that seems to last more than 5 oscillations ($f\delta t = \frac{Q}{\sqrt{\pi}}$), requires a quality factor $Q < 12.5$. This has led us to choose $Q = 10$ for convenient EEG scalograms in Figs.V.3 and V.4. This is a compromise, however, since singular oscillations, like motion artefacts or K-complexes, are not optimally localized in time, while the α waves could be much better localized in frequency.

Noticing that α waves are not concomitant with sleep spindles of the σ band, and that these bands need not be separated; we call α-σ their union, from 8 to 16 Hz. This means that the wavelet bandwidth can be as wide as $\delta\log f < \log\frac{16}{8}$: a lower quality factor can be chosen ($Q > 3.6$), improving a lot the time localization. We thus also use $Q = 5$ to study local EEG structures, such as modulations, with a higher time resolution (see Fig.V.6 (C)), even though the frequency resolution looks degraded.

In fact, the spectral information seemingly missing from the amplitude is contained in the phase, see Fig.V.5 (B) for α waves. Spectral lines associated to regular neural rhythms are retrieved precisely from their instantaneous frequency, as visualized with frequency-reassignement (C). Even at $Q = 5$, the power density (D) estimated with frequency-reassignment (thick black line) has the same spectral width as from the raw Fourier transform (thin grey line), directly caused by the temporal variability of the instantaneous frequency; the bandwidth associated to the wavelet (thick red line) has been squeezed.

In summary, the wavelet representation for the EEG is not only suitable to represent the global structure of sleep, but also to its micro-structures, at a temporal resolution that would be unreachable otherwise: a quarter of a second for the α waves with $Q = 5$.

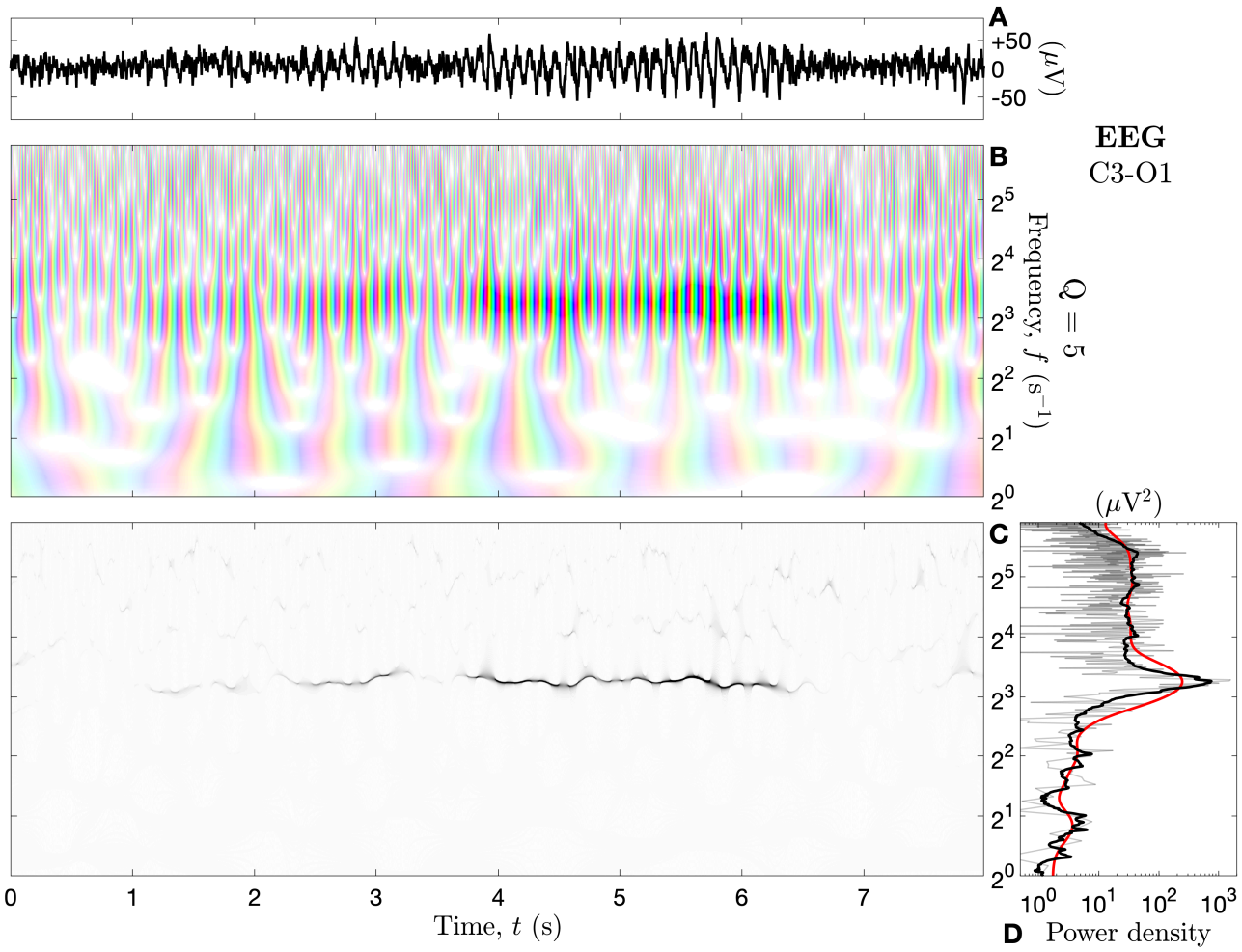


Figure V.5: *EEG with an α wave when subject 3 wakes up (with eyes closed). (A) Signal (mV). (B) Wavelet transform ($Q = 5$) with the hue-saturation colour coding of the phase and amplitude (see Fig.II.1). (C) Frequency-reassigned power density Eq.*(III.17)*. (D) Power log-frequency density Eq.*(II.23) *(in mV^2), estimated from the wavelet transform without and with frequency-reassignment (thick red and black lines), or directly form the Fourier transform (thin grey line).*

1.2.2 EEG modulations as band power, rate or magnitude signals

Like the cardiac and respiratory rhythms, the neural activity is modulated, and we may want to estimate these modulations for further analysis of polysomnography. This suggests using the same complex rate estimator F_x as previously for the EEG recording x. The presence of many neural "rhythms" in different bands, however, is an important difference between an EEG and ECG or AF recordings, that mostly contain a single one. We present below how the common approach of estimating the power variations in each EEG bands can be reproduced in the wavelet framework, and how it is related to F_x.

In Fig.V.6, we select the time interval for which the subject of Fig.V.3 falls asleep (around 20 min) and wakes up (around 50 min). These transitions are well visible in Fig.V.6 (C) from the changes of amplitude at high and low frequency at these times, as well as in the average instantaneous frequency $\Re\{F_x(t;Q)\}$ (E), that summarizes this behaviour in a surprisingly close way to the hypnogram (A). Notice that we can access much more information from the wavelet transform (C) than from the hypnogram, such as micro-states of arousal during sleep at 40 and 46 min, yielding transient high amplitudes at the high frequencies. A conventional way to deal with the complexity of an EEG is to divide it into band-limited signals, computed

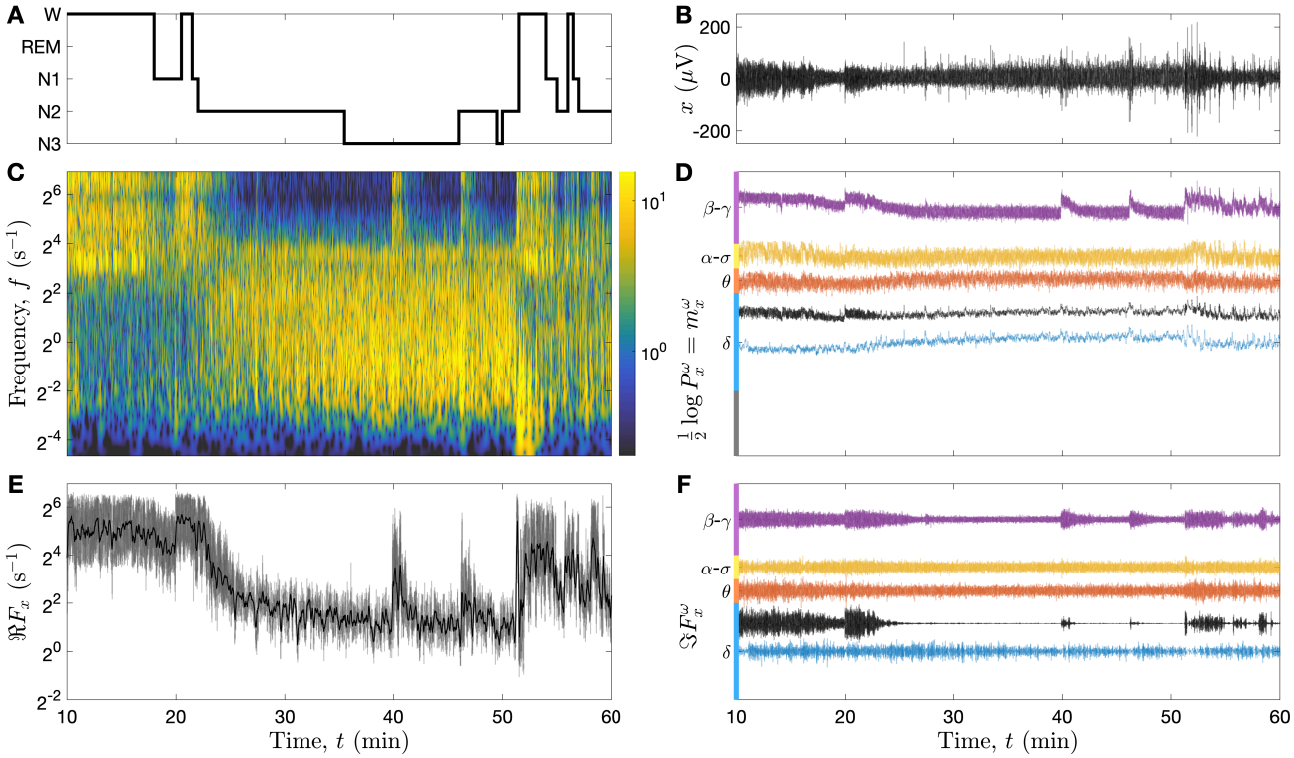


Figure V.6: *Extraction of modulations in the EEG. (A) Hypnogram: subject 3 falls asleep around 20 min and wakes up around 50 min. (B) EEG signal x (C3-O1). (C) Amplitude from the wavelet transform (twice the modulus) with quality factor $Q = 5$, colour scale in microvolt. (D) Natural logarithm of the EEG band magnitudes, Eq.*(V.2)*: m_x^δ in blue, m_x^θ in red, $m_x^{\alpha\text{-}\sigma}$ in orange, $m_x^{\beta\text{-}\gamma}$ in purple. The black line is the magnitude m_x in the full frequency range (from 0.04 to 125 Hz). (E) Average instantaneous frequency $\Re\{F_x\}$ on the full frequency range (grey line, and its 8 s moving average, black line). (F) Magnitude rate $\Im\{F_x^\omega\}$ computed in the same frequency bands as in (D). In (D) and (F), the signals have an arbitrary scale, and a mean value aligned to the position in (C) and (E) of the central band frequency.*

straightforwardly from its wavelet transform as a band power P_x^ω, or as a band magnitude signals m_x^ω:

$$P_x^\omega(t) = \int_{B_\omega} |X(t, f; Q)|^2 \mathrm{d}\log f \, \frac{Q}{\sqrt{\pi}} \tag{V.1}$$

$$m_x^\omega(t) = \tfrac{1}{2}\log P_x^\omega(t) \quad . \tag{V.2}$$

Note that the width of these bands $B_\omega = [f_\omega^-,\ f_\omega^+]$ is effectively limited by the wavelet bandwidth $\delta \log f$. The conventional bands $\omega = \delta, \theta, \alpha, \sigma, \beta, \gamma$ are placed rather evenly on the linear frequency scale (with a width of about 4 Hz), not at all on a logarithmic one. This choice was probably contingent on the linearity of the frequency axis; we take the liberty to slightly adapt the bands as follows: δ from 0.25 to 4 Hz, θ from 4 to 8 Hz, α-σ from 8 to 16 Hz (as discussed previously) and β-γ above 16 Hz (up to the Nyquist frequency limit), see Fig.V.6 (D).

Eq.(IV.11) relates precisely band power or magnitude to a band-limited version of the complex rate estimator:

$$\Im\{F_x^\omega(t; Q)\} = -\partial_t m_x^\omega(t)/\Upsilon \quad . \tag{V.3}$$

The instantaneous frequency (real part of the complex rate) in most band is not used in the following and hence omitted in Fig.V.6. It serves nonetheless to distinguish α waves from sleep spindles in our custom α-σ band. The band signals is Fig.V.6 (D) and (F) contain the same

information on the EEG modulations: they essentially differ by a phase shift in a coherence analysis. Therefore, both kinds of signals (magnitude and its rate) can be used to study the modulation of the amplitude in the EEG bands. We prefer the magnitude, that is more straightforward to interpret.

An average is required to define magnitude signals without divergences (otherwise occurring at phase vortices of the wavelet transform). When multiple EEG channels are available, their average may be used at the cost of losing the spatial information. Here, the variability of the average estimator from the band integration depends on its relative bandwidth compared to the one of the wavelet; the α-σ band signal is more variable than the δ one. However, a small averaging is sufficient to avoid divergences. A continuous and homogeneous approach consists in using the time-dependent log-frequency power density estimator $S_{xx}(t, f_\omega; Q_+, Q_-)|f_\omega|$, locally-averaged version of the scalogram $|X(t, f; Q)|^2$, from which we can define the magnitude field as:

$$m_x(t, f_\omega; Q_+, Q_-) = \tfrac{1}{2} \log \left(S_{xx}(t, f_\omega; Q_+, Q_-)|f_\omega| \right) \quad , \tag{V.4}$$

with a continuous band frequency parameter f_ω, a bandwidth controlled by Q_- and a time resolution by $Q_+ > Q_-$. A practical estimator for this quantity is obtained from a time smoothing over n wavelet durations of the scalogram: $S_{xx}(t, f; nQ, Q)|f|$ ($f = f_\omega$, $Q = Q_-$). At this point, we reach an agreement with the proposal in [Prerau et al. 2017] to use a proper statistical estimation of the EEG power density in the time-frequency plane. For more detail on the correspondence between their multi-taper spectrogram and our wavelet estimator, we refer to chapter II, section 5.

Before comparing the neural magnitude signals to the cardio-respiratory ones, we explore in the next section the possibility to compare two EEG recordings by extending this quadratic approach to cross power density and coherence.

1.3 Wake-sleep patterns in the magnitude and coherence of an EEG pair

In this section, we illustrate the time-frequency coherence between two channels of an EEG recording, and we recall the steps to achieve its computation and graphical representation. For an illustrative purpose, we selected a pair of clean EEG signals, recorded on both sides of the head of subject 6: x is measured between points C3-A2 and y is measured between points C4-A1 (left-right symmetric to C3-A2). One would expect that these recordings from contra-lateral locations would produce a very strong coherence, with little temporal evolution.

Their global coherence (complex correlation coefficient), estimated over the full overnight record (9 hours), is as low as $\rho_{x^+y^+} = 0.061 + i0.029$, the real part of which is the usual Pearson correlation coefficient $\rho_{xy} = 0.061$. In regard to the non-stationary and multi-scale characteristics of EEG, this absence of global linear similitude can be explained. Actually coherent time-frequency subdomains could occur, with different phases that interfere and globally cancel out. Therefore, we propose to localize the correlation analysis in different frequency bands and at different times by applying the time-frequency coherence formalism.

The succession of computations performed on these signals leading to γ_{xy} is represented in Fig.V.7. We show in Fig.V.7(B, C) the signal $x(t)$ and the amplitude of the wavelet transform $2|X(t, f; Q)|$; the second EEG has a very similar aspect, hence not represented. The chosen quality factor $Q = 10$ is sufficient to identify the EEG bands (simultaneous modes of frequency ratio $\exp(\delta \log f) \approx 1.3$ can be distinguished). The wavelet transforms need to be paired, multiplied and smoothed for the preliminary estimation of the power spectral densities. To keep the full frequency resolution offered by the wavelet analysis, we smooth in time only a Gaussian kernel of width $n\delta t$ that lasts $n = 50$ wavelet durations, see Eqs.(II.70, II.71), to estimate the

power log-frequency densities $S_{xy}(t, f, nQ, Q)|f|$. Their geometric mean, $\sqrt{S_{xx}S_{yy}}\,|f|$ in panel (D), gives a clear illustration of sleep stage patterns, summarized in the hypnogram (A) from the original annotations. The use of a large time-smoothing regime sets a low level of the spurious coherence ($\gamma_{\rm sp} \sim 0.14$), allowing to discriminate strong and weak correlations. It has the side effect of a poor time resolution, especially below 0.5 Hz (a quarter of an hour at $f = 0.2$ Hz), where scattered coherent spots of varying phase are most likely due to isolated intense and coherent events such as motion artefacts.

Since the cross power density $S_{xy}|f|$ is a complex-valued map, we employ in Fig.V.7 (E) the amplitude-phase colour coding. Its phase (i.e. the one of the coherence) is an interesting piece of information which we do not discard. It represents the phase difference between the signals: the zero phase shift is coded in green, the phase opposition ($\pm\pi$) is coded in magenta, and the phase quadrature is coded in orange ($-\frac{\pi}{2}$) or light blue ($+\frac{\pi}{2}$), as indicated in the colour bar. The strength of local correlations are finally obtained in Fig.V.7 (F) from the time-frequency coherence estimator $\gamma_{xy} = \gamma_{xy}(t, f; nQ, Q)$, ratio of (E) and (D). The regions of high in-phase coherence (green) during the REM stage (of low power), illustrates well the insensitivity of γ_{xy} to the power density. The phase of the cross-spectrum remains, but its modulus is normalized so that a loss of coherence is only due to destructive phase interference. For a better readability, the saturation of colours is discretized into 5 ranges of coherence modulus $|\gamma_{xy}|$. The low coherence values that cannot be distinguished from the spurious coherence (p-value $> 10^{-1}$) are in white (no saturation). The range represented with the faintest colour saturation is made of low but significant coherence values, with p-values $10^{-3} < p < 10^{-1}$: $|\gamma_{xy}| \in [\gamma(10^{-1}), \gamma(10^{-3})] \approx [0.21, 0.36]$. The high coherence ranges are delimited by the coherence values $|\gamma| = \gamma(10^{-3})$, 0.5, 0.7 and 1, with increasingly saturated colours.

The expected spurious coherence level is illustrated in Fig.V.7 (G), computed between incoherent signals, here the EEG and its phase-randomized surrogate signal [Lancaster et al. 2018]. Thus, the estimation of the significance for low coherence values can be controlled and visualized: the surrogate coherence only exhibits scattered spots of significance $10^{-3} < p < 10^{-1}$ (consistent with a density of false positives of about 10% in the time-frequency plane) with a random phase. The size of these spots is representative of the resolution of the coherence analysis in the time-frequency plane: their area is of the order of n time-frequency atoms. Its comparison to γ_{xy} of Fig.V.7 (F) confirms that both EEG activities are strongly correlated, with various phase shifts, in many time-frequency regions.

Fig.V.8 provides in (A) and (B) a zoom of the EEG power density and inter-EEG coherence of Fig.V.7 (D, F). The selected time-frequency domain contains the three main EEG patterns, structured both in scales (frequency bands) and in time stages: as indicated in the hypnogram, it starts with the end of NREM, then REM, wake (W) and back to another NREM state. In addition to the commonly scrutinized power density in each frequency band, provided in (A) as a continuum, a particularly rich information is found in the coherence and its phase.

The phase-frequency relation in all regions of significant coherence $\gamma_{xy} = |\gamma_{xy}|e^{i\phi_{xy}}$ is well described by $\phi_{xy} = \phi_{\pm} + \Upsilon f\tau$, responsible for the vertical rainbows (phase gradients) at high frequencies in Figs.V.7 (E, F) and Fig.V.8 (B). Remarkably, the first term of this linear relation can only take two angular values, $\phi_+ = 0$ (in-phase, green) and $\phi_- = \pm\pi$ (phase opposition, magenta), and $\tau \approx 10$ ms is a constant delay. Therefore, the left-right symmetry of the correlated EEG activity (that would write $\phi_{yx} = \phi_{xy}$) is only spoiled by a short delay: x is 10ms early compared to y. This delay only corresponds to the third highest local maximum of the global correlation function $\rho_{xy}(\tau)$ (that does not increase much: $\rho_{x^+y^+}(\tau) = 0.085 + i0.019$), confirming the irrelevance of stationary methods. Of the order of one time-step (sampled at 125 Hz), we suspect an instrumental origin for this delay.

The map $\gamma_{xy}(t, f; nQ, Q)e^{-i\Upsilon f\tau}$ in Fig.V.8 (C) provides a visual support to the detailed description of these patterns in Table V.1. It consists in compensating the linear trend of the

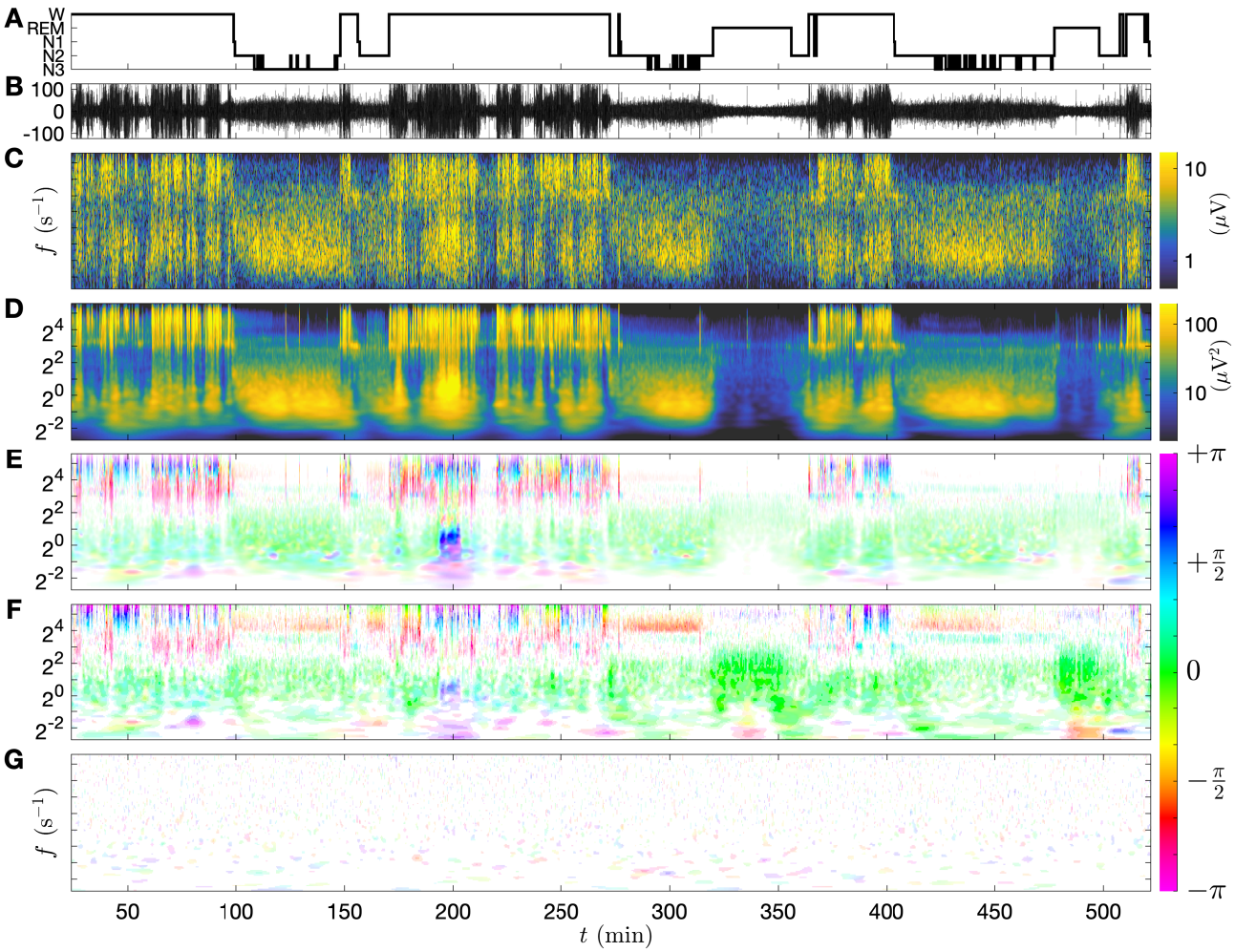


Figure V.7: *Time-frequency coherence analysis of two EEG recordings of subject 6: x corresponding to C3-A2 and y to C4-A1. (A) Hypnogram. (B) Signal y (in μV). (C) Amplitude $2|Y(t,f;Q)|$. (D) Geometric mean of the (time-varying) log-frequency power densities $\sqrt{S_{xx}S_{yy}}\,|f|$. (E) Cross power log-frequency density $S_{xy}(t,f;nQ,Q)|f|$. (F) Coherence $\gamma_{xy}(t,f;nQ,Q)$ (ratio of (E) and (D)). (G) Coherence with a surrogate EEG, to illustrate the spurious coherence level in (F). The ranges of coherence moduli $|\gamma_{xy}|$ for the colour saturation coding are delimited by the lower thresholds $\gamma(10^{-1}) \approx 0.21$, $\gamma(10^{-3}) \approx 0.36,\ 0.5,\ 0.7$. The quality factor $Q = 10$ is used for the computations, together with the Gaussian smoothing window χ_n of temporal width $n = 50$ units of wavelet duration.*

phase-frequency relation, caused by the global delay between the EEG recordings. For such a small delay, we notice that $\gamma_{xy}(t,f;nQ,Q)e^{-i\Upsilon f\tau}$ is a very good approximation of the generalized delayed time-frequency coherence $\gamma_{xy}(f,t-\tau,t)$, between $x(t-\tau)$ and $y(t)$.

In the light of this specific phase relation, the description of the coherence patterns boils down to: (i) the sign of the correlation, (+) for in-phase regions ($\phi_+ = 0$) and (−) for the ones in phase opposition ($\phi_- = \pm\pi$), and (ii) its strength (insignificant to high) everywhere is the time-frequency plane. The time-frequency map of $\phi_\pm$ is obtained by compensating the delay: $\gamma_{xy}(t,f;nQ,Q)e^{-i\Upsilon f\tau}$, illustrated in Fig.V.8 (C). A synthetic description of the inter-EEG coherence in distinct time-frequency regions (EEG bands and sleep stages) for subject 6 is constructed in Table V.1. This confirms the strong capability of the time-frequency coherence method to bring to sleep study otherwise unnoticed inter-EEG correlations.

We recall that this particular phase relation is specific to the compared pair of bipolar EEG (C3-A2 versus C4-A1). In the context of a unipolar multi-channel EEG, the possible pairing

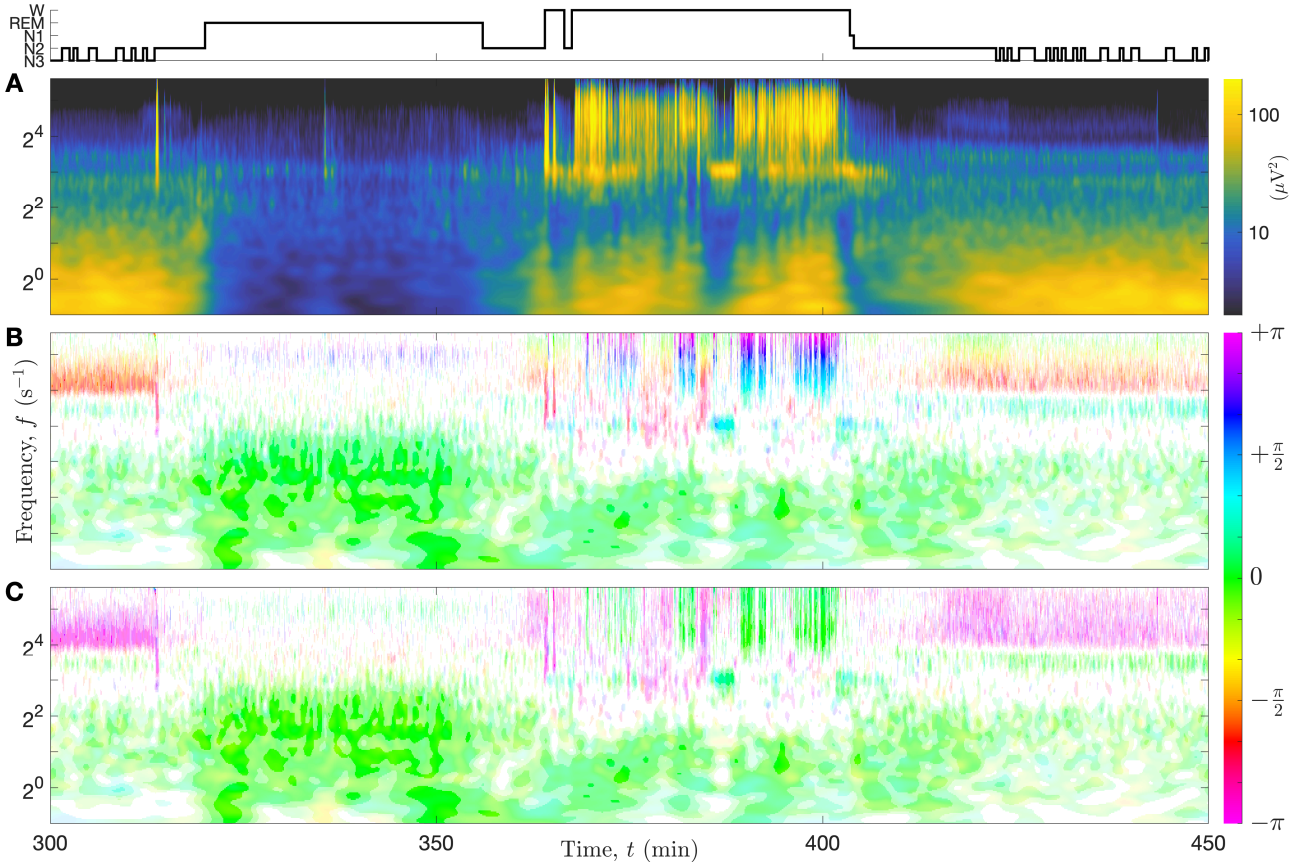


Figure V.8: *Details of the three main sleep-wake patterns in the EEG power density and coherence from Fig. V.7 (D, F). (A) Power density* $\sqrt{S_{xx}S_{yy}}\,|f|$ *(geometric mean). (B) Coherence* γ_{xy}*. (C) Coherence with delay compensation,* $\gamma_{xy}(t,f;nQ,Q)e^{-i\Upsilon f\tau}$*, where* $\tau \approx 10$ *ms.*

band (Hz)	W	NREM	REM
γ 60	$++^P$	$-_p$	$+_p$
β 20	$--^P$	$--_p$	0_p
σ 16	*bursts* 0	$+$ $++$	
α 12	$++^P$ $-$ 0	0_p	*liminal*
θ 8	0 *liminal*	0	
δ^+ 4	$+^P$ 0_p	$+$	$++_p$
δ^- 1 1/4	*bursts*	0^P $+^P$	

Table V.1: *Summary of coherence patterns observed for subject 6, organized by band (row) and stage (column). Positive and negative coherences are denoted by* $\pm$*, doubled when strong,* 0 *when incoherent, and* $^P/_p$ *for especially high/low power density. When heterogeneous, we qualify the fine structure: several symbols for alternating or intermittent, "bursts" for singular events, "liminal" or interstitial when influenced by neighbouring bands.*

of channels are numerous. S_{xy} is only one entry of the corresponding power density matrix $\mathbf{S}$, indexed by the electrode locations. The introduction of these two spatial variables increases the dimensionality of the EEG description, and the difficulty of its practical interpretation. The exhaustive computation of the matrix $\mathbf{S}$ for an 11-channel EEG yields 55 independent entries, each being a time-frequency image. These images and their sleep-patterns are largely redundant, with an important diversity of phase values depending on the pair of electrodes.

Furthermore, a single-channel EEG is already a difference of potentials at distinct spatial locations; this fact appears as a more fundamental issue, this is the reference problem. For instance, the re-referencing of two channels x and y into $x' = x - z$ and $y' = y - z$ (with the common reference signal z) changes the power density from S_{xy} to $S_{x'y'} = S_{xy} - S_{xz} - \overline{S_{yz}} + S_{zz}$, as well as its associated coherence. We could assume that the difference introduced between

candidate references is negligible ($S_{x'y'} \approx S_{xy}$). Preliminary observations suggest that this assumption does not hold: referencing either at Cz or with the spatial average changes entirely the phase value for most pairs of locations. Improving the spatial localization of unipolar EEG, the Laplacian method might help to circumvent this reference problem that clutters the interpretation of the inter-EEG phase.

Summary of section 1

Essential component of a polysomnography, EEG recordings of the neural activity have been introduced, and their main instrumental and time-frequency characteristics have been discussed and illustrated. A constant quality factor representation is especially well suited to observe their multi-scale architecture along time, globally described in terms of wake-sleep temporal stages, and spectral bands spanning over 3 decades. The wavelet transform $X(t, f; Q)$ with quality factor $Q = 5$ is proposed to study EEG phenomena in the finest details, allowing a very high time resolution of the transient modulations of the neural activity, while its phase can be used to retrieve precisely the instantaneous frequency of more regular oscillations. If using an amplitude representation only, for instance the power density $S_{xx}(t, f; Q_+, Q_-)$ without frequency reassignment, $Q_- = 10$ is advised for a better distinction in the narrow intermediary band (θ, α, σ). In correspondence with multi-taper spectrogram representation, the EEG reading from the power density is particularly clarified using a time-smoothing (from $Q_+ > Q_-$), that can be as important as $Q_+ = 500$. Beyond the power density, sleep patterns can also be observed from the inter-EEG time-frequency coherence $\gamma_{xy}(t, f; Q_+, Q_-)$, illustrated for individual contra-lateral recordings. Phase patterns with alternating signs and significance (incoherent to very coherent), varying independently from the power density, bring novel insights into the neural activity in different stages and bands. The magnitude of neural activity can be extracted either from frequency integration of the EEG intensity in each discrete band of interest, or continuously across scales from the amplitude or power density with small quality factors. These neural modulations are compared to cardio-respiratory modulations in the next section.

2 Sleep apnea: a slow and ubiquitous rhythm

In this section, we focus on the impact of sleep apnea on neural (EEG), cardiac (ECG) and respiratory (AF) rhythms recorded during sleep. Their altered dynamics is studied from the extracted physiological modulation estimators, and from their coherence. The significance of our coherence estimator is discussed in relation to the chosen quality factors. This original method, based on a two-step time-frequency decomposition, can be used to capture the rhythm modulations of any physiological signal and requires no signal-specific adjustment, other than the possibility to restrict the spectral range. Applied on the polysomnography of the subject 3, affected by obstructive sleep apnea, this method shows how repeated apnea events during the NREM sleep stage N2 are associated to very coherent slow modulations (∼0.035 Hz) across all possible pairs of physiological signals.

The robustness of this observations is assessed by performing a statistical survey of the large *shhs2* database. Based on the clinician annotations of apneic events, we reconstruct typical HRV power density and coherence spectra for subpopulations of patients with sleep apnea disorders. These spectra not only confirm the statistical validity of the first observation on the selected subject, but also draw our attention to other key elements: (i) the coherence spectra around the slow mode of apnea modulation, in the band [0.01, 0.08] Hz, is made of a fundamental and a harmonic mode, restricting the apneic fundamental frequency to the range

[0.01, 0.04] Hz; (ii) the fast coherent cardio-respiratory modulation (1/4 Hz), associated to respiratory sinus arrhythmia (RSA), almost vanishes for more severe (obstructive or central) sleep apnea; (iii) the very slow coherent mode is also present in neural modulations, in phase with the heart rate and in phase opposition with the breathing rate.

These statistical results are completed by the distribution of the phase shift of slow neural modulations, continuously across EEG band frequencies, compared to a cardiac or respiratory reference modulation. We put into evidence the existence of slow and very slow EEG activity, coherent with cardio-respiratory modulations, both in presence and absence of apnea. These neuro-respiratory and neuro-cardiac interactions are characterized in the continuum of EEG band frequencies, and exhibits a typical phase pattern. We interpret it as the observation of the multi-band phase relation between the cardio-respiratory activity and cyclic alternating patterns [Terzano et al. 2001] that corresponds to slow recurrent patterns of cortical activity.

Eventually, we illustrate preliminary investigations on the detection of sleep spindles from the phase information in the σ band, and the coupling of their amplitude to the phase of slow oscillations during deep sleep (N3 stage) [Purcell et al. 2017]. We motivate the possibility to measure phase-amplitude couplings as the coherence between EEG amplitude signals and the EEG recording itself. We finish on a complete picture of wake-sleep and apnea stages in this individual polysomnography by including the previously computed pulse-respiration quotient.

2.1 Neuro-cardio-respiratory coherence of an apneic subject

The modulations of physiological rhythms are extracted from the EEG, ECG and AF recordings from subject 3 in order to explore their correlations below the frequency of the instrumental high-pass filter. The cardiac and respiratory modulations are estimated with the complex rate signal F_x, and the neural modulations with the previously introduced magnitude signals m_x^ω in the bands ω.

For the sake of simplicity, we use the most generic wide-band estimator $F_x(t;Q)$ for the cardiac and respiratory recordings. The coherence analysis applied to them in the chapter IV suggests that further insights from more sophisticated estimators are limited; we recall arguments for this simplification.

Although these physiological fluctuation rate signals do not compare directly to the idealized cardiac and respiration rates, their spectral richness capture all the modulations that are resolved by the wavelet in the considered frequency range. There are two main quantitative differences, both related to the non-circularity of the rhythmic oscillations. The first one is the echo of the carrier wave oscillation in the estimated rate signal, strong for the ECG, reduced for the aECG14 signal and weaker for the airflow recording. Instead of perturbing the analysis of modulations, the carrier frequency enriches it beyond the spectral range for the frequency and amplitude modulations (FM and AM).

The second difference is a dilation factor between the ideal and the estimated rates, that is caused by the contributions of harmonic components to the waveform. Weak for the airflow rate signal (see Fig.IV.19), this effect is strong for the ECG rate signal. In the case of an invariant waveform, the analysis of the rate modulations is not affected since the resulting factor is constant. To the contrary, modulations of the waveform can be confused with FM by the wide-band estimator $F_x(t;Q)$, as for the voice vibrato in Fig.III.5. For this reason, we can only consider it a *proxy* for the heart rate, sufficient for a coherence analysis because the cardiac waveform does not change dramatically during the recordings, at least at the scale of few minutes. We verified it in Fig.IV.11, where its coherence with the reference heart rate signal is significant, although decreased compared to more sophisticated estimators.

2.1.1 Physiological modulations in time and frequency

The modulation estimators are computed for the full overnight records of subject 3. In particular, we discuss the modulations of the cardiac frequency, respiratory frequency, and EEG magnitude in the δ and β-γ bands. The contributions from multiple scales, superimposed in these modulation estimators, are revealed by their wavelet transform, which is the preliminary step to the time-frequency coherence analysis.

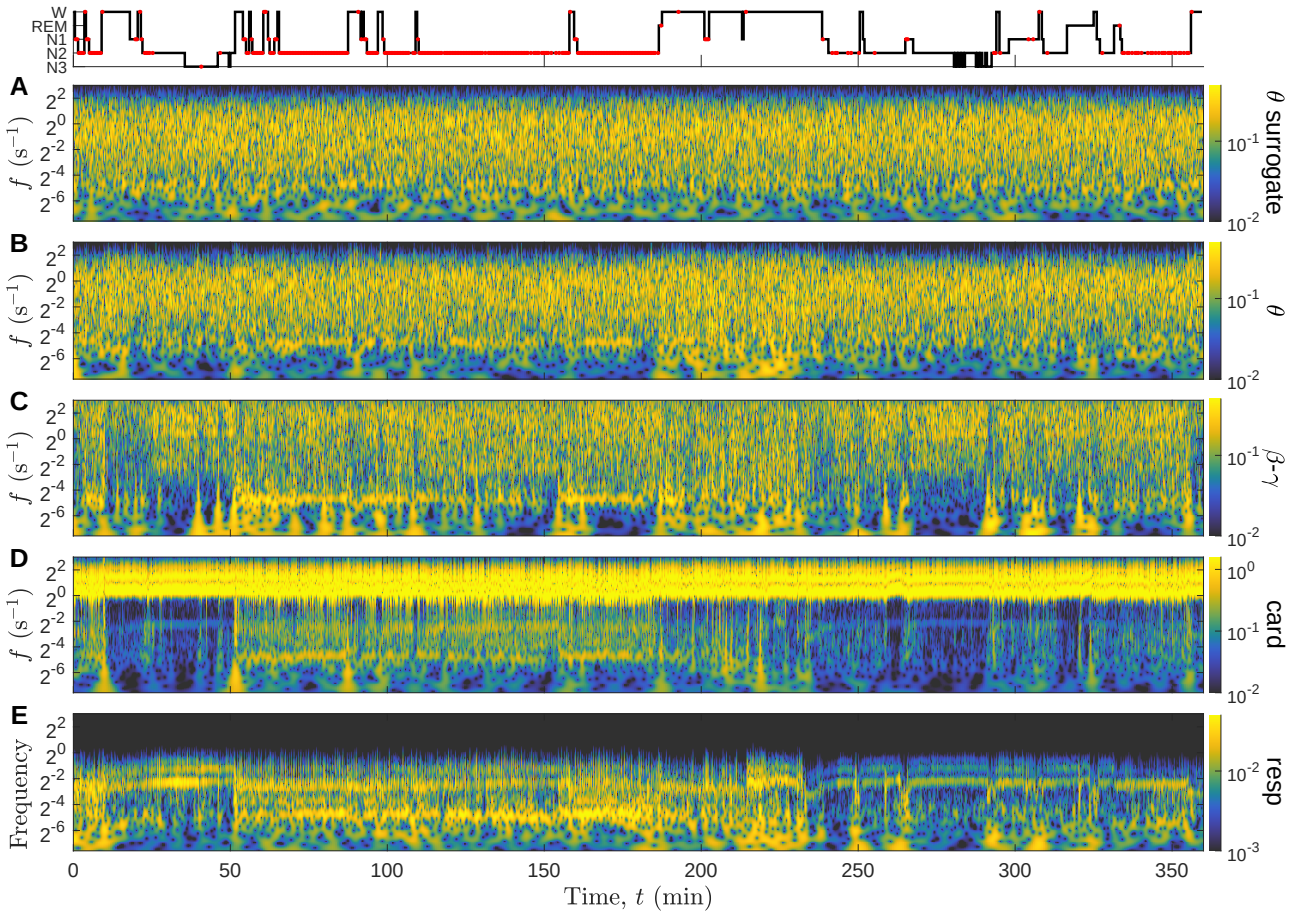


Figure V.9: *Physiological modulations of subject 3. (A) Phase-randomized surrogate of (B) m^{θ}_{EEG}, the magnitude in the θ band, (C) magnitude in the β-γ band $m^{\beta\text{-}\gamma}_{EEG}$, (D) cardiac frequency modulation $\Re\{F_{ECG}\}$, (E) respiratory frequency modulation $\Re\{F_{AF}\}$. The colour codes for the amplitude (twice the modulus) of the wavelet transform, with the quality factor $Q = 5$. The amplitude has the unit of the estimator: no dimension in (A, B, C), in radian per second in (D, E). At the top row, the hypnogram is marked with red dots corresponding to annotated events of obstructive apnea with arousal.*

The wavelet transform is performed on two distinct levels to obtain the time-frequency representations shown in Fig.V.9: a first transform of each recording is required to compute the modulation signals, and a second transform is applied on these new signals. Even though the choices of the parameters could be distinct in these two rounds of wavelet transform, we use for both the quality factor $Q = 5$, appearing as a good compromise between a precise time localization and a sufficient frequency resolution. This fixes the wavelet widths to $f\delta t = \frac{Q}{\sqrt{\pi}} \approx 2$ oscillations and $\delta \log f = \frac{\sqrt{\pi}}{Q} \approx \log 1.65$ (1.65 being the least distinguishable frequency ratio). In addition to these modulations, represented in Fig.V.9 (B - E), we construct a control signal, the phase-randomized surrogate of the neural modulation in the θ band ($m^{\theta}_{\mathrm{EEG}}$) shown in panel (A) for comparison. This is a stationarized version of the original modulation signals obtained

by shuffling uniformly the phase of its Fourier coefficient, operation that leaves the spectral density unchanged.

At this stage of the analysis, $m^{\theta}_{\mathrm{EEG}}$ is hardly distinguishable from its surrogate signal, Fig.V.9 (B) and (A) respectively; they both exhibit a quite homogeneous distribution of the modulation's amplitude in the time-frequency plane. In the heart rate generic estimator, $\Re\{F_{\mathrm{ECG}}\}$, Fig.V.9 (D), the most intense oscillations are localized at the cardiac frequency (and its harmonics): this is the carrier frequency of the cardiac modulations in the strongly non-circular ECG signal. In spite of this dominant component, the information about the HRV is nonetheless present at lower frequencies: a mode of varying amplitude at the breathing frequency 0.2 Hz confirms the presence of RSA. Breathing modulations intensify, become unsteady and extend towards low modulation frequencies in the time interval between 50 and 180 min (stage N2). Apart from the respiration mode due to the non-circular carrier wave frequency, the breathing rate estimator $\Re\{F_{\mathrm{AF}}\}$, Fig.V.9 (E), exhibits in this time-frequency region an intense mode at about 0.035 Hz. The subject 3 is severely affected by sleep apnea, and this time interval corresponds to an uninterrupted sequence of such events; "obstructive apnea with arousal" are marked with red dots in the hypnogram. The presence of a clear mode at ∼0.035 Hz means that the corresponding apneic events occur with a quite regular period: approximately every 30 s. For this reason, we refer to this phenomenon as the "apneic rhythm".

In complement, we show in Fig.V.10 the time-frequency content of neural modulations in the EEG band magnitudes.

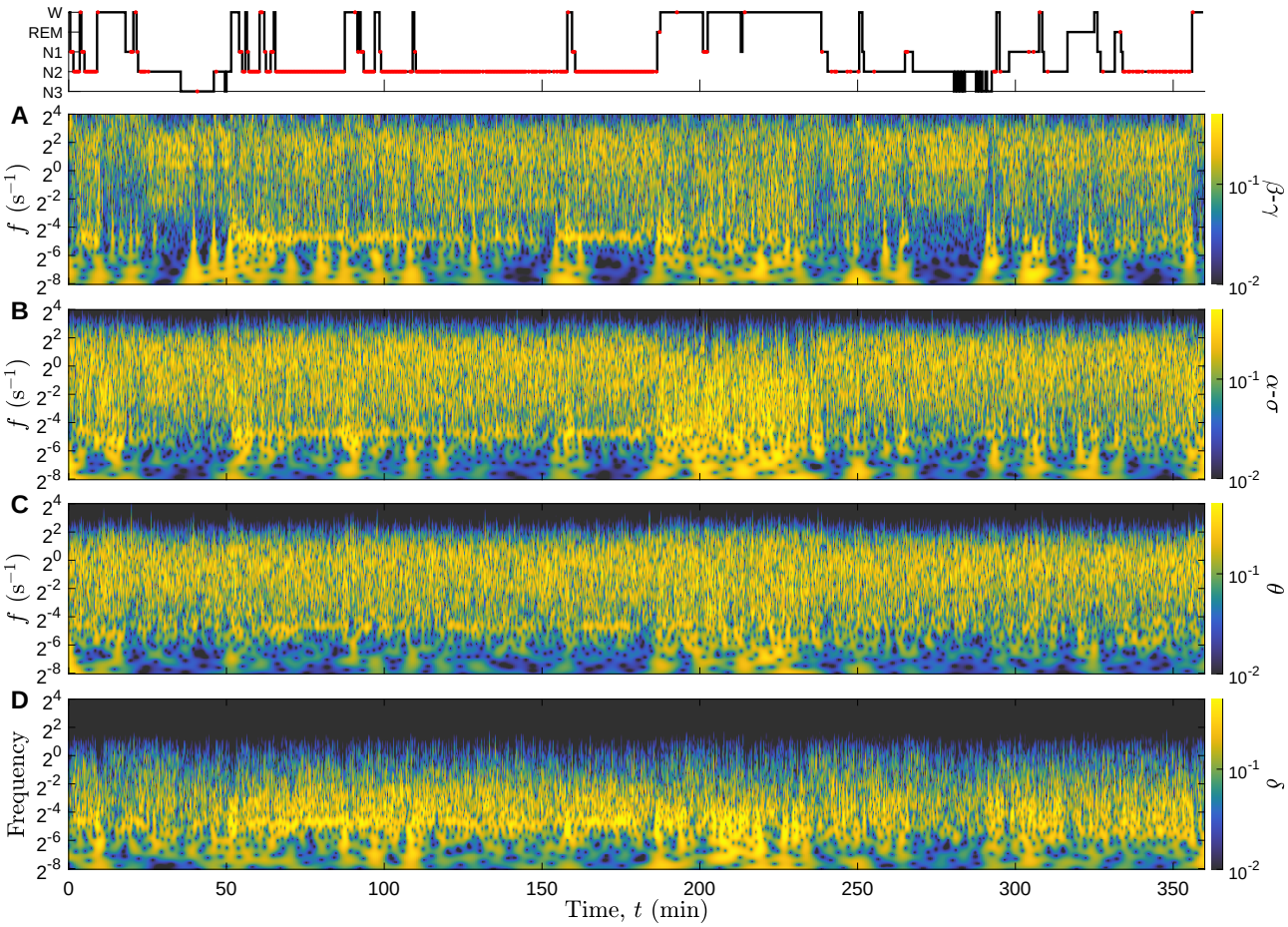


Figure V.10: *Neural modulations only, for subject 3 as in Fig.V.9. EEG magnitude m^{ω}_{EEG} in the band $\omega = \beta$-γ (A) 125 to 16 Hz, α-σ (B) 16 to 8 Hz, θ (C) 8 to 4 Hz, and δ (E) 4 to 1/4 Hz.*

The neural modulations vanish at high frequencies, beyond a cut-off that can be antici-

pated. Indeed, the time resolution is $\delta t \approx 2/f$ s for the wavelet transform with $Q = 5$, so that no modulation is expected above $f_\omega^+/2$ in each EEG band ω, as we can observe. The apneic rhythm is seen in all EEG bands, but appears weaker in the α-σ and θ bands than in the β-γ band, and strongest in the δ bands. Although strong δ waves (and associated modulations) are expected during the N2 stage, β-γ waves are much weaker, see Fig.V.3. This illustrates an important characteristic of the magnitude: related to the power by a logarithm, the extent of its modulation does not depend on its high or low mean value. It only needs to be more important than the background noise, that can be of physiological or instrumental origin.

So far, we can anticipate that this apneic rhythm causes correlations between the neural, cardiac and respiratory activities, since it is noticeable in all physiological modulations.

2.1.2 Neuro-cardio-respiratory coherence in time and frequency

Next, we estimate the power densities $S_{xy}(t, f; Q_+, Q_-)$ (cross and auto) from the wavelet transforms to obtain the time-frequency coherence. Two distinct quality factors are required for a statistically significant estimator: Q_- controls the frequency resolution and Q_+ the time resolution, for a number of statistical degrees of freedom $n_d = \frac{Q_+}{Q_-}$. We use the practical estimator based on the Gaussian time-smoothing kernel χ_n, Eq.(II.70), denoted $S_{xy}(t, f; nQ, Q)$, that approximates the canonical estimator $S_{xy}(t, f; Q_+, Q_-)$. The time-smoothing is over of $n = 10$ wavelet durations, decreasing the resolution to $10\delta t \approx 20/f$ s, sufficient to identify the breathing rate at a resolution of 1 or 2 min and to resolve the variability of the apneic rhythm. However, for only $n_d \approx n$ degrees of freedom, the spurious coherence at a 90% level of significance ($p < 10^{-1}$) associated to this quite local estimator is as high as $\gamma(10^{-1}) \approx 0.46$, see Eq.(II.78). This time-frequency coherence analysis is therefore limited to rather strong correlations. The resulting time-frequency coherence of different pairs of modulation signals for the subject 3 are represented in Fig.V.11.

The most striking observation in Fig.V.11 is a strong coherence in the frequency band near 0.035 Hz, between 50 and 180 min, in all pairs of physiological rate signals (C - H) (which can extend to 200 min, and also visible around 340 min). This coherent component is significant only intermittently in pairs involving the θ modulations (C, F, H). By comparing the time intervals in which this apneic rhythm appears with the annotations of the hypnogram (A), we notice that it only occurs during the N2 stage and that the coherence decreases or disappears when the person wakes up (W). The different colours of this region indicate different phase shifts between modulations. For instance, in panel (C) of Fig.V.11, the EEG β-γ band is $-\frac{2\pi}{3}$ to $-\frac{5\pi}{6}$ radians delayed (late) compared to the EEG θ band. This means that not only these two EEG frequency bands behave coherently, but also that they are quite in phase opposition; while the EEG signal in the θ band reaches its maximum, the EEG signal in the β-γ band increases progressively from its lowest value. In Fig.V.11 (D), the small phase shift between cardiac and respiratory modulations indicates that the decreases and increases of the heart and breathing rates occur quasi in-phase at each cycle of apnea (or the cardio-apneic rate variation slightly precedes the respiratory one). The light green colour of the apneic coherent region in the next panels (E-H) indicates that the cardiac and respiratory modulations evolve nearly in phase with the neural ones in the $\beta - \gamma$ band, while it is rather in phase opposition with the θ band EEG modulations (purple blue colour).

We can also observe in Fig.V.11 (G) a region of strong coherence ($|\gamma| \sim 0.8 - 0.9$) in phase opposition (magenta), from 250 to 340 min at very low frequencies (below $2^{-6} \sim 0.02$ Hz, i.e. at the scale of a few minutes). This region corresponds to isolated events of apnea (at times 250, 265, 292, 303, 306, 324 min), with relatively quick drops and restoration of the respiration frequency and simultaneous rise and disappearance of β-γ magnitude in the EEG. This can be checked in the ECG and AF time-frequency representation Fig.II.6 (A, C). These kinds of

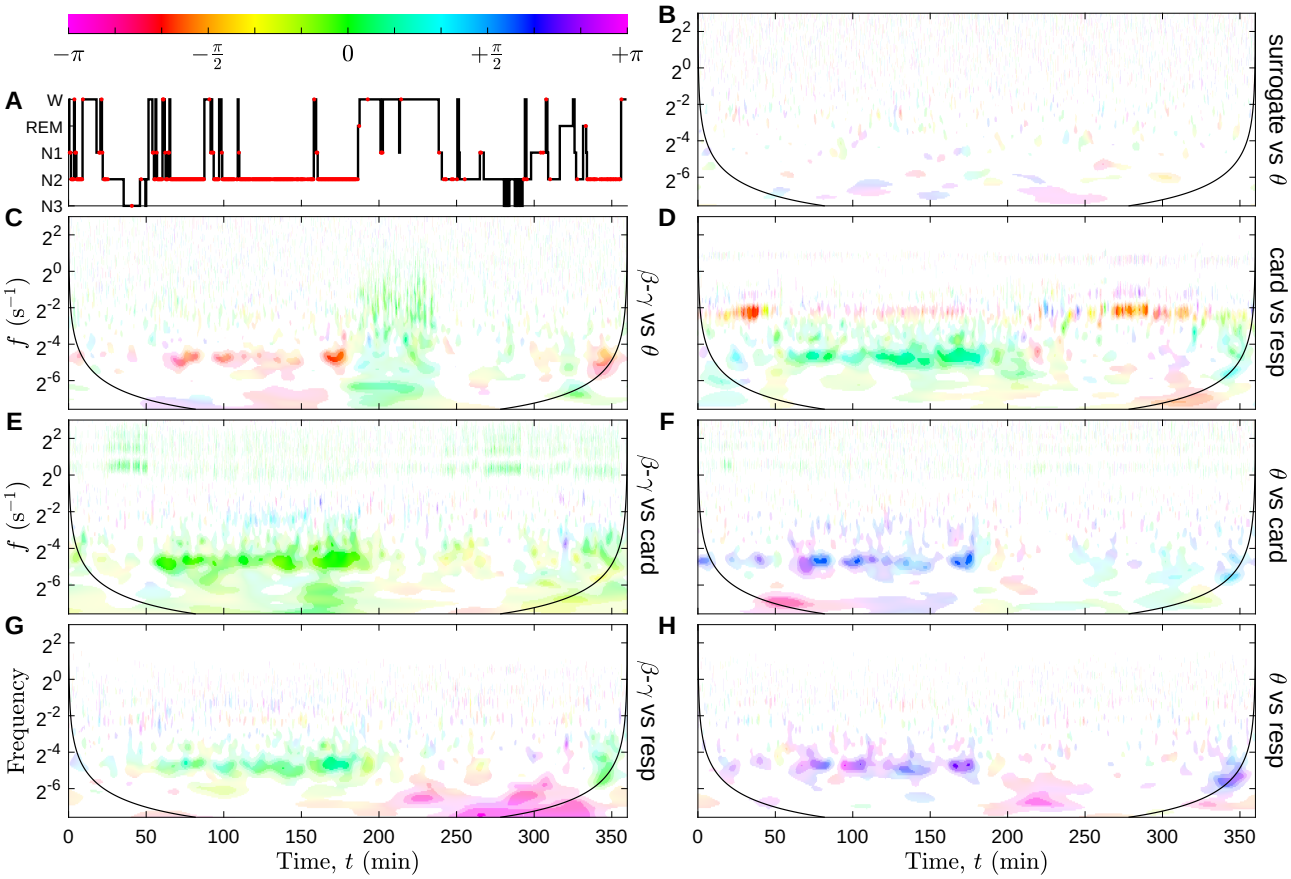


Figure V.11: *Time-frequency coherence $\gamma_{xy}(t, f; nQ, Q)$ between pairs of physiological modulations (x versus y, from Fig.V.9) for subject 3. (A) Hypnogram; the red dots corresponds to annotated events of obstructive apnea with arousal. (B) Band θ surrogate (phase-randomized signal) versus θ. This control coherence illustrates the level of significance of the spurious coherence. (C) β-γ versus θ band, (D) cardiac versus respiratory rate, (E) β-γ band versus cardiac rate, (F) θ band versus cardiac rate. (G) β-γ band versus respiratory rate. (H) θ band versus respiratory rate. The ranges of coherence moduli $|\gamma_{xy}|$ for the colour saturation coding are delimited by the lower thresholds $\gamma(10^{-1}) \approx 0.46$, $\gamma(10^{-3}) \approx 0.71,\ 0.8,\ 0.9$. The parameters are $Q = 5$ and $n = 10$. A black line materializes a distance $n\delta t$ from the initial and final times, beyond which border effects are possible.*

micro wake states may constitute a distinct recovery mechanism, slower than the apneic rhythm around 0.035 Hz.

Other regions of highly significant coherence can be observed. In Fig.V.11 (D), the modulation of the cardiac rate by the respiration (RSA, at the breathing frequency near 0.2 Hz) is also observed in some time intervals (from 30 to 40 min and from 270 to 320 min). During the sleep apnea episodes, the significance of the coherence is lower, even though the corresponding cardiac modulations are stronger, see Fig.V.11 (D).

In Fig.V.11 (E), in-phase coherent lines at the cardiac fundamental and harmonic frequencies highlight the presence of cardiac impulses in the β-γ band of the EEG (also visible but less significant in the θ band). Interestingly, a slight coherence of phase shift $\frac{\pi}{3}$, at the respiratory frequency from 100 to 150 min, also appears between the cardiac rate and the β-γ amplitude.

We can also examine the inter-band EEG coherence, in Fig.V.12. Again, the very coherent apneic mode is ubiquitous in all pairs of band magnitude neural signals, each with a particular phase shift.

The wake state is characterized by more intense modulations in the α-σ band, Fig.V.10

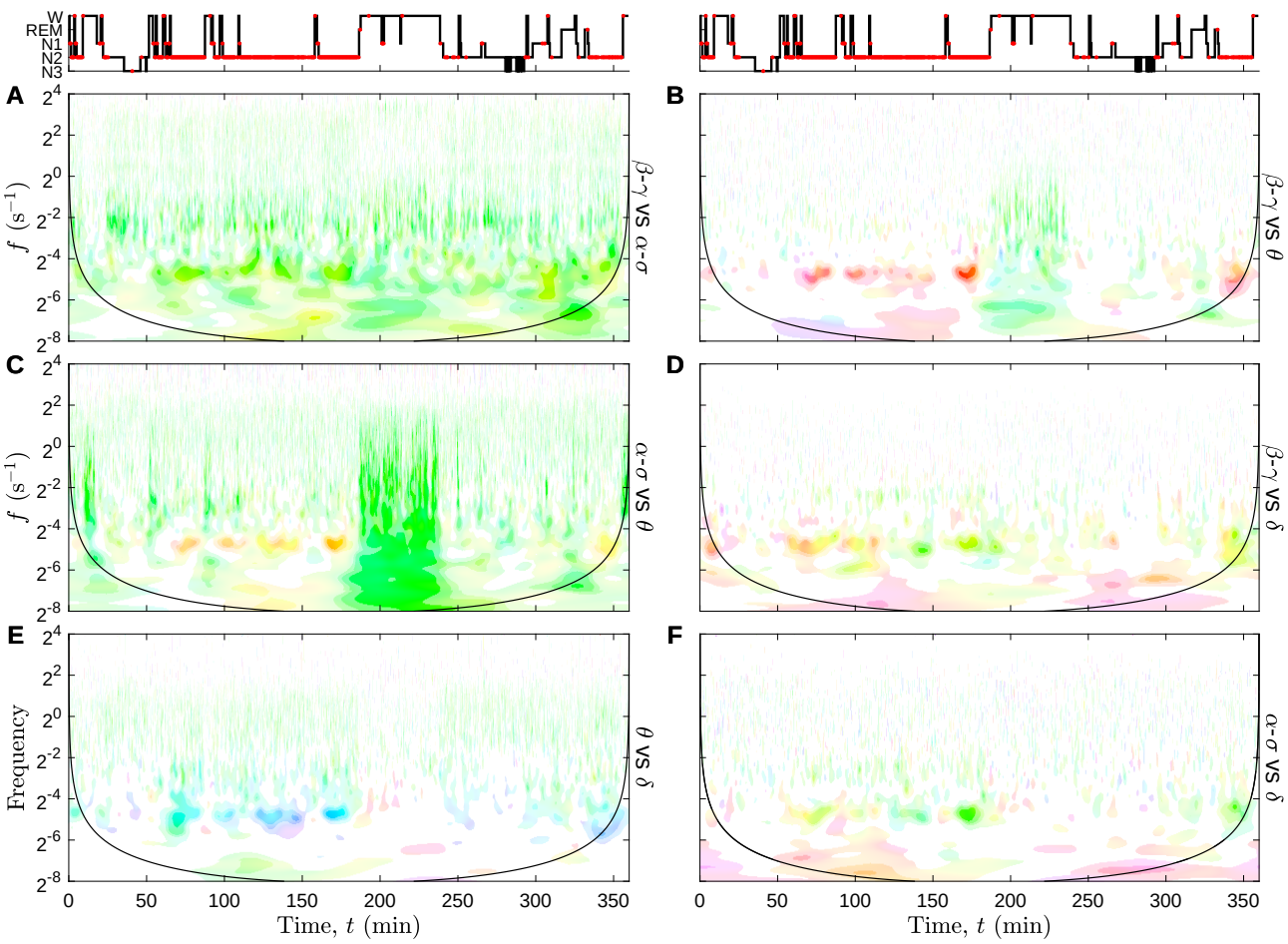


Figure V.12: *EEG inter-band time-frequency coherence for subject 3, between EEG band magnitude modulations shown in Fig.V.10. Same as in Fig.V.11.*

(B), and a strong in-phase coherence (green) between the θ and α-σ bands, Fig.V.12 (C). Such coherence increase is also observed between θ and β-γ, while the δ versus θ pair shows a drop of the inter-band coherence. These wake-related coherent or incoherent events concern a wide range of modulation frequency from 2 to 2^{-8} Hz, that can be partly related to the subject motions.

This inter-band EEG coherence contains a rich phenomenology able to characterize the connectivity between EEG bands, as recently observed using the correlation coefficient ρ_{xy} [Lin, Liu, et al. 2020]. However, it constitutes a high dimensional object to manipulate, considering that both bands should rather be considered as two distinct frequency parameters, in addition to the time and frequency variables of the coherence analysis. The dimensionality of this situation is reduced in the next section to study the evolution of the phase shift of the apneic modulation between the EEG bands.

2.2 Statistical properties of the apneic rhythm in the *shhs2* database

We complete this individual study of the coherence between physiological modulations with a statistical analysis of a large selection of subjects from the *shhs2* database, distributed in five groups corresponding to sleep apnea profiles.

Before selection, an individual analysis of the rates' coherence (Fig.V.11) is performed from all the subjects of the *shhs2* database. The statistics is computed in each group by averaging in time, over the selected and cumulated intervals, the time-frequency coherence $\gamma_{xy}(t, f; nQ, Q)$ and squared coherence. This reduces the huge amount of generated individual

data to the typical correlations between the physiological modulations, across frequencies and apnea groups.

Note that the squared coherence prevents interferences between distant times, when the value of the coherence phase varies, as would be the case when estimating the (global) spectral coherence. It is thus interpreted as an average strength of local coherence. The average phase is studied separately through the simple average coherence, and its full distribution is obtained from the angular density of coherence.

2.2.1 Apnea groups and time selection from annotations

The statistical analysis is obtained from all subjects from the *shhs2* polysomnography database (cf. Appendix 1), 2650 subjects in total. This represents 2.3 years of cumulated sleep time, defined as the duration between the first and the last sleep stage for each subject. The statistical results consist in averaging the produced individual time-frequency analysis over selected time intervals, that correspond to specific apnea profiles. These conditional time averages are interpreted as ergodic estimations in different statistical ensembles.

The selection of typical apneic subjects was based on respiratory events scored by clinicians when the amplitude of the airflow drops for more than 10 seconds, below 70% of the baseline for hypopnea (H) or below 25% of the baseline for obstructive and central sleep apnea. Obstructive sleep apnea (O) is distinguished from central sleep apnea (C) by a greater amplitude in the thoracic or the abdominal effort signal. 469264 of these three types of respiratory events are scored in total. The proportion of annotated apnea time interval in sleep time are as follows: 12.1% of hypopnea, 2.3% of obstructive apnea and 0.4% of central apnea.

We could have stopped there the selection of apneic time intervals of each type, regardless of the subjects. Instead, we refine it retaining only time intervals from subjects mainly affected by the type of apnea of interest. This has the effect of enhancing the specificity of the profiles, thus helping their interpretation.

Apnea group	Selection criteria	Number of subjects	Selected / total durations (h)	Number of apnea events
H	$p_H > 0.2$ $p_O + p_C < 0.01$	87	161 / 647	25277
O	$p_O > 0.1$	153	217 / 1152	25627
C	$p_C > 0.01$ $p_O < 0.1$	189	61 / 1463	10480
control	$p_H + p_O + p_C < 0.03$	129	938 / 957	0
all	none	2650	20114	469264

Table V.2: *Composition of the statistical ensembles, constructed from selected time intervals in distinct apnea groups of subjects in the shhs2 polysomnography database.*

For each person, we compute the proportion of the cumulated duration of sleep apnea in the total sleep duration, p_{ap}, ap = H, O, C. The groups are constructed as selection among the 2650 persons in the *shhs2* database using criteria on these apnea proportions. Group H corresponds to the 87 subjects affected by hypopnea more than 20% of their sleep time while the other apneas last less than 1%. Group O corresponds to the 153 subjects affected by obstructive apnea more than 10% of their sleep time. Group C corresponds to the 189 subjects affected by central apnea more than 1% of their sleep time and obstructive apnea less than 10%. These three groups have comparable sizes: group H contains 25277 hypopneas lasting 161 h out of 647 h of total sleep time, group O contains 25627 obstructive apnea lasting 217 h out of 1152

h of total sleep time, and group C contains 10480 central apnea lasting 61 h out of 1463 h of total sleep time.

Two other groups are defined. A fourth control group (no apnea) of is made of 129 subjects which are very little affected by any type of apnea: less than 3% of sleep time (3851 events lasting 19 h over a total of 957 h), and we retain only time intervals without apnea. We ensured that there is no subject in the intersection of any pair of these groups (that is not clear from the criteria alone). The fifth group (all) includes the whole *shhs2* database without any selection, and cumulates the 20114 h of sleep time over 2650 subjects. The table V.2 summarizes the characteristics of the five resulting time ensembles: $T_{\rm ap}$, ap = H, O, C, control, all.

2.2.2 HRV power density profiles as a point of comparison

Before applying the statistical analysis over time to the coherence between physiological signals, we apply it first on the simpler power density of the heart rate. Starting from the conventional heart rate estimator from the ECG's R-R intervals, $x(t) = f_c(t)$, the power spectral density S_{xx} is indeed a common metric for the heart rate variability (HRV). We use it as a point of comparison for the analysis that follows.

As in chapter IV, we use the PhysioNet cardiovascular signal toolbox [Vest, Da Poian, et al. 2018] to compute the HRV signals from the ECG's R-R intervals (*jqrs* algorithm), corrected automatically for ectopic and non-normal beats (also called N-N intervals). We then compute the typical power density, corresponding to each apnea profile, as the time averaged scalogram Eq.(II.23) conditioned to all selected time intervals, $t \in T_{\rm ap}$, in this apnea group (ap):

$$S^{\rm ap}_{xx}(f;Q)|f| = \langle |X(t,f;Q)|^2 \mid T_{\rm ap}\rangle_t \frac{Q}{\sqrt{\top}} \quad . \tag{V.5}$$

Not to complicate too much the notations, we write it here as a single time average, as if we compiled each individual recording $X(t,f;Q)$ one after the other in time, which is the meaning of the time ensembles $T_{\rm ap}$. Notice these are power log-frequency densities, represented in Fig.V.13 (A).

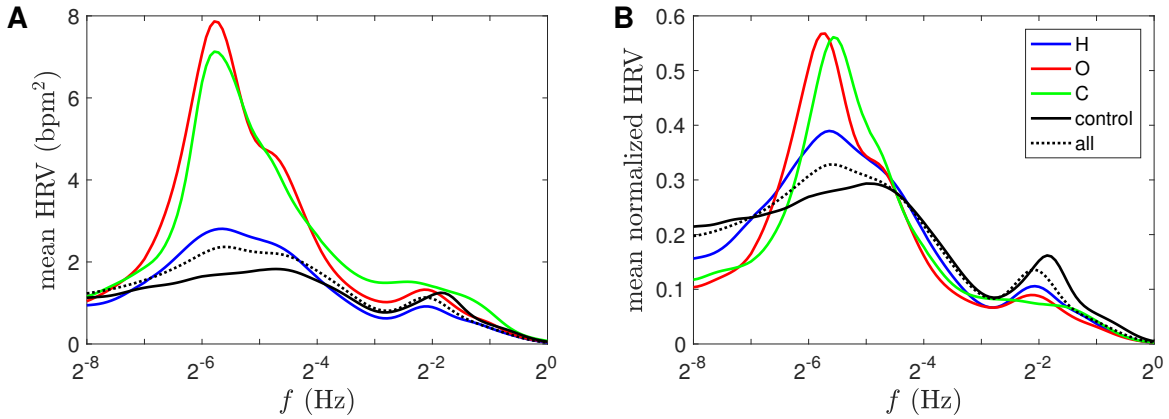


Figure V.13: *Comparison of typical heart rate variabilities (HRV) for different apnea profiles (ap): hypopnea (H), obstructive apnea (O), central apnea (C), no apnea (control) or total (all), see Table V.2. (A) power log-frequency density $S^{ap}_{xx}(f;Q)|f|$ ($Q = 5$) of the heart rate, in bpm² (min^{-2}), conditioned to selected time intervals in each apnea groups. (B) same time-averaged spectra, normalized for each subject prior to the group average (weighted by individual durations). The heart rate signals are estimated for each subject's ECG from the corrected R-R interval.*

In practice, the computation goes as follows: (i) we obtain the scalogram as the squared modulus of the wavelet transform ($Q = 5$) of the heart rate signal for each subject in *shhs2*, (ii)

we select the subjects for each group and the time intervals with sleep apnea of interest, (iii) we perform the (conditional or unconditional) time averages of the scalogram for each subject. The power density profiles Fig.V.13 (A) are then obtained by averaging individual spectra in each group, weighted by the duration of each individual time selection, whereas in Fig.V.13 (B) the individual power spectra are normalized (by the power) prior to the group average (also weighted by individual duration). As a result, the profiles in (A) give mean absolute values for these HRV spectra, whereas (B) shows the mean profiles relative to the strength of the HRV (by normalizing out strong or weak individual HRV).

In fact, we have also excluded from the time ensembles described in Table V.2 all 10 seconds epochs of non-physiological R-R intervals (outside 0.375 to 2 s during more than 15% of the duration, before correction). This mask, specific to this heart rate estimator, reduces the total durations in each group by about 5 to 10%. The mean normalized HRV spectra in Fig.V.13 (B) are nearly insensitive to this procedure compared the ones obtained without any data exclusion or with a stricter selection criterion (signal quality index $sqi > 0.9$, excluding 25 to 30% of the heart rate duration that does not coincide to the alternative estimation from the *sqrs* algorithm). These selections, supposedly affected by detection artefacts, tend to have strong amplitudes, so that their exclusion leads to a global decrease of the values of the mean HRV spectra in Fig.V.13 (A).

The presence of a peak at low frequencies in Fig.V.13, especially prominent in the case of obstructive and central sleep apnea (groups O and C) but much flatter in case of hypopnea (H) and without apnea (control), confirms the existence of the apneic rhythm in the heart rate [Penzel et al. 2016]. The apneic frequency at maximum intensity, near 0.02 Hz (most common value), is lower than previously observed on subject 3 (0.035 Hz). A much less prominent peak is observed at the faster respiratory frequency, sign of the RSA modulation of the heart rate. Its low intensity is comparable between apneic and non-apneic groups, see Fig.V.13 (A). The normalized HRV spectra in Fig.V.13 (B) suggests that RSA is a typically higher proportion of the total HRV power for non-apneic subjects compared to apneic ones.

We comment below the asymmetric shape of the apneic component. First, the bandwidth of the wavelet $\delta \log f \approx \log 1.65$ covers less than an octave (from the quality factor $Q = 5$), and imposes a log-normal (Gaussian) shape and a minimum width for an isolated peak. Then, the quality factor $Q = 5$ allows to distinguish the fundamental mode of the rhythm from its harmonics (of order 2) one octave higher. Finally, the variability of the apneic frequency over times and subjects in each group is likely to spread the averaged peaks on larger widths (more than one octave). These remarks explain the asymmetry of the apneic peak for the profiles O and C: the harmonics has a lower intensity and the inter-individual variability makes it merge with the fundamental component. This bimodal shape is simply a sign of non-circularity of the apneic modulation in the heart rate (it is not sinusoidal).

2.2.3 Several averages for the cardio-respiratory coherences

We now introduce again the respiratory signals and compare them to the heart rate with a coherence analysis, as in section 2.4, except that we distinguish the apneic profiles. We also use wavelet estimators only, for the modulations of the cardiac frequency, and the respiratory frequency and amplitude (FM and AM).

Instead of displaying all phase distributions as in Figs.IV.16 and IV.21 (all subjects and sleep times) for each apneic group, we focus here on the averaged coherence profiles. Two types of average coherence can be defined, the simple one (of complex value), or the squared one (positive):

$$\gamma_{xy}^{\mathrm{ap}}(f;Q,n) = \langle \gamma_{xy}(t,f;nQ,Q) \mid T_{\mathrm{ap}} \rangle_t \tag{V.6}$$

$$\gamma_{xy}^{2\,\mathrm{ap}}(f;Q,n) = \langle |\gamma_{xy}(t,f;nQ,Q)|^2 \mid T_{\mathrm{ap}} \rangle_t \quad . \tag{V.7}$$

Little information is lost in the case of cardio-respiratory coherences, since we have shown that their phase distributions are essentially unimodal, hence the mean phase is retrieved as the one of $\gamma^{\rm ap}_{xy}$. The quantity $\gamma^{2\,\rm ap}_{xy}$ averages local coherence intensities, regardless of their phase. This complementary information is valuable because it is not subject to destructive interference in case of broad phase distribution across times and subjects, contrary to $\gamma^{\rm ap}_{xy}$, so that it can detect locally significant coherence. The overall significance is straightforwardly determined from the number of statistical degrees of freedom $n_d \approx (1+n^2 e^{\frac{1}{Q^2}})^{\frac{1}{2}} \sim n$; in particular, the spurious coherence is $\gamma^2_{\rm sp} \approx \frac{1}{n+1}$.

Alternatively, an average coherence could be computed from the cross spectrum $S^{\rm ap}_{xy}$ with a conditional time average ($t \in T_{\rm ap}$) as in Eq.(V.5) and normalized by the corresponding auto spectra. The result is expected to be similar to $\gamma^{\rm ap}_{xy}$ and $\gamma^{2\,\rm ap}_{xy}$, but the significance of the latter is not controlled any more with this approach. artefacts, often sudden, intense and incoherent, are also better handled in Eqs.(V.6, V.7), that normalize before the average: pollution from intense artefacts thus only propagates over the short time resolution $n\delta t$, rather than dominating the averaged spectra before computing coherence.

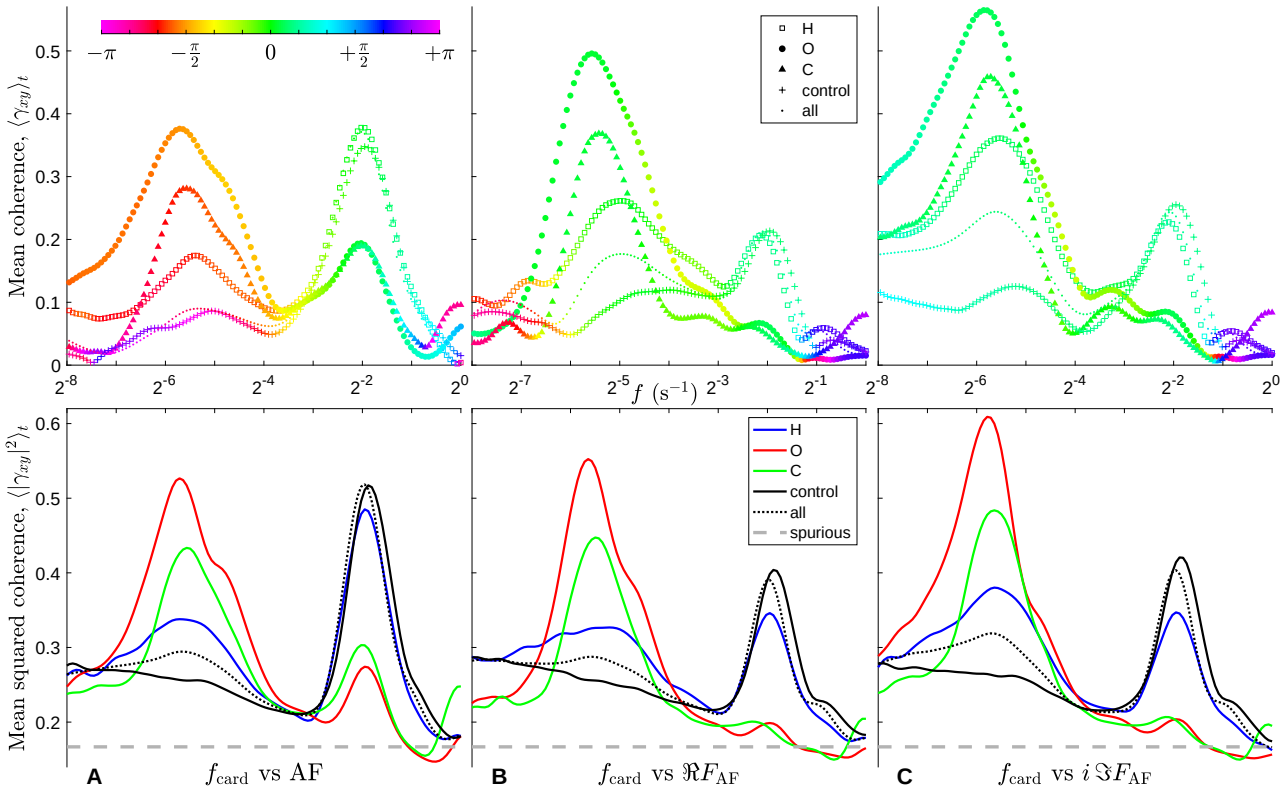


Figure V.14: *Average profiles of the coherence (top) and squared coherence (bottom) between the heart rate and different respiratory signals. Cardiac frequency (precise estimator) versus: (A) airflow recording AF, (B) breathing frequency $\Re\{F_{AF}\}$ and (C) amplitude modulation $i\Im\{F_{AF}\}$, generic estimators. Each profile corresponds to a conditional time average among subjects strongly affected by hypopnea "H", obstructive "O" or central "C" sleep apnea, without conditioning "all" or in the "control" group (without sleep apnea). See legends and text for details. The grey thick dashed line traces the expected level of spurious squared coherence $\gamma^2_{\rm sp} \approx \frac{1}{n+1}$, where $n = 5$.*

In Fig.V.14, the heart rate modulation (FM), precise wavelet estimator, is compared to the airflow recording (A) and to the generic wavelet estimators for the respiratory FM (B) and AM (C). The smoothing parameter has been chosen as small as possible, here $n = 5$, to capture transient episodes of coherence at a time resolution $n\delta t = \frac{nQ}{\sqrt{\Gamma} f} \approx 10 f^{-1}$ s. The duration rises

to 40 min at the slowest frequency, so that we excluded the first and last $n\delta t$-long time interval in each recording to avoid border effects. The total durations of the averaging time domains $T_{\rm ap}$ are much longer, making these profiles very robust. The expected spurious level for the squared coherence is $\gamma^2_{\rm sp} \approx 0.17$, grey thick dashed line in Fig.V.14.

There is a clear correspondence between these coherence profiles and the ones for the HRV spectrum at low frequencies: the peak of the apneic rhythm is prominent in the groups O and C (also asymmetric due to harmonics), inexistent in the control group and intermediary in the H group in all panels of Fig.V.14. However, the coherence approach brings novel insights in the respiratory frequency band: the RSA mode, very coherent both in the control and H group, is severely decreased in the O and C groups, even though the amplitude of this modulation remains (see Fig.V.13 (A)). This is best observed between the heart rate and the airflow in Fig.V.14 (A) where their mean coherence gets halved and their mean squared coherence drops from 0.35 to 0.1 above the spurious level. The peak of RSA coherence in Fig.V.14 (B, C) is nothing but a mere reflect of the one in panel (A), since the breathing modulation at the breathing frequency is in fact the breathing carrier wave (raw recording) itself, incidentally present in the rate estimators.

Therefore, a clear inversion occurs from the fast coherence without apnea to the slow coherence of apnea, the hypopnea being the pivotal state in between. Based on all O and C profiles, we have estimated the localization of the fundamental apneic (low frequency) mode: the global maximum lies at 0.019 ± 0.002 Hz (i.e. a period of 12 to 15 breathing cycles), and the widths suggests a variability of this rhythm among subjects ranging from 0.011 Hz to 0.038 Hz (i.e. 1.8 octave). There is close to no phase shift between apneic modulations of the heart rate (FM) and the ones of the breathing FM and AM, while they appear in-between phase quadrature (group O) and opposition (control group) for the airflow.

A slow cardio-respiratory coherence still exists in the control and hypopnea groups, with an increasing significance towards the lowest frequencies. We know from Fig.IV.21 that the coherence between cardiac and respiratory FM in Fig.V.14 (B) switches sign around 0.015 Hz, while its phase only slightly drifts between cardiac FM and respiratory AM (C). The cardio-respiratory coherence settles much more in the FM-AM pair (C) than in the FM-FM pair (B) during apnea (O and C groups).

All these observations can be formulated from the most generic (wide-band) estimator for the cardiac FM, based directly on the ECG, instead of the precise one (narrow-band) based on the amplitude of the ECG amplitude at 14 Hz (aECG14). This is shown in Fig.B.25 of Appendix 3.2, together with the effect of increasing the smoothing parameter and approximate degrees of freedom to $n = 10$ (as done in the next section). Although the coherence is slightly diminished, the profiles are essentially unchanged by these choices.

2.2.4 Average neuro-respiratory and neuro-cardiac coherences

Now, we introduce the neural magnitude signals $m^{\omega}_{\rm EEG}$, as extracted in Fig.V.6, that represent modulations in the EEG bands $\omega = \delta, \theta, \alpha$-$\sigma$ and β-γ. We compute the statistics in different apnea groups, for their coherence both with the breathing FM and cardiac FM, see Eqs.(V.6, V.7). These are performed with the generic FM estimators to reduce computational cost. The choice of the smoothing parameter $n = 10$ lowers the expected level of spurious squared coherence to $\gamma^2_{\rm sp} \approx 0.09$.

Starting with the squared coherence $\gamma^{2\,\rm ap}_{xy}(f; Q, n)$ in Fig.V.15, we confirm the ubiquitous appearance of the apneic coherent mode near 0.02 Hz in both neuro-respiratory and neuro-cardiac pairs of modulation signals. The most prominent peaks are obtained for obstructive apnea (O), then central apnea (C), and finally hypopnea (H), as different degrees of the same phenomenon. Again, the group without apnea (control) group exhibits a flat but significant

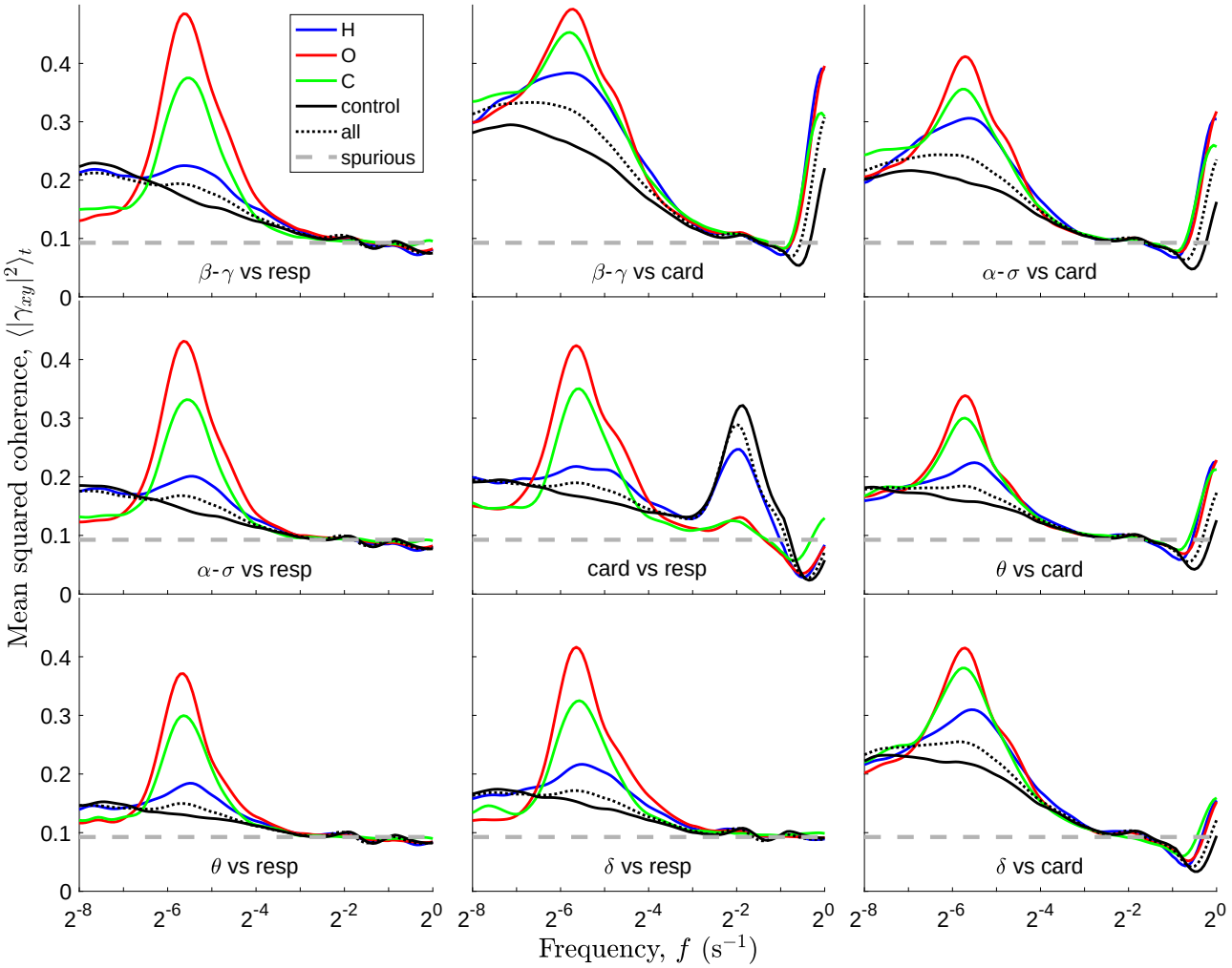


Figure V.15: *Average profiles of the squared coherence between neural band magnitude and respiratory (resp) or cardiac (card) frequency modulations (generic estimators). The neural band signals are denoted β-γ above 16 Hz, α-σ from 16 to 8 Hz, θ from 8 to 4 Hz, and δ from 4 to 1/4 Hz. Each profile correspond to a conditional time average among subjects strongly affected by hypopnea "H" (blue line), obstructive "O" (red line) or central "C" sleep apnea (green line), without conditioning "all" (black dotted line) or in the "control" group (without sleep apnea, black line). The expected level of spurious squared coherence is $\gamma^2_{\text{sp}} \approx \frac{1}{n+1}$, where $n = 10$ (thick grey line). See non-averaged individual coherences in Fig. V.11.*

coherence profile at low frequency, increasing towards very low frequency, and decreasing to the spurious level at high frequency. The result obtained for the full *shhs2* database (all) is intermediate between control and hypopnea (most common respiratory event).

In neural modulations, no coherent mode appears at the breathing frequency: the RSA modulation of the parasympathetic activity does not involve cortical neurons (but neurons in the medulla). However, a cardiac mode is visible at 1 Hz in all neuro-cardiac coherence profiles. This is a cross-talk between EEG and ECG, i.e. the occurrence of cardiac pulses in the EEG which is widespread in the *shhs2* database and of instrumental origin. A neural response to the cardiac activity also exists, called the heart-beat evoked response or potential [Park and Blanke 2019].

Neural modulations at very low frequencies, thus at long timescales (about 2 to 5 min), are possibly associated to changes of sleep-wake stages or micro states such as arousal. Their coherence with the heart and breathing rate vary with the EEG band, for instance it is especially important in the β-γ band but less significant in the θ band. We can notice that the very slow

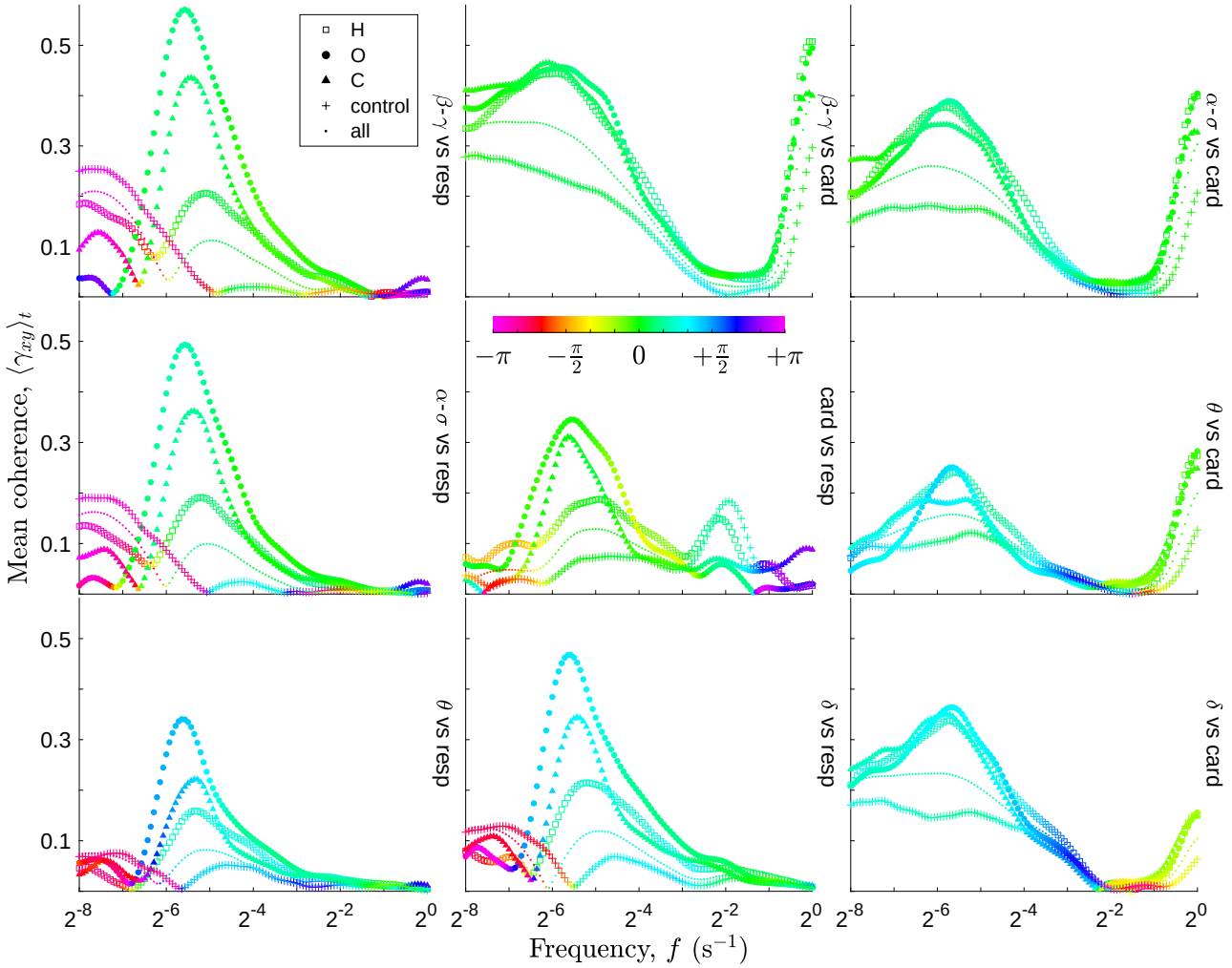


Figure V.16: *Average profiles of the coherence between neural band magnitude and respiratory (resp) or cardiac (card) frequency modulations (generic estimators). The neural band signals are denoted β-γ above 16 Hz, α-σ from 16 to 8 Hz, θ from 8 to 4 Hz, and δ from 4 to 1/4 Hz. Each profile correspond to a conditional time average among subjects strongly affected by hypopnea "H" (square), obstructive "O" (circle) or central "C" sleep apnea (triangle), without conditioning "all" (dot) or in the "control" group (plus). See non-averaged individual coherences in Fig.V.11.*

neuro-respiratory coherence is significantly lower, in all EEG bands, in the O and C apnea groups compared to the H and control groups, whereas no such difference is observed in the very slow neuro-cardiac coherence. For apneic subjects, their lack of correlation with breathing rate modulations may reflects the persistence in the N2 stage observed in Figs.V.3 during severe obstructive sleep apnea.

Examining the mean coherence in Fig.V.16, it is clear that very slow neuro-respiratory modulations constitute a distinct coherent component in phase opposition, as for the very slow cardio-respiratory FM modulations. It is stronger in the control and hypopnea groups than in the apnea groups, that coexist with the faster modulations in the slow range which are in-phase. In contrast, the same phase is observed for the slow and very slow neuro-cardiac modulations, thus appearing as a single and broader coherent component.

Finally, the phase shifts between neural modulations and cardio-respiratory ones clearly varies with the EEG band. This could also be observed on a single subject in Figs.V.11 and V.12. This dependence on the EEG band frequency is studied in the following.

2.2.5 Restoring the EEG band frequency continuum

In order to investigate the evolution of the phase between neural and cardio-respiratory modulations across bands ω of the EEG $x(t)$, we introduce again the continuum of scales from the time-frequency representation of the EEG intensity $X(t, f_1; Q_1)$, with band frequency $f_1 = f_\omega$ and resolution $\delta \log f_1 = \frac{\sqrt{\tau}}{Q_1}$, where the index 1 distinguishes the EEG band frequency continuum from the modulation frequencies ($f_2 = f$) of the subsequent coherence analysis.

The input signal representing neural modulations is the magnitude field $m_x(t, f_1; n_1 Q_1, Q_1)$, defined in Eq.(V.4) from the EEG power density. We use for this a short time-smoothing operation, $n_1 = 2$, that is sufficient to avoid instabilities while conserving a precise time resolution. For a coherence analysis, this field is then wavelet-transformed along time, with respect to a second quality factor Q_2. This yields the second frequency variable $f_2 = f$, that selects a scale for neural modulations in each band, as previously. We consider that $Q_2 = 5$ is sufficiently high to observe potential rhythms in physiological modulations. It may be denoted $M_x(t, f_1, f_2; n_1 Q_1, Q_1, Q_2)$, where the successive quality factors controls the resolutions of the time-frequency variables (t, f_1, f_2). Neural modulations are necessarily slower than the time resolution, so that the range of scales for modulations is bounded by a high frequency cut-off:

$$f_2 < \frac{\sqrt{\tau}}{n_1 Q_1} f_1 \quad , \tag{V.8}$$

i.e. essentially by the EEG band frequency $f_1 = f_\omega$, as confirmed in Fig.V.10. The quality factor $Q_1 = 5$, putting the stress on a precise time resolution, leads to $f_2 < f_1/4$. A more precise band frequency resolution with $Q_1 = 15$ yields $f_2 < f_1/12$; it requires a more important scale separation between carrier frequency f_1 and modulation frequency f_2. Therefore, to study the apneic rhythm $f_2 \approx 0.02$ Hz in a very low EEG band $f_1 = 0.1$ Hz, $Q_1 = 5$ is suitable while $Q_1 = 15$ is not: Eq.(V.8) can also be understood as a limitation on the resolution of neural modulations that can be achieved, controlled by the scale separation f_1/f_2.

The time-frequency coherence as in Fig.V.12, for a pair EEG magnitude fields with distinct continuous band frequencies, contains a lot of information about inter-scale neural dynamics. However, its practical computation is cumbersome, because of its high dimensionality (1 time, 3 frequencies). We compute instead the time-frequency coherence as in Fig.V.11, between the EEG magnitude field m_x and a reference cardiac or respiratory modulation signal y, that focuses on neuro-cardio and neuro-respiratory modulations. Computed once again from the practical time-smoothing operator, it may be written explicitly as $\gamma_{m_x y}(t, f_1, f_2; n_1 Q_1, Q_1; n_2 Q_2, Q_2)$, but we denote this coherence more simply as γ_{xy} in the following (where x is now the neural magnitude field). Some time resolution is given up in the way (in exchange for coherence significance), with a final duration δt greater than $\delta t_j = \frac{n_j Q_j}{\sqrt{\tau} f_j}$ for both $j = 1, 2$: $\delta t = \sqrt{\delta t_1^2 + \delta t_2^2}$.

Therefore, the duration of the polysomnography T (in average 455 min in the *shhs2* database) constrains the slowest scales for both the carrier waves and the modulations. However, the EEG frequency $f_1 = f_\omega$ cannot be extended much below the cut-off of the instrumental low-pass filter at 0.15 Hz. Therefore, we get the lowest modulation frequency in the coherence analysis:

$$f_2 > \frac{n_2 Q_2}{\sqrt{\tau}} T^{-1} \quad . \tag{V.9}$$

The time-smoothing parameter is fixed to a small value, $n_2 = 5$, for the coherence analysis, so that we could —at the very best— reach ultra-slow scales with cycles lasting $f_2^{-1} = 90$ min (such as the ultradian rhythm), using a quality factor as low as $Q = 2.5$.

The resulting coherence is finally studied statistically over time for the full *shhs2* database, divided into temporal ensembles for each apnea profile (ap) as explained in section 2.2.1. In

particular, the conditional time average Eq.(V.6) yields the coherence profile $\gamma^{\rm ap}_{xy}(f_1, f_2; Q_1, Q_2)$, where the time-smoothing parameters n_1, n_2 are omitted for the sake of simplicity. Indeed, they are only important to predict the range of f_2 and the level of spurious coherence before the time average, or after the time average for the squared coherence $\gamma^{2\,\rm ap}_{xy}$ Eq.(V.7).

2.2.6 Phase relation between neural and cardio-respiratory modulations

The phase $\phi_{xy} = \phi_x - \phi_y$ between neural modulations x (the magnitude field $m_{\rm EEG}$) and cardio-respiratory modulations y (derived from the complex rate estimator $F_{\rm ECG}$ or $F_{\rm AF}$) is the angle (complex argument) of the coherence, so that the average phase is the angle of $\gamma^{\rm ap}_{xy}(f_1, f_2; Q_1, Q_2)$, detailed for apnea each profile ap = H, O, C, control, all.

This phase depends on the continuum of EEG band frequency $f_1 = f_\omega$, and the scale of the modulations $f_2 = f_{\rm mod}$. Based on previous observations, we simplify the investigation by considering only two types of modulations: slow modulations and very slow modulations. The latter is found in a wide band around the frequency $f_2 = f_{\rm vs} = 0.005$ Hz (very slow), so that we choose a low quality factor $Q_{\rm vs} = 2.5$. The former one corresponds to apnea modulations, $0.01 < f_{\rm apnea} < 0.04$ Hz (slow) depending on the subject, that can be quite rhythmic (narrow-band), so that we choose a higher quality factor $Q_{\rm apnea} = 5$. To take into account the inter-individual variability of the apnea frequency, we select the value $f_2 = f_{\rm apnea}$ in the range $[0.01, 0.04]$ Hz that maximizes the coherence for each subject and at each EEG band $f_1 = f_\omega$. When the apnea rhythm is not present, this procedure may sometimes pick the broad-band very slow modulation instead.

At the apnea frequency $f_{\rm apnea}$, the statistics of ϕ_{xy} is illustrated in Fig.V.17 as a function of f_ω, for all subjects and over all times (ap = all), and for different cardio-respiratory modulations y (panels A to D). The average phase is plotted as a grey line when it represents a unique component, i.e. when its distribution is unimodal. We know it from the coherence-based distribution of the phase $G_{xy}(\phi\,; f_\omega, f_{\rm mod})$ Eq.(IV.5), also represented, and providing much more details on the spread of the phase and the associated density of coherence, generalizing the modulus of coherence, see Eq.(IV.6). To unambiguously differentiate this new situation, where the modulation frequency $f_1 = f_{\rm apnea}$ is fixed and the continuum is $f_2 = f_\omega$, we change the colour coding for the angular density of coherence compared to previous figures.

These distributions show a non-trivial evolution of the coherence (and its phase) between slow modulations (potentially linked to apnea) from the slowest neural waves (denoted δ^- in Table V.1) to the fastest (γ). In Fig.V.17 (A, B), the phase between neural magnitude and cardiac FM has a well-defined value; its distribution is unimodal with little dispersion. The phase is close to zero and the coherence most intense above 8 Hz (α-σ and β-γ bands), and the neural modulation in the slow band, below 2 Hz, has a slight advance and an important coherence with cardiac FM. This coherence drops significantly and the mean phase (grey line) shifts almost to quadrature ($\frac{\pi}{2}$) in between (partly δ^+ and θ). This behaviour is unchanged from the generic (A) to the precise (B) cardiac FM estimation, apart from a global enhancement of the contrast in the phase distribution. Overall, this shows that the neural magnitude are mainly positively correlated with the slow heart rate variations, up to some specificities in some EEG bands.

Similar features are observed in Fig.V.17 (D) where the neural magnitude is compared to respiratory AM. The coherence is twice weaker (see colour scale), but its angular density is also unimodal. The phase delay of respiratory AM with neural modulations is greater than the cardiac FM, increasing almost to phase opposition (π) near 3 Hz. A variation of the phase can also be noticed in the α-σ band (8 to 16 Hz).

Eventually a similar mode of the phase distribution is observed between neural modulations and respiratory FM in Fig.V.17 (C), with an even weaker coherence, together with a second

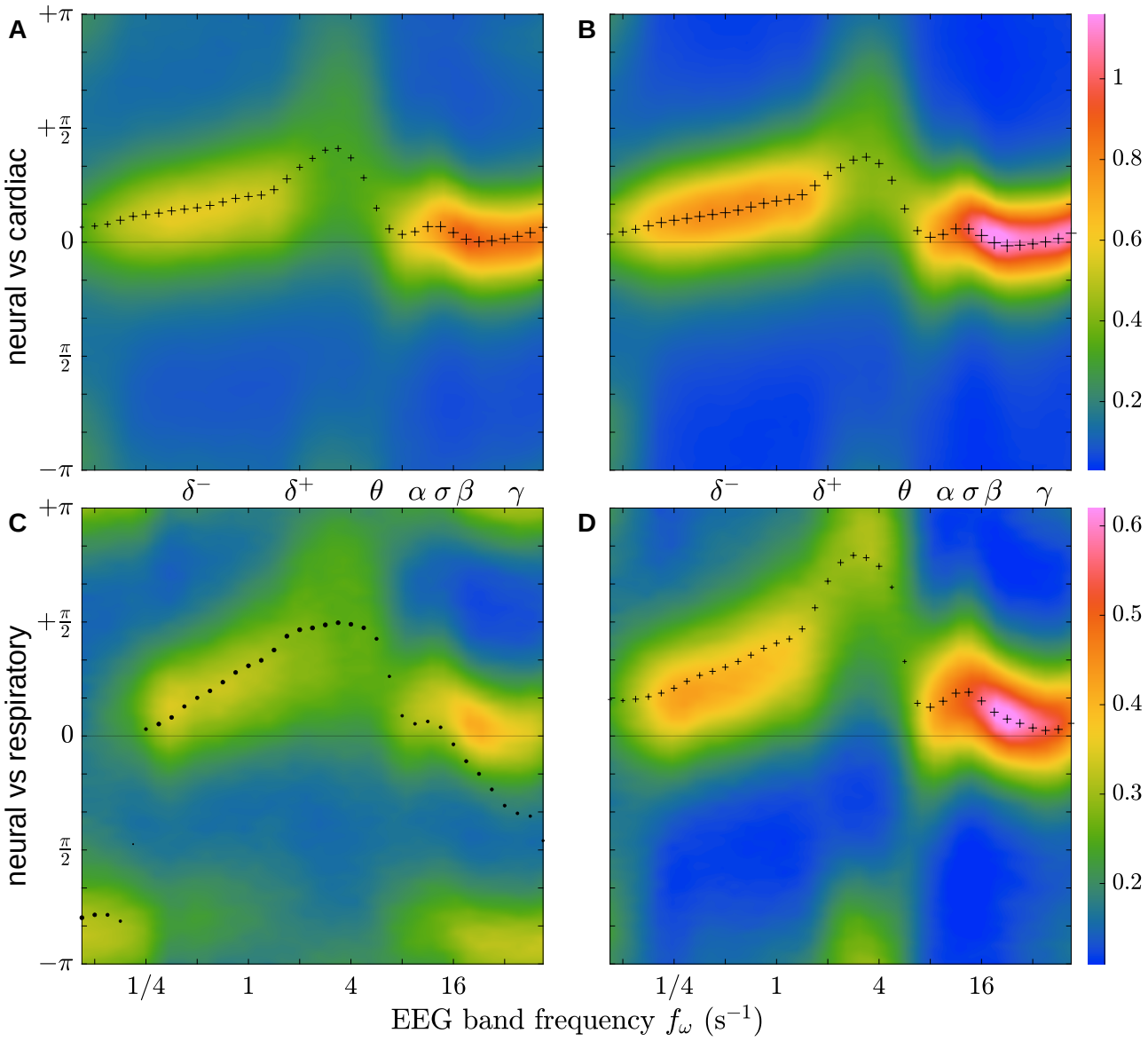


Figure V.17: *Coherence-based distribution (rad^{-1}) of the phase $\phi_{xy} = \phi_x - \phi_y$ between the neural modulations (x) and the cardio-respiratory modulations (y) in the slow range (0.01 to 0.04 Hz, that contains a potential apneic rhythm), as a function of the EEG band frequency (s^{-1}), over the full database shhs2. The modulation estimators for the cardiac frequency, generic in (A), precise in (B), and for the respiratory frequency (C) and amplitude (D), both generic, are compared to the EEG band magnitude with a continuous band frequency and a constant relative bandwidth controlled by $Q = 5$. Same colour scale in (A, B), and in (C, D). The mean phase is marked with black plus signs when the distribution is unimodal (dots when it is not).*

mode in phase opposition, clearly visible in the very low and very high EEG band. Based on the previous observations of a slow and a very slow components in the respiratory FM, we interpret this second component as the very slow one, wide-band, that leaks into the slow range $[0.01, 0.04]$ Hz, especially in the absence of apnea. We recall that this component of the respiratory FM is in phase opposition with the (very) slow neural modulations, cardiac FM and respiratory AM. The first one is likely to be the slow modulation that is enhanced during apnea. This suggests to characterize the interaction between neural magnitude field and breathing rate, Fig.V.17 (D), in the different apnea groups.

This is done in Fig.V.18 by conditioning the time average in Eq.(IV.5) to the selected time ensembles for each apnea profile. The two coherent modes are indeed separated between control

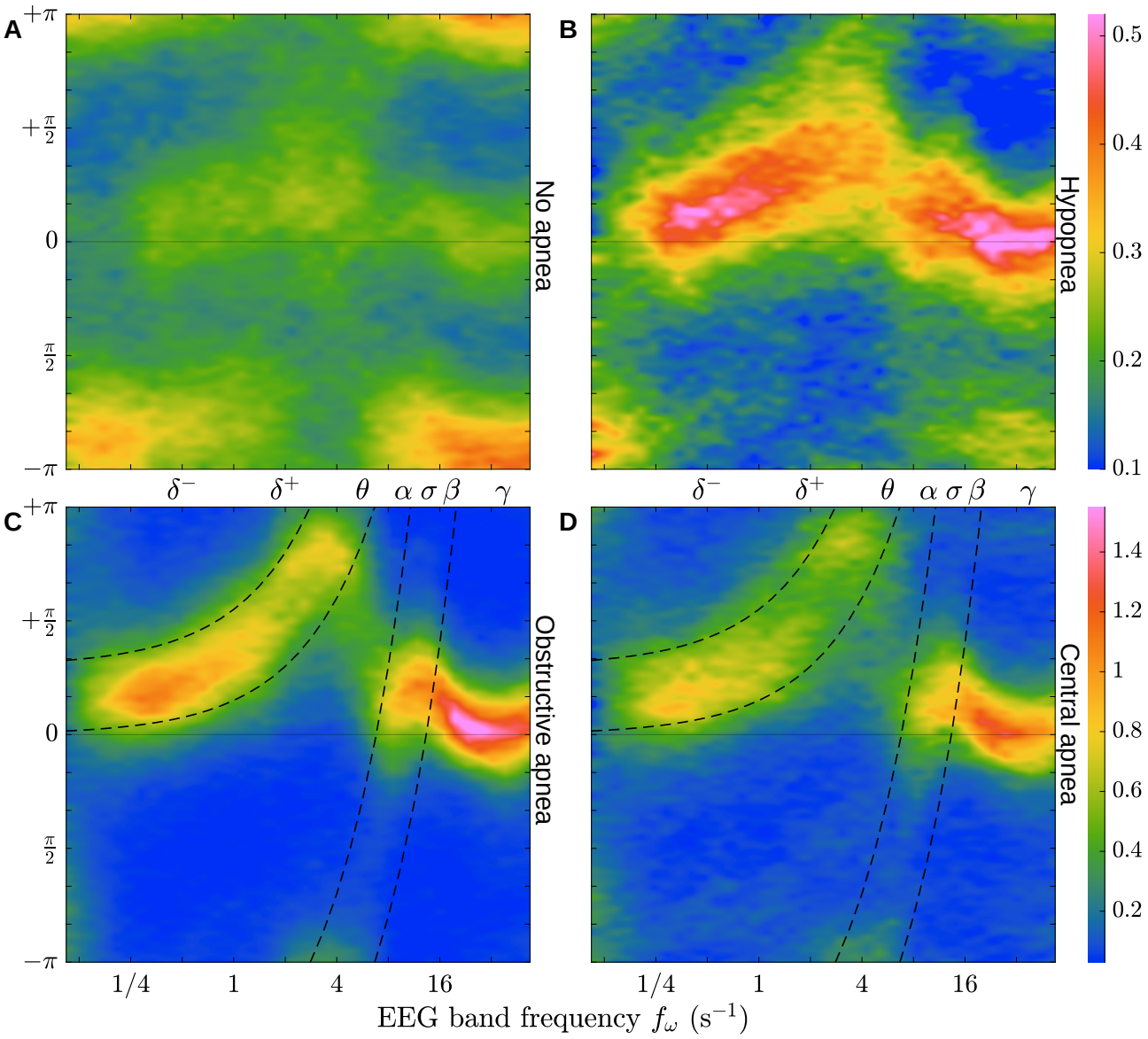


Figure V.18: *Comparison for different apneic profiles of the distribution (rad^{-1}) of the phase ϕ_{xy} between the neural modulations x and the respiratory frequency modulations y at the apnea frequency ~ 0.02 Hz, as a function of the EEG band frequency f_ω (s^{-1}). (A) Control group without sleep apnea. Group with (B) hypopnea, (C) obstructive or (D) central sleep apnea. Same colour scale in (A, B), and in (C, D). The phase-frequency relation $\phi_{xy}(f_\omega)/\Upsilon = \tau_\delta f_\omega + \beta$ (dashed black line), materializes a clear delay $\tau_\delta = 0.10 \pm 0.02$ s and a phase shift $\beta = 0.08 \pm 0.08$ of breathing compared to neural modulations (mainly in the δ band) during sleep apnea (C, D). See Fig.V.17 (C) for the unconditional distribution over shhs2.*

without apnea, and apnea groups. Without apnea (A), the coherence of the dominant mode in phase opposition also vanishes between 2 and 8 Hz, and a residual in-phase mode remains. We interpret it as an effect of undetected hypopnea episodes, that are missing in the available respiratory annotations (on which the time ensemble without apnea is built). Indeed, during hypopnea (B) this "in-phase" mode is dominant with a higher coherence, with again a specific alteration near the θ band, and the mode in phase opposition is residual in very low and high bands.

During more severe obstructive or central apnea, Fig.V.18 (C, D), the apneic mode is the only that remains with an intense coherence and a precise phase. It can no longer be described as "in-phase", only true in the γ band. The important progression of the phase

towards opposition in the δ (below 4 Hz) has a clear trend:

$$\phi_{xy}(f_\omega)/\Upsilon = \tau_\delta f_\omega + \beta \quad , \tag{V.10}$$

where $\tau_\delta = 0.10 \pm 0.02$ s and $\beta = 0.08 \pm 0.08$. The uncertainty on this affine relation between phase and EEG band frequency is represented with black dashed lines, folded into a second branches when crossing $\pm\pi$. Surprisingly, the second branch also explains quite well the specific phase distribution in the α-σ band (8 to 16 Hz), that alters the precisely in-phase γ modulations at high frequencies. The θ band (4 to 8 Hz) constitutes a "hole" in this phase relation: almost no coherence density is observed at the expected phase. Instead, there is some leakage of the coherence from the neighbouring bands that we interpret as an effect of the small quality factor for the EEG bands, $Q_1 = 5$ (the bandwidth is 3/4 of an octave). This interpretation seems in contradiction with the observation in Fig.V.16 of a clear apneic rhythm in the θ band; this points to an important limitation of the discrete neural band signals whose leakage in frequency is even worse due to the band integration. Note that apnea is often associated to a non-REM phase (N2), during which the θ band contains very incoherent EEG oscillations, see Table V.1. Although observed directly between two EEG contra-lateral recordings, it may support this interpretation of a θ "hole" of coherence.

What is the interpretation of the phase-frequency relation and its characteristic time $\tau_\delta = 0.10$ s?

We could first think of a kind of delay between the signals, either of instrumental origin (airflow signal), or as an event-related EEG potential (P100), evoked by apnea. This is a misconception, because such delay would be associated to a similar phase relation with respect to the (fixed) modulation frequency f_{apnea}: no delay can be observed between neural and breathing signals in Fig.V.18. Assuming a delay of 1 s exists between the recordings, it would only explain a small phase shift in Eq.(V.10), $\beta' \approx \beta + 0.02$, without influencing the value of the characteristic time τ_δ.

Let us expand its meaning: consider a typical subject of the O group, severely affected by an obstructive sleep apnea every 50 s (0.02 Hz). When the breath stops completely, this translates into a minimum of the respiratory frequency. At this time (and scale), the neural activity in the EEG band $f_\delta^+ = 4$ Hz reaches its maximum magnitude, that we can call apnea impulse. The apnea impulse then propagates towards lower EEG frequencies in the δ band down to 1/4 Hz. At this phase of the apnea cycle, the apnea impulse also appears in the α-σ band. Shortly later, the subject has totally recovered from the breathing obstruction, reaching a maximum of the respiratory frequency but also a maximum magnitude in the high β-γ EEG band (arousal, opposing deep sleep). Apnea impulses are indeed observed in the EEG as vertical high amplitude events during the long N2 stage in Fig.V.3, see also Fig.12 in [Prerau et al. 2017].

Therefore, the characteristic time $\tau_\delta = 0.10$ s in Eq.(V.10) describes the propagation "rate" of the apnea impulse from high δ^+ (4 Hz) to low δ^- (1/4 Hz), 4 octaves: it takes 10% of an apneic cycle for the δ impulse to propagate 1 Hz below. In other words, the neural response in the δ band of the EEG that is coherent with each apnea impulse is a linear chirp of rate $\tau_\delta^{-1} = 10$ Hz per apneic cycle (or rather 5 Hz in half a cycle).

This particular coherence phase between modulations of the neural activity and the breathing rate is in fact not specific to apnea. This is verified by reproducing the same study for the very slow modulations at $f_{\text{vs}} = 0.005$ Hz, coherent without apnea with the same characteristics and incoherent with severe apnea. In the control group (no apnea), neural modulations appear in coherence with the very slow breathing rate modulations, with the same linear evolution of the phase in the frequency range of the δ band, and constant in the β-γ band. The only difference is a constant phase shift $\pm\pi$ compared to the modulation as the apneic frequency, as observed previously in Figs.IV.21 (A) and V.16 (left column). Finally, the neural phase

response curve across EEG band frequencies is similar for all heart and breathing rate (FM and AM) modulation signals, as show in Fig.V.17 for all subjects.

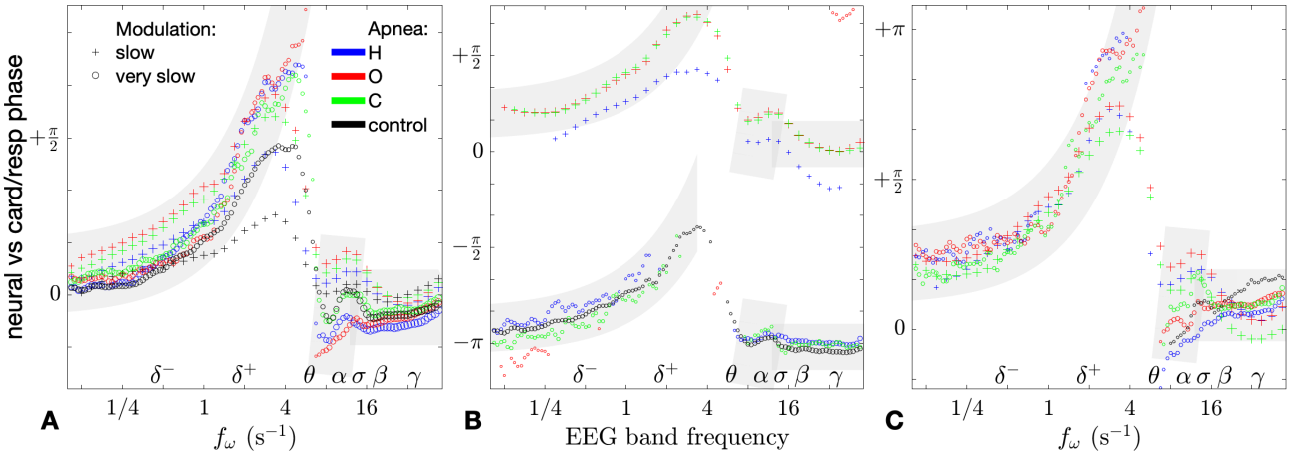


Figure V.19: *Average phase difference $\phi_{xy}(f_\omega, f_{mod})$ between the neural modulations x, across EEG bands f_ω, and 3 cardio-respiratory modulations signals y: (A) cardiac FM (precise estimator of the heart rate), (B) respiratory FM and (C) AM (both from the generic estimator). Very slow modulations (circle) correspond to $f_{mod} = 0.005$ Hz, and to $f_{mod} = f_{apnea} \in [0.01, 0.04]$ Hz for slow modulations (plus). The typical phase-frequency pattern in highlighted (light grey). The size of the marker is proportional to the modulus of the average coherence γ^{ap}_{xy} (not shown below 0.05 for very slow modulations and below 0.15 for slow ones); colours refer to the apnea profiles.*

All these observations are summarized in Fig.V.19, that represents only the mean phase $\phi_{xy}(f_\omega)$ from the average coherence $\gamma^{\rm ap}_{xy}(f_\omega, f_{\rm mod}; Q_1, Q_2)$ between the neural modulations (x) and different cardio-respiratory modulations (y, different in each panel A, B, C), for the each apnea profile and for both slow and very slow modulations $f_{\rm mod} = f_{\rm apnea}, f_{\rm vs}$. The main difference is a global phase shift, together with slight variations in the chirp rate τ_δ^{-1}.

Some difference specific to apnea can be spotted. In panel (A), the chirp rate τ_δ^{-1} (or chirpyness) is clearly lower without any apnea or hypopnea, for both slow and very slow neuro-cardiac modulations in the δ band. Interestingly, the central apnea is similar to no apnea and distinct from both hypopnea and obstructive apnea for very slow modulations in the α-σ band. Panel (B) corresponds to the observations of Fig.V.18. In panel (C) without apnea (black), no coherence is found between neural modulations and respiratory AM, except for very slow modulations in the higher EEG bands. In these bands, the neuro-cardiac phase for the different apnea profiles are subtly but clearly distinct when varying both f_ω and $f_{\rm mod}$.

We have essentially observed the typical phase response at multiple bands of the neural activity to any kind of slow cardio-respiratory modulations (apnea or not), characterized by a chirping δ modulation, a θ incoherence, an α-σ "echo" of the δ chirp, and a β-γ coherent modulation. We expect modulations of the oxygen saturation recorded from pulse oximetry to exhibit this same phase relation to neural modulations. Sleep apnea does affect the scale of modulations and the intensity of their coherence, the phase structure is essentially preserved.

We hypothesize that the slow broad-band neural modulations that are described from this coherence analysis corresponds to cyclic alternating patterns [Terzano et al. 2001], a cortical activity at multiple bands during NREM sleep composed of two "on/off" states alternating more or less irregularity in the slow range. Sleep stages and episodes with and without cyclic alternating pattern has been related to the dynamics of slow and very slow cardio-respiratory coupling [Thomas, Mietus, Peng, and Goldberger 2005]. These stages have not been studied yet in the present analysis, that could be readily extended based on the hypnogram. The power in

the δ band is known to correlate to the cardio-respiratory coupling, considered as an observable in [Thomas, Mietus, Peng, Guo, et al. 2014]. Although our approach is slightly different, these previous observations are completed extensively, although globally in time, from the pairwise characterization of the coherence with cardio-respiratory modulations in each EEG band, also uncovered the structure of their phase.

2.3 Sleep spindles, phase-amplitude coupling and stages

In this section, we choose an individual polysomnography to discuss investigations of different aspects of the EEG analysis that remains on a preliminary stage of development: phase-amplitude coupling, spindle detection, influence of sleep stages, EEG high-pass filtering and referencing. Only few individual observations have been collected, their statistical confirmation over a database have not been realized yet. The polysomnography has been selected for its rich phenomenology, with balanced sleep stages, the intermittent occurrence of sleep apnea, and for its clean multi-channel EEG recording.

The first aspect is the interaction that exists between the amplitude or the magnitude in some EEG bands, and the oscillation in a lower band: a certain phase of the slow EEG oscillation can trigger the intensity in the higher band. This phenomenon is known as phase-amplitude coupling, that can be implemented from different time-frequency perspectives able to distinguish the phase and the amplitude in different frequency bands [Munia and Aviyente 2019]. We describe it below as yet another type of intra-EEG coherence, between the amplitude in a band and the EEG recording itself. It can happen between slow δ waves ($\sim$1 Hz) and sleep spindles [Purcell et al. 2017], that are short and intermittent wave trains made of up to a dozen regular oscillations, precisely located in the σ band (between 12 and 16 Hz). This is illustrated in Fig.V.20, where the EEG (Pz) is shown in panel (A), and the slow oscillations and sleep spindles are isolated and compared in panel (B). For sleep spindles, we choose the quality factor $Q_1 = 5$ to cover the full σ band (around $f_\sigma = 13$) with a good time resolution; slow oscillations have a broader band (they are arrhythmic) around $f_\sigma = 0.9$, so that we choose $Q_2 = 2$. Sleep spindles appear preferentially in the rising front of the slow oscillation, maximum near the phase $-\frac{\pi}{2}$.

In Fig.V.20 (C), we recall that we can measure precisely the instantaneous frequency of the sleep spindle from a rate estimator, here restricted in the σ band, together with the associated uncertainty ΔF_x, see Eq.(III.37). Notice that the instantaneous frequency of the spindles can vary a lot, even though they are short-lived: they chirp. The uncertainty is important at times without spindles and vanishes when there is a spindle, so that we can use it as an indicator to detect their presence. For that purpose, we compare in panel (D) the (root-mean-square) amplitude in this band to the inverse of the relative uncertainty:

$$\frac{|F_x|}{\Delta F_x} = \frac{|\gamma_{\dot{x}x}|}{\sqrt{1-|\gamma_{\dot{x}x}|^2}} \quad , \tag{V.11}$$

that we call the *certainty* of the oscillations. It may be interpreted as the number of cycles that can be distinguished unambiguously (before doubting of one cycle). The global coherence between certainty (x) and amplitude (y) signals is about $\rho_{x^+y^+} = 0.64e^{i0.04}$: they are strongly correlated positively (the phase is negligible), but not perfectly correlated, hence substantially different.

For detecting and isolating spindles, we suggest using the certainty signal rather than an amplitude signal, because it is not expected to be affected by the electrode impedance, nor suffer from a baseline wander (as is the amplitude): applied in a narrow band, the certainty directly represents the degree of regularity of the wave, between circular oscillation and stochastic

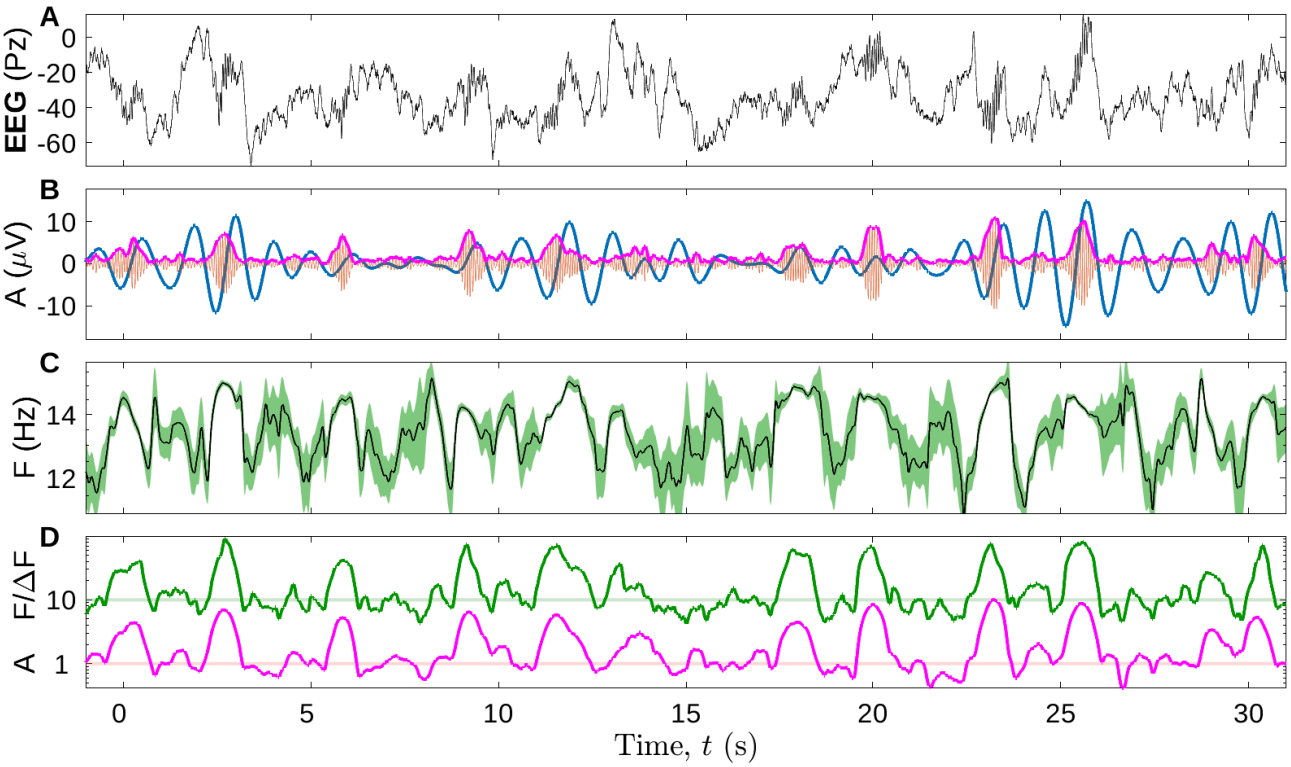


Figure V.20: *Sleep spindles and slow oscillations during deep sleep (N3) for subject 7. (A) EEG signal $x(t)$ from the electrode Pz against the average reference (in μV). (B) Band-limited oscillations (μV), computed as $2X(t, f; Q)$, with $Q = 2$ and $f_\delta = 0.88$ Hz for slow oscillations (thick blue line), with $Q = 5$ and $f_\sigma = 13.2$ Hz for sleep spindles (thin red line) and their amplitude (aEEG13, thick magenta line). (C) Instantaneous frequency in the σ band (12 to 16 Hz), $\Re\{F_x\}$ (black line), and its uncertainty ΔF_x (green area) reported in (D) as a "certainty" ratio $\frac{|F_x|}{\Delta F_x}$ (thick green line) compared to the amplitude of the spindles (thick magenta line, in μV) computed under the same conditions.*

fluctuation, whatever its amplitude. As a proof of concept, we compute the certainty in the α-σ band, and we use it to segment the EEG band signal into oscillations that exceed a threshold certainty of 9. In order to alleviate the thresholding effects, we merge close but fragmented segments whenever their overall certainty is above the threshold, and we discard segments that last less than 5 oscillations. The threshold should be adjusted depending on the choice of the quality factors (here $Q_+ = 7$ and $Q_- \approx 3$). The result during a full polysomnography is represented as a scatter plot in Fig.V.21 (A), where each marker represents an isolated oscillatory event, with its mean instantaneous frequency and number of oscillations (size). Two alternating populations, spindles of the σ band and α waves, appear with densities that depend on the sleep stages: spindles are dense during deep sleep (N3), less during N2, while α are dense during wake, less during REM, so that the EEG activity in this octave (between 8 and 16 Hz) reflects the characteristics of the whole polysomnography.

In the next panels of Fig.V.21, we describe the phase-amplitude coupling between not only sleep spindles but also θ waves slow EEG oscillations: while the modulus of coherence measures the strength of the interaction, the angle of coherence provides the phase of the slow oscillation at which the amplitude in the higher band is maximum. Several types of signals can be used as the "amplitude" of the sleep spindles, from the sophisticated certainty signal to magnitude and power estimators or simply the amplitude of the wavelet transform. We notice that they provide quite similar results (except the magnitude, that amplifies the noisy background), so that we simply choose the aEEG13 signal (EEG amplitude at 13.2 Hz for $Q_1 = 5$) as in

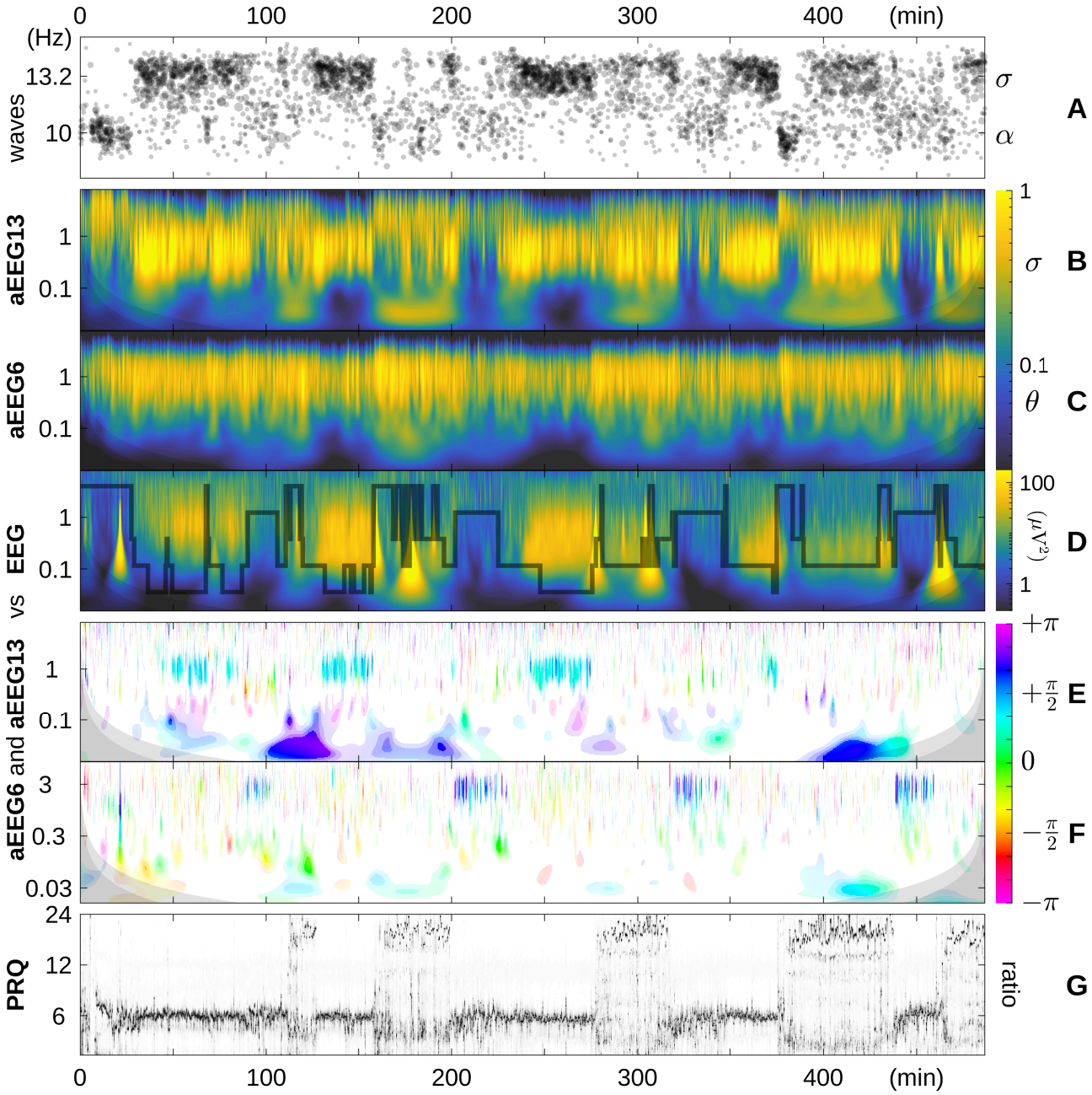


Figure V.21: *Neural σ, α, θ waves, their coupling to slow δ waves and PRQ, during the full polysomnography of subject 7. (A) Neural activity in the α-σ band, segmented as waves with certainty > 9 and more than 5 oscillations. (B, C, D) Power log-frequency density $S_{xx}(t, f; Q_+ = 100, Q_- = 2)|f|$ (μV) in the range $f \in [0.016, 8]$ Hz, where x is: (B) the amplitude of sleep spindles around 13.2 Hz (aEEG13), (C) the amplitude of θ waves around 6 Hz (aEEG6), and (D) the slow oscillations from the EEG signal. The hypnogram is superimposed to the EEG power density. (E, F) Amplitude-phase coupling (x vs y) from their time-frequency coherence $\gamma_{xy}(t, f; Q_+, Q_-)$: (E) aEEG13 versus EEG, (F) aEEG6 versus EEG. The 4 thresholds of colour saturation, at $|\gamma_{xy}| = 0.21, 0.30, 0.36, 0.41$ correspond to increasing significance levels, with p-values 10^{-1}, 10^{-2}, 10^{-3} and 10^{-4} respectively. (G) Pulse-respiration quotient as the distribution of the cardio-respiratory rates ratio from Fig.IV.23.*

Fig.V.20 and the aEEG6 signal for θ waves. The phase for the slow oscillation, is the one of the wavelet transform of the EEG at the relevant low frequency ($Q = 2$). The amplitude signals are decomposed in the same way, with $Q = 2$ to focus on singular bursts. Their power densities are presented in Fig.V.21 (B, C, D), followed by the time-frequency coherence analysis of the amplitude modulations (AM) versus the EEG recording, for the spindles (E) and the θ waves (F).

The amplitude-phase coupling appears as very distinct spots of very significant ($p < 10^{-4}$) but intermittent coherence (vertical bursts), around 1 Hz and systematically during the N3 stage for sleep spindles, that indeed arrive with a phase advance of $\frac{\pi}{2}$ compared to slow oscillations, and around 3 Hz during REM stage for θ waves with a phase $\frac{2\pi}{3}$, that seems to switch to $-\frac{\pi}{3}$ during N3 (less significantly, $p > 10^{-2}$). Note that the scale separation is only one octave between θ waves and δ^+ oscillations.

We can also notice a slight but systematic phase gradients of less significant coherence below these spots, clearly positive in panel (F), sign of a delay (rather than simply a phase) between θ amplitude and δ^+ oscillations. This gradient seems negative for sleep spindles (they would be late) from the angle π near 0.2 Hz (E); however, we must interpret it carefully, since the EEG is typically high-pass filtered around this scale.

In spite of the instrumental filtering, strong coherence regions are visible a decade below (around 0.03 Hz), in relation to the occurrence of sleep apnea during the N2 stage. This apneic modulation is especially important in the σ band, as can be guessed from low horizontal lines in the aEEG13 power density, in Fig.V.21 (B). This is confirmed from the pulse-respiration quotient (PRQ) represented in the last panel (G), that oscillates repetitively between 3 or 4 to about 20 heart beats per breathing cycle, sign of obstructive sleep apnea. Out of these apneic phases, the PRQ is steadily locked to 6 during N3 and fluctuates more broadly during REM.

From these observations, we can summarize the sequence of main sleep patterns of subject 7 as $wSRASARSARSwARA$, where w stands for wake (and sleep onset), S for spindles-slow waves coupling / synchronous heart-breath / deep sleep (N3 stage), R for rêve (REM, dreaming) / θ-δ^+ coupling / heart-breath fluctuations, and A for apnea oscillations of heart-breath / N2 stage.

Eventually, we discuss the dependence of this individual observation to the choice of the EEG electrodes locations, here Pz (parietal on the sagittal plane) and referenced against the average of all 11 available channels. The quality of the detection of sleep spindles can degrade when varying its location, and the observation of the amplitude-phase coupling can simply disappear. These observations can be done exclusively from electrodes C3, Cz and Pz, against A1, A2 and the average reference. When the polarity is respected (the former minus the latter), their inclusion into a rough surface Laplacian (linear combination between neighbours) also leads to the same observations, stressing the spatial localization of the phenomenon. Interestingly, minus the average reference alone also works: the (unlabelled) ground electrode has certainly been placed near Pz.

Summary of section 2

This last section starts and finishes with the study of coherence in the polysomnography of two subjects, severely affected by obstructive sleep apnea for the former, with a lighter apnea for the latter. In-between, an extensive statistical analysis of coherence in neuro-cardio-respiratory modulations has been performed over the 2650 subjects and 2.3 years or cumulated sleep recordings in the large *shhs2* database. The time-frequency power density and coherence is first computed individually between pairs of physiological observables: cardiac FM, respiratory FM and AM (from rate estimators) and neural band AM (magnitude signals), as illustrated for the first subject. Then, typical coherence profiles are computed from conditional time averages over subgroups and selected intervals, based on the clinician annotations of apnea events of the different types: hypopnea, obstructive and central apnea, or their absence. The presence of apnea, especially obstructive and central, results in the ubiquitous appearance of slow narrow-band coherent modulations in all pairs of observables, localized between 0.01 and 0.04 Hz depending on the subjects. This rhythm corresponds to the repetitive occurrence of apnea and recovery cycles in a particularly regular manner, increasing both the amplitude and coherence of corresponding modulations, appearing as spectral peaks compared the flat baseline without apnea and the intermediary hypopnea profile. This slow apneic component in the physiological variability is shown to be concomitant with an important coherence reduction, not only in the fast and narrow-band RSA cardio-respiratory interaction, but also for the very slow and broad-band (irregular) neuro-respiratory and cardio-respiratory component in phase opposition.

The phase between neural modulations associated to apnea are observed to vary with the considered EEG band, as well as their coherence with other modulations. By retrieving the continuum of EEG bands, we observe how the coherence and its phase evolves with the band frequency. Although the coherence varies with the apnea profiles, the phase-frequency relation is shown to be essentially conserved, up to a global phase shift, for all apnea profiles, slow and very slow components and for all pairs of neuro-cardiac and neuro-respiratory coherence. It is composed of a chirp in the δ band, a hole in the θ band, an echo of the chirp in the α-σ band and a constant phase in the β-γ band. The chirp has a linear rate of about 10 Hz, that starts at high δ^+ during apnea (or equivalent cardio-respiratory slow-down) to low δ^- during recovery. This neural activity is hypothesized to be cyclic alternating patterns of the cortical activity, whose multi-band phase relation to cardio-respiratory modulations has been uncovered.

Eventually, the last individual polysomnography illustrates the specific relation that exists between slow δ oscillations and sleep spindles of the σ band during NREM sleep, together with other stage-specific phenomena. The discussion comprises the detection of sleep spindles and α waves from the phase information in the α-σ band, the coupling of their amplitude to the phase of slow oscillations during deep sleep (N3 stage), and the possibility to describe this phase-amplitude coupling as the coherence between EEG amplitude signals and the EEG recording itself. We show that a similar phenomenon exists for θ waves during REM sleep. The inclusion of the stage-dependent variations of the heart-breath ratio (or pulse-respiration quotient), completes the picture of the sequence of wake-sleep and apnea stages in this individual polysomnography.

Appendix B

1 Sources of biophysical recordings: description of the databases

The human physiological recordings that are analysed in part B are sourced in existing databases. In the results presented in this thesis, these polysomnography databases have been used in two ways: either as a source of individual recordings to study the dynamics of single trajectories, or as a pool of numerous recordings for collective statistics over time.

We detail here the sources of the selected individual recordings, and the databases that have served for collective statistics.

1.1 Polysomnography ensemble for collective statistics

Collective statistics have been obtained from two datasets: 5.6 hours of recording from the 42 subjects of *capnobase*, and more than 2.3 years of recording from 2650 subjects of *shhs2*.

The CapnoBase TBME R-R benchmark dataset (*capnobase.org*) contains 42 cases of 8-min recordings obtained on children and adults aged 0 to 75, from anaesthesia monitors during elective surgery. For each subject, they are composed of CO_2 waveforms (capnograms), photoplethysmograms from pulse oximetry (PPG) and electrocardiograms (ECG), together with expert annotations. Although it was intended to assess the characterization of the breathing rhythm from the cardiac recordings [Charlton et al. 2018; Karlen et al. 2013], we only used the simultaneous PPG and ECG recordings to vary the input cardiac recording for heart rate estimation.

The Sleep Heart Health Study is a multi-cohort study focused on sleep-disordered breathing and cardiovascular outcomes, that compiles polysomnographic signals from 5804 participants aged 40 and older, recruited from existing studies with pre-collection of cardiovascular risk factors [Lind et al. 2003; Quan et al. 1997; Zhang et al. 2018]. Available from the National Sleep Research Resource (*sleepdata.org*), all the samples were annotated with sleep events such as sleep stage, arousal and apnea. We used the entire dataset of the visit 2 (*shhs2*) composed of polysomnography from 2651 subjects, but one polysomnography (the 687^{th}), that has been excluded because its annotation file has been lost. Polysomnograms were obtained in an unattended setting, usually in the homes of the participants, by trained and certified technicians. The recording montage consisted of two bipolar electroencephalograms (EEG), right and left electrooculograms (EOG), an electromyogram (EMG), thoracic and abdominal excursions (THOR and ABDO) recorded by inductive plethysmography bands, "airflow" (AF) detected by a nasal-oral thermocouple, fingertip pulse oximetry, ECG from a bipolar lead and derived heart rate (PR), body position and ambient light.

1.2 Selected polysomnography for individual dynamics

A manual screening of individual recordings has been an essential first step to learn from observation the important aspects of both polysomnography and wavelet representations. The individual examples that are reproduced in this thesis aim at giving a taste of this exploration. They have been selected based on different considerations, including the quality of the measurement, the richness of the phenomena they contain, and their representativity of the considerable amount of raw images that has been computed.

Subjects 1 and 2 correspond respectively to the recording 0331_8*min* (32 years old, 60 kg), and 0009_8*min* (10 years old, 46 kg), from *capnobase*.

Subject 3 (male, 40 years old, 108 kg) corresponds to the recording *slp04* from the *MIT-BIH* polysomnographic dataset [Goldberger et al. 2000; Ichimaru and Moody 1999], available from the PhysioNet Research Resource for Complex Physiological Signals (*physionet.org*). It includes 16 male subjects, aged 32 to 56 (mean age 43), with weights ranging from 89 to 152 kg (mean weight 119 kg), and most of them were affected by a severe obstructive sleep apnea. Subject 3 has an apnea-hypopnea index (AHI) of 59.8 (among the most severe sleep apnea), and its polysomnography is composed of an ECG, an EEG, an invasive blood pressure signal (BP measured using a catheter in the radial artery), and an "airflow" signal (AF) from a nasal thermistor.

Subjects 4, 5 and 6 correspond respectively to recordings *shhs2-200079* (female, 56 years old), *shhs2-200901* (female, 54 years old) and *shhs2-205136* (female, 61 years old) from the second visit of the Sleep Heart Health Study (*shhs2*).

Subject 7 is an elderly adult that participated in the ongoing *FUSO* project, that aims at developing an automatic sleep spindle analyser, see a previous work that includes sleep spindles detection [Taillard et al. 2019].

Subject 8, 9 and 10 correspond respectively to patients *p33, p20* and control *n14* in a study of dynamic IR thermograms to assist in early breast cancer diagnosis from a multifractal analysis [Gerasimova et al. 2014; Gerasimova-Chechkina et al. 2016], and propose an alternative to the more invasive X-ray mammography. These were the first dataset examined in this thesis, but also the most difficult to process from their spatiotemporal nature. Signals are extracted from spatial average and standard deviation in a region of interest, represented of Fig.B.22. We have partially decoupled thermal fluctuations from geometric fluctuations of the body, by correcting small body deformations with an affine transformation. The use of spatial wavelets is expected to improve this analysis.

2 Cepstral approach, a high quality factor perspective on the heart rate

We present the case of a physiological signal, derived from the infrared thermograms of subject 9. These recordings were studied under the angle of the multifractal analysis, i.e. at a small quality factor to localized vertical structures. Here, we take the high quality factor perspective to put into evidence the presence of physiological rhythms, characterized by horizontal harmonic structures.

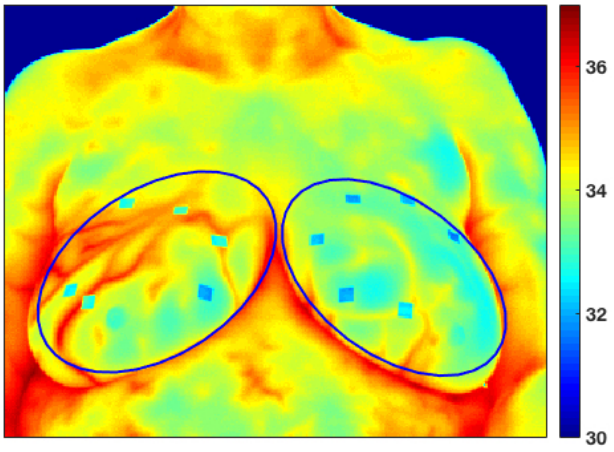


Figure B.22: *Temperature map (°C) of the chest of subject 9 corresponding to a single frame of the dynamic infrared thermogram. Colder squared marks placed on the warmer skin constitute references to track motions and geometry. Elliptical regions of interest are enclosing each breast: the higher temperature and apparent vascularization of the right one (plotted on the left) is related to the presence of a tumour, whereas the left (colder) one in healthy.*

The signal $x(t)$ in Fig.B.23 (A) is the standard deviation of the skin temperature, obtained as a spatial average over the right breast of a subject. The panel (B) represents its power density $S_{xx}(t, f; Q_+, Q_-)|f|$ with a quality factor as high as $Q_- = 20 \approx Q_+$. This allows to put into evidence any regular rhythm as a harmonic spectrum up to the 8^{th} order. Among fluctuations, this signal is influenced by the breathing rhythm, visible as a slight oscillation in the signal and a horizontal structure at low frequency, that can be identified as the fundamental respiratory component. Some fragments at below 1 Hz might be interpreted as a slight non-circularity (order 2 and 3 harmonics).

The observation of interest here is the clear but tenuous harmonic lines observed above 2 Hz. They are separated by about 1 Hz, sign of the presence of the cardiac rhythm. Similarly to the ECG, many harmonic lines are presents, except for the fundamental line. The strategy of the ECG consists in choosing a small quality factor to detect beats in a band-limited amplitude signal (like the aECG14 signal). In this case, we choose the band frequency $f = 4$ Hz and a broad band with a quality factor as small as $Q = 2$, to cover all available harmonic intensity. The generic narrow-band quadratic rate estimator is then applied near 1 Hz, for $Q_- = 5$ and $Q_+ \approx 10$. Apart from occasional perturbations from vertical singularities, the heart rate extraction works in this low signal-to-noise situation, as shown as a black dotted line in Fig.B.23 (C).

The rate signal estimated from the low quality factor perspective is superimposed in Fig.B.23 (C) to the outcome of an alternative high quality factor perspective on the temperature signal, starting with $Q = 20$ instead of $Q = 2$, and based on the cepstral analysis [Randall 2017].

It consists in detecting periodicities in linear frequency, sign of the harmonic series, from a Fourier analysis of the magnitude of the spectrum. The (real) *cepstrum* may be define generically from a positive spectrum of $x(t)$ as:

$$\mathcal{C}_x(\tau) = \int_0^\infty \log(S_x(f)) \cos(\Upsilon f \tau) \mathrm{d}f \quad , \tag{B.12}$$

where τ is called the *quefrency* and its value at a peak represents the period of a rhythm in the signal, equivalently τ^{-1} is its fundamental frequency. A time-dependent power spectral density as in Fig.B.23 (B) turns it into a time-quefrency object. However, its direct application does not achieve the detection of the heart period τ_c, due to the strong amplitude decay at high frequency.

We propose to adapt it in a way that is insensitive to this spectral trend $\sim f^{-\alpha}$, that consists in looking for the harmonic spectral periodicity in $\partial_{\log f} \log S_x(f)$ instead. It also requires adapting cos into $-\sin$ due to the derivative. Finally, we prefer a log-frequency integral, more compliant with the wavelet transform, yielding the following dimensionless time-quefrency

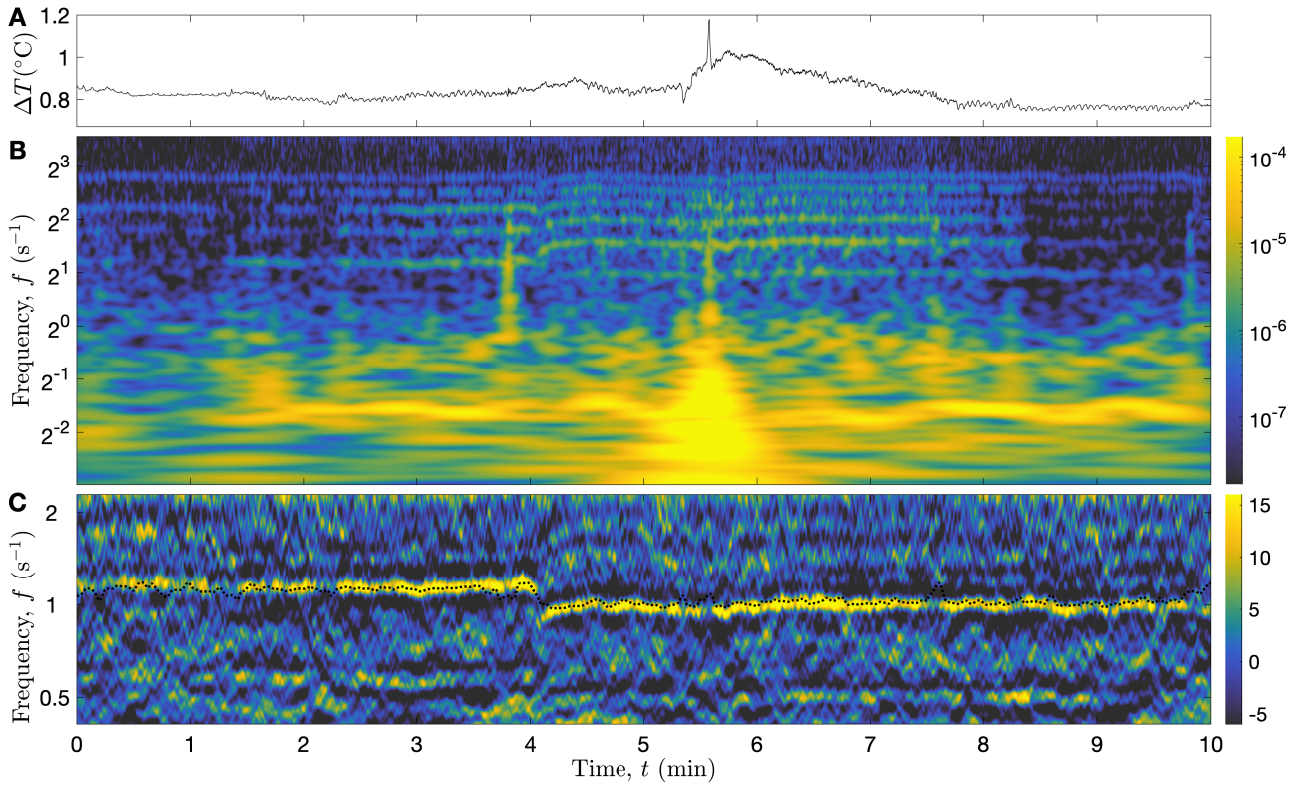


Figure B.23: *Heart rate detection at high quality factor from a cepstral strategy. (A) Temperature signal $x(t)$ (°C) derived from the dynamic infrared thermogram of subject 10. (B) Power log-frequency density $S_{xx}(t, f; Q_+, Q_-)$ estimated at high quality factor ($Q_- = 20$) and from a short time smoothing ($Q_+ \approx Q_-$). (C) Wavelet-based cepstrum defined in Eq.*(B.13)*, where the inverse of the quefrency has been identified with a frequency $f = \tau^{-1}$. The resulting cardiac line is compared to a narrow-band wavelet estimator applied to broad-band amplitude signal, $2|X(t, f; Q)|$ for $f = 4$ Hz and $Q = 2$, that collects heart beats.*

quantity:

$$\mathcal{C}_x(t,\tau) = -\int_0^\infty \frac{\partial \log(S_{xx}(f,t;Q_+,Q_-)|f|)}{\partial \log f} \sin(\Upsilon f\tau) \mathrm{d}\log f \quad , \tag{B.13}$$

The use of the power log-frequency density instead of the power spectral density does not matter, due to the derivation. However, the additional smoothing from $Q_+ > Q_-$ is important to avoid divergences (Q_+ only needs to be slightly larger than Q_-, in order not to lose too much time resolution).

The result $\mathcal{C}_x(t,\tau)$ is represented in Fig.B.23 (C), where we have set $f = \tau^{-1}$ for a time-frequency interpretation. Although insufficient to extract a cardiac FM signal, the heart rate can be read at any time as the fundamental line of the cardiac rhythm. In particular, we observe a sudden drop of the heart rate at 4 min. The heart rate estimator obtained from the low quality factor approach follows quite well this line of fundamental cardiac frequency. An additional step would be required to extract the heart rate signal from the cepstrum.

Even though a time-quefrency analysis is here satisfyingly adapted to the wavelet transform, the Gabor paradigm has many characteristics that makes it preferable to the wavelet paradigm: its homogeneous resolution in linear frequency is a clear advantage to detect the harmonic series, whose order is unlimited, even at a rather small time resolution. This is confirmed by recent applications to physiological signals combining Gabor transform and cepstrum [Li et al. 2019; Lin, Su, and Wu 2018; Su and Wu 2017].

Finally, notice that the rate estimator not only extracts a heart rate signal, but can also

be given an uncertainty from the formalism described in Eq.(III.37).

3 Cardio-respiratory coherence

We give here further details on the estimated cardio-respiratory coherence. In particular, the figures below illustrate the alternative result when varying the choice for the heart rate estimator. The results are very similar, up to a slight decrease or increase in estimated coherence, showing that the most generic (wide-band) estimator is sufficient for a coherence analysis.

3.1 Phase distribution: verification with a precise heart rate detection

We show here that a more precise estimation of the physiological rates does not improve much the measure of the cardio-respiratory coherence. More specifically, the coherence-based distribution of the phase $G_{xy}(\phi\,;f)$ between heart rate (x) and breathing rates (y) shown in Fig.B.24 is computed with the precise wavelet estimator of the heart rate $\Re\{F_{aECG14}(t,f_c;Q)\}$, as in the first section of chapter IV.

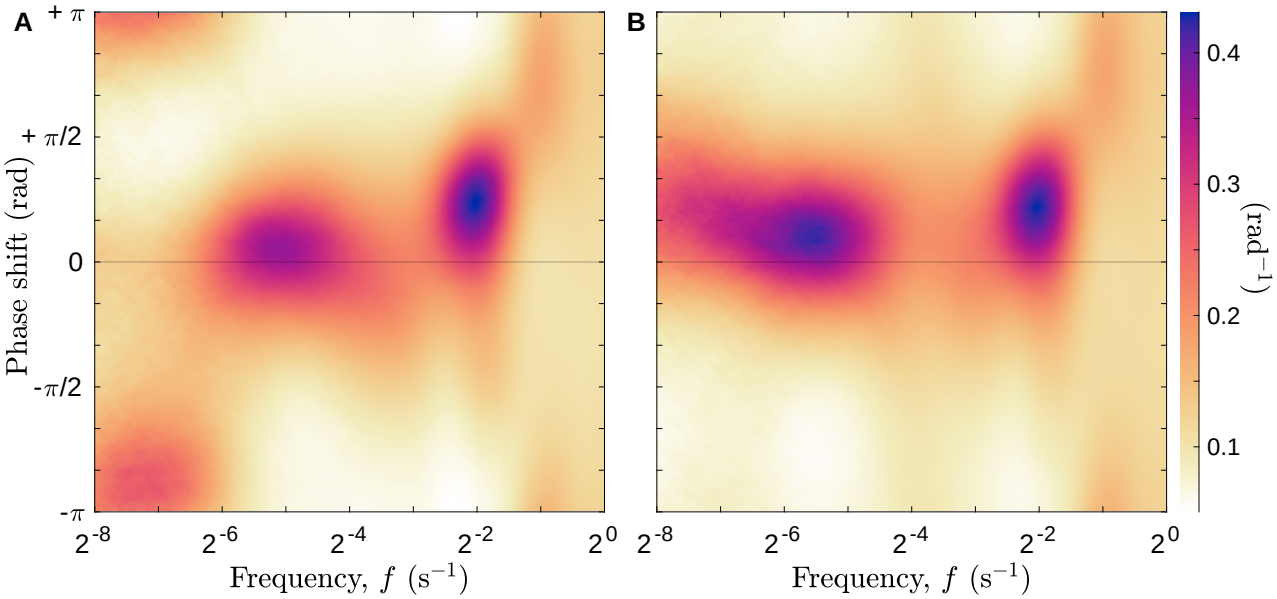


Figure B.24: *Comparison of the phase distributions $G_{xy}(\phi\,;f)$ in shhs2 between heart rate and breathing rates modulations. (A) Cardiac frequency ($\Re\{F_{aECG14}(t,f_c;Q)\}$) vs. respiratory frequency ($F_{AF}(t;Q)$). (B) Cardiac frequency vs. respiratory amplitude ($i\Im\{F_{AF}(t;Q)\}$).*

Compared to Fig.IV.21, that is computed from generic estimators, the coherence in Fig.B.24 is enhanced in both panels (of up to 30% at the density peaks), while the angular spread remains identical. This does not affect the previous identification of the fast, slow and very slow coherent components.

Since the precise heart rate estimators attenuates the oscillations at the cardiac frequency (strong for the generic estimator), the only important difference in the resulting cardio-respiratory coherence lies near the heart frequency (1 Hz). We do not interpret breathing modulations in this frequency range since it exceeds the respiratory frequency (0.25 Hz).

3.2 Apnea profiles: comparison with a generic heart rate estimator

In Fig.B.25, we show the statistical analysis in the *shhs2* of the cardio-respiratory coherence for different apnea profiles for a generic cardiac frequency modulation estimator, $\Re\{F_{ECG}(t,Q)\}$, instead of the precise and sophisticated narrow-band estimator $\Re\{F_{aECG14}(t,f_c;Q_+,Q_-)\}$, as in Fig.V.14.

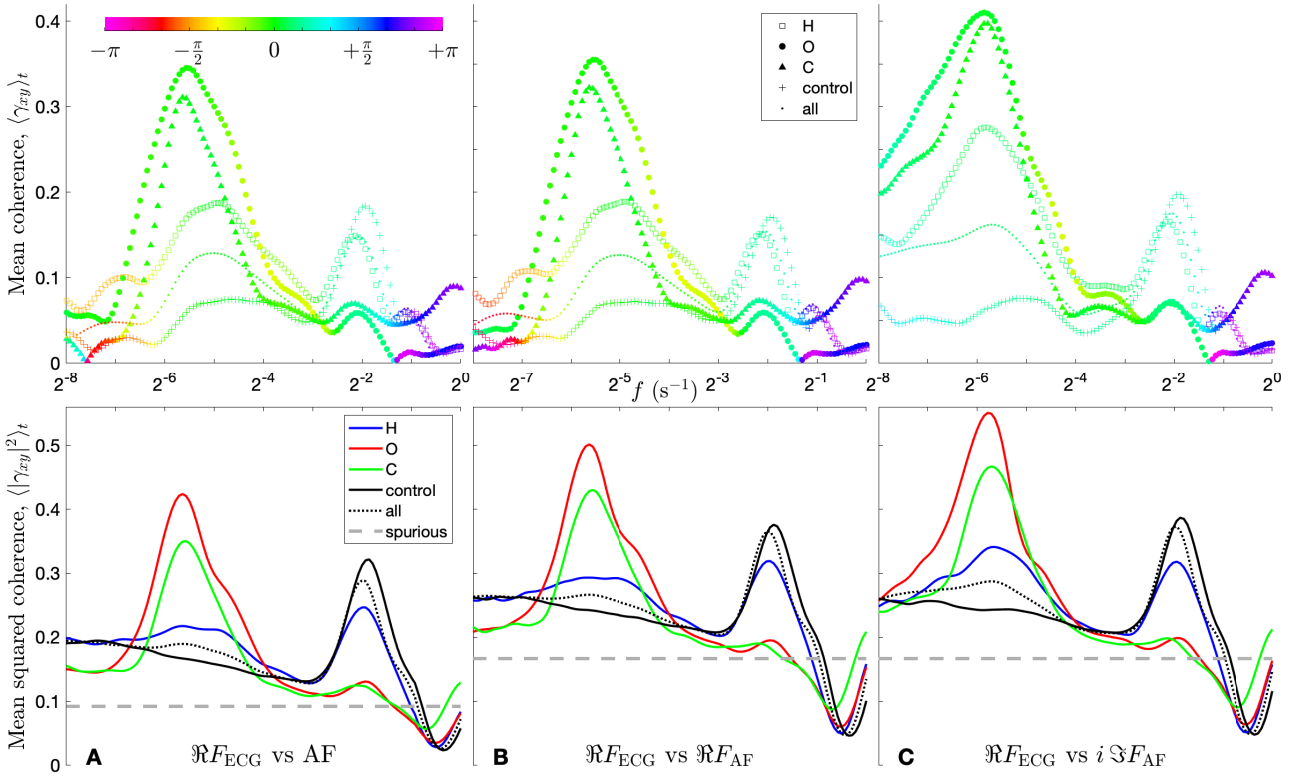


Figure B.25: *Average profiles of the coherence (top) and squared coherence (bottom) between the heart rate and different respiratory signals. Cardiac frequency (generic estimator)* $\Re\{F_{ECG}\}$ *versus: (A) airflow recording AF, (B) breathing frequency* $\Re\{F_{AF}\}$ *and (C) amplitude modulation* $i\Im\{F_{AF}\}$*, generic estimators. Each profile correspond to a conditional time average among subjects strongly affected by hypopnea "H", obstructive "O" or central "C" sleep apnea, without conditioning "all" or in the "control" group (without sleep apnea). See legends and text for details. The grey thick dashed line traces the expected level of spurious squared coherence* $\gamma^2_{\rm sp} \approx \frac{1}{n+1}$*, where* $n = 10$ *in (A) and* $n = 5$ *in (B, C).*

Apart from the slight decrease in coherence caused by the lower precision of the cardiac FM, the only important difference lies near the heart frequency (1 Hz) that are not interpreted as breathing modulations. As previously, this reflects the intense cardiac modulation that remains in the generic cardiac FM estimator due to the strong non-circularity of the ECG. The depletion of the coherence below the spurious level is caused by the accumulation of phase vortices (and amplitude zeros) in the neighbourhood of any intense mode.

4 Wake-sleep stage durations and transitions

We use the clinician scoring of wake-sleep stages for the 2650 subjects in the *shhs2* database to provide basic characteristics of sleep dynamics. The average sleep duration (from falling asleep to last waking) is 455 min, composed of wake (W) at 18%, REM at 17%, N1 at 5%, N2 at 48% and N3 at 13%. Their typical durations, between The important proportion of wake is certainly

representative of the sleep perturbation caused by the polysomnography apparatus. Denoting $X_n \in \{W, REM, N1, N2, N3\}$ the stage at step n and neglecting timings $t(n-1 \to n)$, we already can get an idea of the network of transitions between states by estimating transition probabilities. In Fig.B.26, we represent matrices for the probability of the previous stage knowing the current one $\Pr(X_{n-1}|X_n)$ (left), and the probability of the next stage knowing the current one $\Pr(X_{n+1}|X_n)$ (right). The distribution of the previous and next stages on a line of the matrices is normalized.

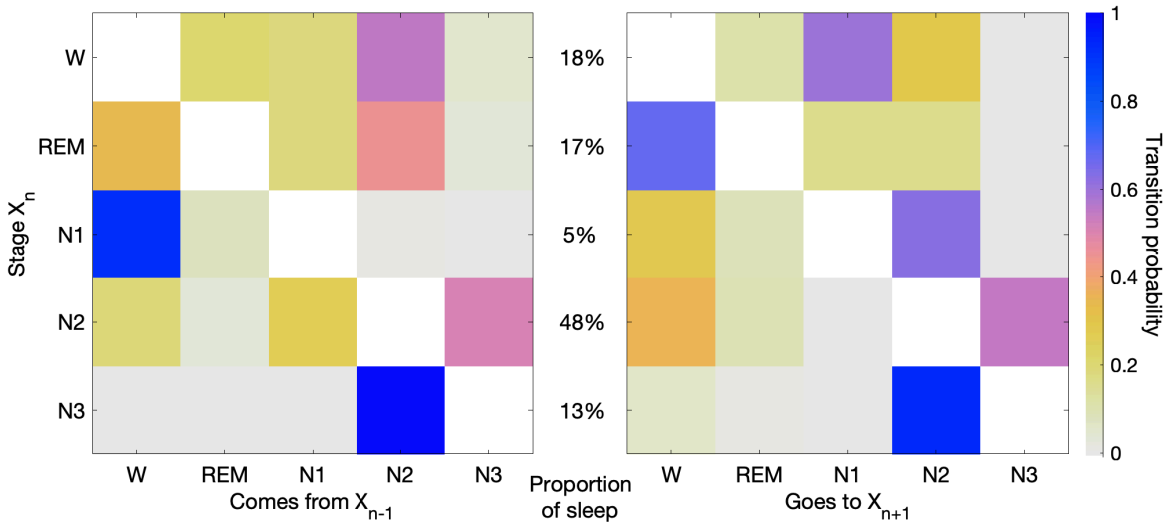


Figure B.26: *First order transition probabilities* $\Pr(X_{n-1}|X_n)$ *(left) and* $\Pr(X_{n+1}|X_n)$ *(right) between wake-sleep stages estimated form annotations in the shhs2 database. The proportion of each stage in the total sleep duration is also reported (middle).*

For instance, we know from the line associated to the N1 stage (light sleep), that it is essentially preceded by the wake stage (left), and most probably followed by the N2 stage, or less probably by wake (sometimes by REM), but it constitutes only 5% of the sleep duration (middle). We also observe that the N3 stage (deep sleep) is almost only connected to N2, compatible with our observations that N3 is a deeper *degree* of non-REM sleep, not clearly separated from N2.

Conclusion

An analytical tool to assist the observation of the dynamics of living organisms has been developed in this thesis, and applied to the study of interactions between the respiratory, cardiac and neural systems based on polysomnographic recordings. The development of this thesis is summarized below and perspectives are then outlined.

Summary

Part A

In the first part, we have exposed how to decipher complex natural recordings via the characterization of both their oscillating and fluctuating content. This is achieved by representing recordings in time and frequency, before performing statistical estimations in these two dimensions, via a combination of wavelet decomposition and recomposition.

We have started with the oscillatory approach of recordings, in particular with the temporal intuition associated to oscillating signals: a wave evolves in time with a certain amplitude and a phase cycling with a certain time period. The analytic version of the signal, a very suitable complex helical representation of an oscillation with its radius and angle, is of limited use when oscillations at multiple timescales are superimposed. From the analytic continuation of the signal in the complex plane, the separation of components at distinct scales naturally leads to the continuous wavelet transform with respect to an analytic wavelet, shifted and dilated in time. This time-scale decomposition has the specificity of being independent of any arbitrary reference scale, contrary to the alternative paradigm of the Gabor or short-time Fourier transform. This property is particularly interesting to analyse multiple simultaneous oscillatory events, *each from its own scale*, possibly separated from others by several decades.

Yet, the Fourier description of periodicity in the recording, from its spectrum, is indispensable; the wavelet scale variable only appears as a reminiscence of the frequency domain. In order to recover this complementary spectral intuition, we have revisited the choice of the analytic wavelet. Once its positive spectrum is standardized into a dimensionless shape, the precision of the scale-to-frequency correspondence narrows down to one essential parameter, its quality factor Q. Proportional to the number of wavelet oscillations and to the inverse of its bandwidth, it fixes a resolution trade-off, that shares the atom of time-frequency uncertainty. In other words, the quality factor is a cursor between the time and frequency interpretations. For the latter one, a finer control of the wavelet shape is superfluous, as long as the quality factor is high enough. The log-normal function, a central limit case in the general family of Morse wavelets, stands out as especially convenient for its regularity, with a faster than polynomial decay in both domains, and for its symmetry, crystallizing the idea of the logarithmic frequency scale.

Signal and spectrum are retrieved from the wavelet transform in two equivalent ways: on one hand as time and frequency marginals, on the other hand as opposite asymptotics for the quality factor. The probabilistic approach of fluctuations is then introduced, by considering the recording as a realization of a stochastic process. Applying the above operations to a quadratic representation results in wavelet-based statistical estimations: for a single trajectory, the time-dependent variance and spectral density, or for a product of trajectories, the correlation coefficient declined as a temporal and a spectral coherence. Their localization both in time and frequency, however, requires a second parameter to estimate the statistical average, that was identified with an effective number of orthogonal "tapers" (here wavelets), interpreted as statistical degrees of freedom. The alternative but equally effective approach consists in reducing the time-frequency resolution from a smoothing operation.

In the log-normal wavelet framework, these probabilistic and time-frequency perspectives

connect into a time-frequency statistical estimator that we qualify as canonical from its very special role. The introduction of several degrees of freedom in the statistical estimation coincides with the distinction of two different quality factors in the time-frequency analysis: $n_d = \frac{Q_+}{Q_-}$, the higher one controlling the time resolution, independently from the lower one for the log-frequency resolution. This correspondence is a bridge between two viewpoints on information, balancing significance and resolution. Following the first decomposition step of oscillations into single time-frequency atoms, the second step recomposes n_d uncertainty atoms for more stability against fluctuations.

We had to make two compromises, one on each side of the bridge, to make practical applications of these two essential parameters. The first one is the numerical implementation of the canonical estimator by a Gaussian time-smoothing, that is shown to be a correct Laplace approximation at first order. The second one is the simplified prediction of expected levels of statistical significance for the canonical coherence estimation. We motivate and verify form the simulation of stationary and independent Gaussian noises that the squared spurious coherence is nearly beta-distributed. The resulting power density and coherence are respectively a local variance and correlation coefficient detailed for all time shifts and scales. From the estimated levels of coherence significance, we construct a synthetic colour-coding from hue and saturation for the angle and modulus of significant time-frequency coherence regions.

The problem of estimating the modulations of a fluctuating rhythm is finally introduced. The repeated non-circular waveform is decomposed by the wavelet transform into horizontal and vertical time-frequency structures, consistently interpreted as harmonics up to a certain order $\sim \frac{Q}{\sqrt{\uparrow}}$ and beats above them. The cycling phase (and varying amplitude) of harmonic lines, in particular the one of the first (fundamental) order, is related to the relevant rate by a time-derivative, which happens to be equivalent to a frequency scaling. A certain flexibility exists in the spectral selection of the cycling phase, in relation to the concepts of stable and unstable horizontal ridges and frequency-reassignment, so that we posit a relevant frequency band. Modulations in this band can be extracted from a generic family of rate estimators. In particular, a quadratic version is compliant with the statistical interpretation: the introduction of a second quality factor (in relation to the bandwidth) ensures the stability of the estimated rate. For a voice recording, the different outcomes, in the narrow and wide-band limits, of this wavelet-estimator of a third type, are estimating the pitch and the vibrato. The uncertainty on the pitch estimation is expressed in terms of coherence.

Part B

The second part is applied to the analysis of the cardiac, respiratory and neural activities, as recorded by the ECG or PPG, AF and EEG signals. Physiological data, polysomnography for most of them, originate from existing databases. We endeavour to translate their processing into biophysical observables of interest using the language developed in the first part.

We start with the heart rate variability, extracted from the cardiac rhythm as its frequency modulation. In spite of the very different recording techniques, the heart rate can be estimated precisely with generic time-frequency estimators of the fluctuating rate in a relevant frequency band. We show that the wide-band version of the rate estimator, that only requires a single parameter, is sufficient to extract modulations of the cardiac frequency, even though it is biased, especially for strongly non-circular ECG oscillations. The most precise heart rate estimation is obtained by applying the narrow-band rate estimator to a new amplitude signal, derived from the wavelet transform of the ECG signal at 14 Hz (aECG14). Estimators are applied to the PPG, ECG and aECG14 signals of all subjects in a database. The precision of the cardiac frequency modulations is evaluated in terms of coherence with a reference heart rate estimated from R-R intervals in the ECG. This coherence is significant in all cases and at each

modulation frequency, and it increases with the circularity of the input cardiac signal and with the specificity of the estimator.

Sign of the cardio-respiratory interaction, the heart rate variability is correlated to breathing, introduced as the AF (airflow) signal. Comparing it to the heart rate, first in the time or frequency domains separately, then in the time-frequency plane, we observe two distinct components of coherence: a fast one at the fundamental frequency of the breathing rhythm, known as the respiratory sinus arrhythmia (RSA), and a slower and less rhythmic one (with a broader band). These interactions are characterized by their phase and modulus of coherence, both individually and collectively. The statistical significance is assessed for the individual time-frequency estimation, and the collective estimation is computed from a second level of statistics over all times and for all subjects of a large database. In particular, the extent of fluctuations in the phase difference between heart rate and breathing oscillations is summarized into an angular density of coherence. Reproducing this method with another type of respiratory signal and comparing the results, we have highlighted difficulties of interpretation related to instrumental limitations, visible in the reduction of coherence and the uncertainty on the phase of slow oscillations caused by high-pass filtering.

The study of the low frequency range is facilitated with breathing rate signals, instead of a raw recording, by extracting the slow activity from its frequency modulations (FM) and amplitude modulations (AM). Proved sufficient to study coherence for the cardiac rhythm, the generic wide-band rate estimator is particularly appropriate for the rather circular oscillations of breathing, whose period can fluctuate greatly. The coherence analysis highlights the fact that slow breathing FM and AM evolve nearly in-phase (positively correlated) with the heart rate modulations (FM); coherence extends to very slow frequencies with a slight phase shift for breathing AM, but with a phase switch to opposition (anti-correlation) for breathing FM below 0.01 Hz. This illustrates the presence of a distinct very slow mode of cardio-respiratory regulation, antagonist to the slow one.

The time-frequency coherence analysis of cardiac and respiratory FM is limited to depicting the similarity of their coevolution. In particular, it cannot account for their synchronization at an integer frequency ratio, nor characterize the evolution of this so-called pulse respiration quotient. In the wavelet framework, we construct a time-dependent distribution of this frequency ratio, enhanced to a very high time resolution and precise spectral concentration with frequency-reassignment and a low quality factor. Applied to polysomnography, we have illustrated individually its ability to track sleep phenomena such as apnea as regular ratio oscillations, REM with its strong fluctuations, and deep sleep with a steady ratio.

The observation of clear synchronizations episodes, also called cardio-respiratory coupling, and the musical analogy with two voices singing a consonant chord have motivated the formulation of a multi-frequency synchronization index. Called the sonance, it measures rationality or irrationality in the ratio distribution. Its development, currently limited to a musical application, is detailed more extensively in a distinct bonus part.

Important third member of the physiological orchestra, the activity of the neural system is introduced as recorded from EEG in a polysomnography. A constant quality factor representation is especially well suited to observe its multi-scale band structure and temporal stages in relation to wake and sleep. The selection of the unique or the pair of quality factors is discussed depending on the investigated phenomenon. $Q_- = 5$ is a minimum for spectral distinction of neural waves, fixing a maximum time-resolution, that can be reduced, either slightly ($Q_+ \approx Q_-$) to define a stable neural magnitude field, or drastically to obtain a synthetic representation of the sleep-wake density patterns, or a very significant estimation of the time-frequency coherence. For $Q_+ = 500$, we illustrate inter-EEG coherence patterns, between contra-lateral recordings, with alternating sign and modulus (incoherent to very coherent), varying independently from the power density. Rates modulations relate to power and magnitude modulations in each EEG

band; the latter are preferred as neural observables.

In order to evaluate physiological interactions, we gather cardio-respiratory rates and neural magnitude signals in distinct bands, and we compare them in pairs as a time-frequency coherence. Applied in the case of a subject affected by a severe obstructive sleep apnea, we identify the signature of apnea as a narrow-band and very coherent component present in all pairs of modulations between 0.01 and 0.04 Hz. An extensive coherence analysis of neuro-cardio-respiratory modulations is performed individually for each of the 2650 subjects in a large polysomnography database. Then, typical coherence spectra are computed from conditional time averages over the 2.3 years or cumulated sleep recordings, by dividing it into selected intervals and subgroups, based on the clinician annotations of apnea events of the different types: hypopnea, obstructive or central apnea, or their absence. The apneic modulations are characterized by an increased amplitude and coherence, appearing as spectral peaks compared to the flat baseline without apnea and the intermediary hypopnea profile. Apnea is also concomitant with an important coherence reduction, not only in the fast and narrow-band RSA cardio-respiratory interaction, but also for the very slow and broad-band (irregular) neuro-respiratory and cardio-respiratory component.

The phase of neural modulations compared to cardio-respiratory ones is then investigated across EEG bands. We stumble upon a specific phase-frequency relation that is essentially conserved, up to a global phase shift, for all profiles (with and without apnea), slow and very slow modulations and for all neuro-cardiac and neuro-respiratory coherence pairs. This interaction is composed of a linear chirp in the δ band, incoherence in the θ band, an echo of the chirp in the α-σ band and a constant phase in the β-γ band.

With an individual polysomnography, we eventually illustrate phase-amplitude couplings in different sleep stages, described as an intra-EEG coherence, in particular the one between slow δ oscillations and sleep spindles of the σ band during deep sleep. We also present a strategy to detect sleep spindles from the phase in the σ band. The instantaneous heart-breath ratio (or pulse-respiration quotient) completes the nocturnal picture of the sequence of sleep stages.

Discussion and perspectives

It is now clear that a description of the dynamics of living systems must be able to discriminate a great diversity of behaviours. The composition of the spontaneous physiological noise ranges from completely stochastic fluctuations, to pulsatile, aperiodic, quasi-periodic rhythms and regular oscillations, that can be concomitant at distinct frequencies, and evolve in a sequence of temporal stages of varying duration. This rich phenomenology is not only distributed among simultaneous recordings, but also within each of them in the form of modulations.

The practice of physics, in this interdisciplinary subject, has consisted in ensuring the visibility, control and soundness of the elementary notions of quality factor, frequency, time, phase and amplitude of the oscillations. In order to build a statistical physics of these physiological signals, a time-frequency perspective is indispensable: these natural oscillating phenomena are correctly described when their intrinsic quality factor is matched by the one of the representation, achieving their localization. Decoupling time from frequency corresponds to dissociating their consistent description into two distorted corner views. The adoption of the adequate middle view amounts to choosing the shape of the uncertainty atom, or wavelet, in the time-frequency decomposition.

The basic probabilistic tool of pair-wise correlations corresponds to the recomposition of atoms from two distinct signals in a local average. The resulting circular (not to say complex) coefficient expresses their coherence. This statistical information is as precise as the atoms are numerous, to the detriment of the localization information, that is reduced accordingly. The probabilistic interpretation is therefore enabled by the separation of slow evolutionary scales and

fast spectral scales. Different paradigms are available to balance these two types of information, each leading to a specific geometry of the composition in the time-frequency plane. Without introducing any arbitrary reference scales, the formalism of the log-normal wavelet proposed in this thesis keeps track of the shape and number of atoms in the composition as a ratio of two quality factors.

All pairs of physiological signals form a complete graph, whose links are time-frequency coherence maps; they would reduce to (circular) correlation coefficients when the information is purely statistical and non-local. The specification of two control parameters (the quality factors) is sufficient to derive from these signals a prototype of the network of physiological interactions, whose structure is being investigated [Bartsch, Liu, Bashan, et al. 2015; Ivanov, Liu, and Bartsch 2016]. Our synthetic colour-coded visualization processes for these time-frequency relations will assist observations, and may be appropriately reduced into principal components.

We expect a fully developed physical approach to specify a directionality in the interactions. Namely, causality must be excavated from this network of interactions, to bring to light the sharpness of the inter-regulation mechanisms between physiological functions. Reduced to its simplest form as a pure delay $e^{i\tau f \tau}$, causality is contained in the spectral direction as a cycling phase. An efficient spectral retrieval of causality becomes possible [Ephremidze, Saied, and Spitkovsky 2018] and its statistical interpretation is well-supported [Faes, Pernice, et al. 2021; Faes, Stramaglia, and Marinazzo 2017]. The time-evolving and multi-scale characterization of causal interactions in the physiological orchestra is close to be achieved.

The next step towards a statistical physics of physiological signals is farther away, but the question of their irreversibility is in sight [Arneodo, Argoul, et al. 1993] and the ground may already be in preparation [Sekimoto 2010].

References

Abreu, L. D. and M. Speckbacher (Dec. 2020). *Affine Density and von Neumann Dimension*. arXiv: 2012.11314.

Abry, P., P. Gonçalvès, and P. Flandrin (1995). "Wavelets, Spectrum Analysis and 1/f Processes". In: *Wavelets and Statistics*. Ed. by P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg, I. Olkin, N. Wermuth, S. Zeger, A. Antoniadis, and G. Oppenheim. Vol. 103. New York, NY: Springer New York, pp. 15–29. ISBN: 978-0-387-94564-4 978-1-4612-2544-7.

Aguiar-Conraria, L., M. Soares, and R. Sousa (Aug. 2018). "California's Carbon Market and Energy Prices: A Wavelet Analysis". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376.2126, p. 20170256. DOI: 10/gjvpmn.

Aguiar-Conraria, L. and M. J. Soares (Apr. 2014). "The Continuous Wavelet Transform: Moving Beyond Uni- and Bivariate Analysis: The Continuous Wavelet Transform". In: *Journal of Economic Surveys* 28.2, pp. 344–375. DOI: 10/ghm2c8.

Alkauskas, G. (2008). "Integral Transforms of the Minkowski Question Mark Function". PhD thesis. University of Nottingham.

Allegrini, P., D. Menicucci, R. Bedini, L. Fronzoni, A. Gemignani, P. Grigolini, B. J. West, and P. Paradisi (2009). "Spontaneous Brain Activity as a Source of Ideal 1/f Noise". In: *Physical Review E* 80.6, p. 061914. DOI: 10/cgvbpf.

Altes, R. A. (Jan. 1976). "Sonar for Generalized Target Description and Its Similarity to Animal Echolocation Systems". In: *The Journal of the Acoustical Society of America* 59.1, pp. 97–105. DOI: 10/d2c3bb.

Argoul, F., A. Arneodo, J. Elezgaray, G. Grasseau, and R. Murenzi (May 1990). "Wavelet Analysis of the Self-Similarity of Diffusion-Limited Aggregates and Electrodeposition Clusters". In: *Physical Review A* 41.10, pp. 5537–5560. DOI: 10/dgjfdg.

Argoul, F., A. Arneodo, J. Elezgaray, G. Grasseau, and R. Murenzi (1989). "Wavelet Transform of Fractal Aggregates". In: *Physics Letters A* 135.6-7, pp. 327–336. DOI: 10/chtmm4.

Argoul, F., A. Arneodo, G. Grasseau, Y. Gagne, E. J. Hopfinger, and U. Frisch (1989). "Wavelet Analysis of Turbulence Reveals the Multifractal Nature of the Richardson Cascade". In: *Nature* 338.6210, pp. 51–53. DOI: 10/d3rr9n.

Arneodo, A., F. Argoul, J. Elezgaray, and P. Richetti (1993). "Homoclinic Chaos in Chemical Systems". In: *Physica D: Nonlinear Phenomena* 62.1-4, pp. 134–169. DOI: 10/cgrwq4.

Arneodo, A., B. Audit, E. Bacry, S. Manneville, J. Muzy, and S. Roux (1997). "Scale Invariance and Beyond: What Can We Learn from Wavelet Analysis?" In: *Scale Invariance and Beyond*. Springer, pp. 37–51.

— (1998). "Thermodynamics of Fractal Signals Based on Wavelet Analysis: Application to Fully Developed Turbulence Data and DNA Sequences". In: *Physica A: Statistical Mechanics and its Applications* 254.1-2, pp. 24–45. DOI: 10/fh35pw.

Arneodo, A., E. Bacry, S. Manneville, and J. Muzy (1998). "Analysis of Random Cascades Using Space-Scale Correlation Functions". In: *Physical review letters* 80.4, p. 708. DOI: 10/c6hrqj.

Arneodo, A., S. Manneville, and J. Muzy (1998). "Towards Log-Normal Statistics in High Reynolds Number Turbulence". In: *The European Physical Journal B-Condensed Matter and Complex Systems* 1.1, pp. 129–140. DOI: 10/chc3db.

Arneodo, A., G. Grasseau, and M. Holschneider (Nov. 1988). "Wavelet Transform of Multifractals". In: *Physical Review Letters* 61.20, pp. 2281–2284. DOI: 10/djzv8v.

Arneodo, A., E. Bacry, and J. Muzy (1995). "The Thermodynamics of Fractals Revisited with Wavelets". In: *Physica A: Statistical Mechanics and its Applications* 213.1-2, pp. 232–275. DOI: 10/fbjw6w.

Asmussen, S., J. L. Jensen, and L. Rojas-Nandayapa (June 2016). "On the Laplace Transform of the Lognormal Distribution". In: *Methodology and Computing in Applied Probability* 18.2, pp. 441–458. DOI: 10/gjvnrb.

Auger, F. and P. Flandrin (1994). "The Why and How of Time-Frequency Reassignment". In: *Proceedings of IEEE-SP International Symposium on Time- Frequency and Time-Scale Analysis*. Philadelphia, PA, USA: IEEE, pp. 197–200. ISBN: 978-0-7803-2127-4. DOI: 10/cfzx9n.

Auger, F., P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu (Nov. 2013). "Time-Frequency Reassignment and Synchrosqueezing: An Overview". In: *IEEE Signal Processing Magazine* 30.6, pp. 32–41. DOI: 10/gftgnq.

Auger, F. and P. Flandrin (1995). "Improving the Readability of Time-Frequency and Time-Scale Representations by the Reassignment Method". In: *IEEE Transactions on signal processing* 43.5, pp. 1068–1089. DOI: 10/dpmrhp.

Aviyente, S., E. M. Bernat, W. S. Evans, and S. R. Sponheim (Jan. 2011). "A Phase Synchrony Measure for Quantifying Dynamic Functional Integration in the Brain". In: *Human Brain Mapping* 32.1, pp. 80–93. DOI: 10/bwknwt.

Aviyente, S. and A. Y. Mutlu (July 2011). "A Time-Frequency-Based Approach to Phase and Phase Synchrony Estimation". In: *IEEE Transactions on Signal Processing* 59.7, pp. 3086–3098. DOI: 10/fmkn9t.

Aviyente, S., A. Tootell, and E. M. Bernat (Jan. 2017). "Time-Frequency Phase-Synchrony Approaches with ERPs". In: *International Journal of Psychophysiology* 111, pp. 88–97. DOI: 10/f9phvp.

Babadi, B. and E. N. Brown (May 2014). "A Review of Multitaper Spectral Analysis". In: *IEEE Transactions on Biomedical Engineering* 61.5, pp. 1555–1564. DOI: 10/gdm2rk.

Babenko, K. I. (1961). "An Inequality in the Theory of Fourier Integrals". In: *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya* 25.4, pp. 531–542.

Babiloni, F., F. Cincotti, F. Carducci, P. M. Rossini, and C. Babiloni (May 2001). "Spatial Enhancement of EEG Data by Surface Laplacian Estimation: The Use of Magnetic Resonance Imaging-Based Head Models". In: *Clinical Neurophysiology* 112.5, pp. 724–727. DOI: 10/frsj26.

Bansal, D., M. Khan, and A. Salhan (Mar. 2009). "A Review of Measurement and Analysis of Heart Rate Variability". In: *2009 International Conference on Computer and Automation Engineering*. Bangkok: IEEE, pp. 243–246. ISBN: 978-0-7695-3569-2. DOI: 10/dg7fmb.

Baraniuk, R. G., P. Flandrin, A. J. Janssen, and O. J. Michel (2001). "Measuring Time-Frequency Information Content Using the Rényi Entropies". In: *IEEE Transactions on Information theory* 47.4, pp. 1391–1409. DOI: 10/dcsvwh.

Bardenet, R., J. Flamant, and P. Chainais (Mar. 2020). "On the Zeros of the Spectrogram of White Noise". In: *Applied and Computational Harmonic Analysis* 48.2, pp. 682–705. DOI: 10/gjvnrw.

Bardenet, R. and A. Hardy (Jan. 2021). "Time-Frequency Transforms of White Noises and Gaussian Analytic Functions". In: *Applied and Computational Harmonic Analysis* 50, pp. 73–104. DOI: 10/gj8mts.

Bartsch, R. P., K. K. Liu, A. Bashan, and P. C. Ivanov (2015). "Network Physiology: How Organ Systems Dynamically Interact". In: *PloS one* 10.11, e0142143. DOI: 10/f8b4vm.

Bartsch, R. P., K. K. Liu, Q. D. Ma, and P. C. Ivanov (2014). "Three Independent Forms of Cardio-Respiratory Coupling: Transitions across Sleep Stages". In: *Computing in Cardiology 2014*. IEEE, pp. 781–784. ISBN: 1-4799-4347-9.

Bartsch, R. P., A. Y. Schumann, J. W. Kantelhardt, T. Penzel, and P. C. Ivanov (2012). "Phase Transitions in Physiologic Coupling". In: *Proceedings of the National Academy of Sciences* 109.26, pp. 10181–10186. DOI: 10/f334c9.

Bashan, A., R. P. Bartsch, J. W. Kantelhardt, S. Havlin, and P. C. Ivanov (Jan. 2012). "Network Physiology Reveals Relations between Network Topology and Physiological Function". In: *Nature Communications* 3.1, p. 702. DOI: 10/f4f8v7.

Bayram, M. and R. Baraniuk (1996). "Multiple Window Time-Frequency Analysis". In: *Proceedings of Third International Symposium on Time-Frequency and Time-Scale Analysis (TFTS-96)*. Paris, France: IEEE, pp. 173–176. ISBN: 978-0-7803-3512-7. DOI: 10/dzsqxh.

Bayram, M. and R. G. Baraniuk (2000). "Multiple Window Time Varying Spectrum Estimation". In: *Cambridge University Press*.

Beckner, W. (July 1975). "Inequalities in Fourier Analysis". In: *The Annals of Mathematics* 102.1, p. 159. DOI: 10/csps7s.

Bedrosian, E. (1963). "A Product Theorem for Hilbert Transforms". In: *Proceedings of the IEEE* 51.5, pp. 868–869. DOI: 10/dmdfwm.

Behar, J., A. Johnson, G. D. Clifford, and J. Oster (2014). "A Comparison of Single Channel Fetal ECG Extraction Methods". In: *Annals of biomedical engineering* 42.6, pp. 1340–1353. DOI: 10/f524wq.

Borgnat, P., P. Flandrin, and P.-O. Amblard (2002). "Generalized Lamperti Transformation of Broken Scale Invariance". In: *Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, 2002.* Vol. 2. Pacific Grove, CA, USA: IEEE, pp. 1576–1580. ISBN: 978-0-7803-7576-5. DOI: 10/cvpbc6.

Brittain, J.-S., D. Halliday, B. Conway, and Jens Bo Nielsen (May 2007). "Single-Trial Multiwavelet Coherence in Application to Neurophysiological Time Series". In: *IEEE Transactions on Biomedical Engineering* 54.5, pp. 854–862. DOI: 10/cww5w6.

Burle, B., L. Spieser, C. Roger, L. Casini, T. Hasbroucq, and F. Vidal (Sept. 2015). "Spatial and Temporal Resolutions of EEG: Is It Really Black and White? A Scalp Current Density View". In: *International Journal of Psychophysiology* 97.3, pp. 210–220. DOI: 10/f7qpr6.

Buzsáki, G. and A. Draguhn (2004). "Neuronal Oscillations in Cortical Networks". In: *science* 304.5679, pp. 1926–1929. DOI: 10/bwtf9q.

Calkin, N. and H. S. Wilf (2000). "Recounting the Rationals". In: *The American Mathematical Monthly* 107.4, pp. 360–363. DOI: 10/gph8gk.

Carmona, R., W. Hwang, and B. Torrésani (Oct. 1997). "Characterization of Signals by the Ridges of Their Wavelet Transforms". In: *IEEE Transactions on Signal Processing* 45.10, pp. 2586–2590. DOI: 10/cdnsc8.

Carmona, R., W. L. Hwang, and B. Torrésani (1995). "Identification of Chirps with Continuous Wavelet Transform". In: *Wavelets and Statistics*. Ed. by P. Bickel, P. Diggle, S. Fienberg, K. Krickeberg, I. Olkin, N. Wermuth, S. Zeger, A. Antoniadis, and G. Oppenheim. Vol. 103. New York, NY: Springer New York, pp. 95–108. ISBN: 978-0-387-94564-4 978-1-4612-2544-7.

Carmona, R., W.-L. Hwang, and B. Torrésani (1998). *Practical Time-Frequency Analysis: Gabor and Wavelet Transforms, with an Implementation in S*. Academic Press. ISBN: 0-08-053942-4.

Carter, G. (1987). "Coherence and Time Delay Estimation". In: *Proceedings of the IEEE* 75.2, pp. 236–255. DOI: 10/fptk6c.

Cazelles, B., M. Chavez, D. Berteaux, F. Ménard, J. O. Vik, S. Jenouvrier, and N. C. Stenseth (May 2008). "Wavelet Analysis of Ecological Time Series". In: *Oecologia* 156.2, pp. 287–304. DOI: 10/dmp6dc.

Chang, C. and G. H. Glover (Mar. 2010). "Time–Frequency Dynamics of Resting-State Brain Connectivity Measured with fMRI". In: *NeuroImage* 50.1, pp. 81–98. DOI: 10/ckdfb6.

Charlton, P. H., D. A. Birrenkott, T. Bonnici, M. A. F. Pimentel, A. E. W. Johnson, J. Alastruey, L. Tarassenko, P. J. Watkinson, R. Beale, and D. A. Clifton (2018). "Breathing Rate Estimation From the Electrocardiogram and Photoplethysmogram: A Review". In: *IEEE Reviews in Biomedical Engineering* 11, pp. 2–20. DOI: 10/gf3m44.

Chavez, M. and B. Cazelles (Dec. 2019). "Detecting Dynamic Spatial Correlation Patterns with Generalized Wavelet Coherence and Non-Stationary Surrogate Data". In: *Scientific Reports* 9.1, p. 7389. DOI: 10/gjvpmk.

Chevalier, G., K. Mori, and J. L. Oschman (2006). "The Effect of Earthing (Grounding) on Human Physiology". In: *European Biology and Bioelectromagnetics* 2.1, pp. 600–621.

Chui, C. K. (1992). *An Introduction to Wavelets*. Wavelet Analysis and Its Applications v. 1. Boston: Academic Press. ISBN: 978-0-12-174584-4.

Chui, C. K. and H. Mhaskar (Jan. 2016). "Signal Decomposition and Analysis via Extraction of Frequencies". In: *Applied and Computational Harmonic Analysis* 40.1, pp. 97–136. DOI: 10/f73xwr.

Cicone, A. and H.-T. Wu (Sept. 2017). "How Nonlinear-Type Time-Frequency Analysis Can Help in Sensing Instantaneous Heart Rate and Instantaneous Respiratory Rate from Photoplethysmography in a Reliable Way". In: *Frontiers in Physiology* 8, p. 701. DOI: 10/gjvnrx.

Claerbout, J. F. (1985). *Fundamentals of Geophysical Data Processing: With Applications to Petroleum Prospecting*. Palo Alto, CA: Blackwell Scientific Publications. ISBN: 978-0-86542-305-3.

Cohen, E. A. K. and A. T. Walden (June 2010a). "A Statistical Study of Temporally Smoothed Wavelet Coherence". In: *IEEE Transactions on Signal Processing* 58.6, pp. 2964–2973. DOI: 10/fcngnq.

— (Mar. 2010b). "A Statistical Analysis of Morse Wavelet Coherence". In: *IEEE Transactions on Signal Processing* 58.3, pp. 980–989. DOI: 10/bt572x.

Cohen, L. and C. Lee (1990). "Instantaneous Bandwidth for Signals and Spectrogram". In: *International Conference on Acoustics, Speech, and Signal Processing*. Albuquerque, NM, USA: IEEE, pp. 2451–2454. DOI: 10/c2wj9m.

Combes, J.-M., A. Grossmann, and P. Tchamitchian (1989). *Wavelets: Time-Frequency Methods and Phase Space Proceedings of the International Conference, Marseille, France, December 14–18, 1987*. Inverse Problems and Theoretical Imaging. Berlin Heidelberg: Springer. ISBN: 978-3-540-51159-5 978-0-387-51159-7.

Cuiwei Li, Chongxun Zheng, and Changfeng Tai (Jan. 1995). "Detection of ECG Characteristic Points Using Wavelet Transforms". In: *IEEE Transactions on Biomedical Engineering* 42.1, pp. 21–28. DOI: 10/bwd854.

Daly, E., C. Brown, C. P. Stark, and C. J. Ebinger (Nov. 2004). "Wavelet and Multitaper Coherence Methods for Assessing the Elastic Thickness of the Irish Atlantic Margin". In: *Geophysical Journal International* 159.2, pp. 445–459. DOI: 10/dv9wvs.

Daubechies, I. and T. Paul (Aug. 1988). "Time-Frequency Localisation Operators-a Geometric Phase Space Approach: II. The Use of Dilations". In: *Inverse Problems* 4.3, pp. 661–680. DOI: 10/d2b8cw.

Daubechies, I. (July 1988). "Time-Frequency Localization Operators: A Geometric Phase Space Approach". In: *IEEE Transactions on Information Theory* 34.4, pp. 605–612. DOI: 10/b873rs.

Daubechies, I. (Jan. 1992). *Ten Lectures on Wavelets*. CBMS-NSF Regional Conference Series in Applied Mathematics 61. Philadelphia, Pa: Society for Industrial and Applied Mathematics. ISBN: 978-0-89871-274-2 978-1-61197-010-4.

Daubechies, I., J. Lu, and H.-T. Wu (Mar. 2011). "Synchrosqueezed Wavelet Transforms: An Empirical Mode Decomposition-like Tool". In: *Applied and Computational Harmonic Analysis* 30.2, pp. 243–261. DOI: 10/dnnqvt.

Daubechies, I. and S. Maes (1996). "A Nonlinear Squeezing of the Continuous Wavelet Transform Based on Auditory Nerve Models". In: *Wavelets in Medicine and Biology*. Vol. 7. CRC press Boca Raton, FL.

de Geus, E. J. C., P. J. Gianaros, R. C. Brindle, J. R. Jennings, and G. G. Berntson (Feb. 2019). "Should Heart Rate Variability Be "Corrected" for Heart Rate? Biological, Quantitative, and Interpretive Considerations". In: *Psychophysiology* 56.2, e13287. DOI: 10/gfhfcc.

Delprat, N., B. Escudie, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrésani (Mar. 1992). "Asymptotic Wavelet and Gabor Analysis: Extraction of Instantaneous Frequencies". In: *IEEE Transactions on Information Theory* 38.2, pp. 644–664. DOI: 10/fsc36p.

Duplantier, B., R. Rhodes, S. Sheffield, and V. Vargas (2017). "Log-Correlated Gaussian Fields: An Overview". In: *Geometry, Analysis and Probability*. Ed. by J.-B. Bost, H. Hofer, F. Labourie, Y. Le Jan, X. Ma, and W. Zhang. Vol. 310. Cham: Springer International Publishing, pp. 191–216. ISBN: 978-3-319-49636-8 978-3-319-49638-2.

Dwyer, P. S. (Sept. 1940). "The Evaluation of Multiple and Partial Correlation Coefficients from the Factorial Matrix". In: *Psychometrika* 5.3, pp. 211–232. DOI: 10/bhqv7n.

Elgendi, M., I. Norton, M. Brearley, S. Dokos, D. Abbott, and D. Schuurmans (Sept. 2016). "A Pilot Study: Can Heart Rate Variability (HRV) Be Determined Using Short-Term Photoplethysmograms?" In: *F1000Research* 5, p. 2354. DOI: 10/gp7n6b.

Ephremidze, L., F. Saied, and I. M. Spitkovsky (Feb. 2018). "On the Algorithmization of Janashia-Lagvilava Matrix Spectral Factorization Method". In: *IEEE Transactions on Information Theory* 64.2, pp. 728–737. DOI: 10/gjvpdt.

Faes, L., G. Pinna, A. Porta, R. Maestri, and G. Nollo (July 2004). "Surrogate Data Analysis for Assessing the Significance of the Coherence Function". In: *IEEE Transactions on Biomedical Engineering* 51.7, pp. 1156–1166. DOI: 10/dtrzcj.

Faes, L., R. Pernice, G. Mijatovic, Y. Antonacci, J. C. Krohova, M. Javorka, and A. Porta (Dec. 2021). "Information Decomposition in the Frequency Domain: A New Framework to Study Cardiovascular and Cardiorespiratory Oscillations". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 379.2212, p. 20200250. DOI: 10/gnjdcd.

Faes, L., S. Stramaglia, and D. Marinazzo (Sept. 2017). "On the Interpretability and Computational Reliability of Frequency-Domain Granger Causality". In: *F1000Research* 6, p. 1710. DOI: 10/gckrp6.

Farre, R. (Dec. 2004). "Noninvasive Monitoring of Respiratory Mechanics during Sleep". In: *European Respiratory Journal* 24.6, pp. 1052–1060. DOI: 10/bb5cjh.

Farré, R., J. Montserrat, M. Rotger, E. Ballester, and D. Navajas (Jan. 1998). "Accuracy of Thermistors and Thermocouples as Flow-Measuring Devices for Detecting Hypopnoeas". In: *European Respiratory Journal* 11.1, pp. 179–182. DOI: 10/cnhpfv.

Ferree, T. C., P. Luu, G. S. Russell, and D. M. Tucker (Mar. 2001). "Scalp Electrode Impedance, Infection Risk, and EEG Data Quality". In: *Clinical Neurophysiology* 112.3, pp. 536–544. DOI: 10/dgd588.

Flandrin, P., F. Auger, and E. Chassande-Mottin (Oct. 2018). "Time-Frequency Reassignment: From Principles to Algorithms". In: *Applications in Time-Frequency Signal Processing*. First. CRC Press, pp. 179–204. ISBN: 978-1-315-22001-7.

Flandrin, P., R. Baraniuk, and O. Michel (1994). "Time-Frequency Complexity and Information". In: *Proceedings of ICASSP '94. IEEE International Conference on Acoustics, Speech and Signal Processing*. Vol. iii. Adelaide, SA, Australia: IEEE, pp. III/329–III/332. ISBN: 978-0-7803-1775-8. DOI: 10/bqmd3g.

Flandrin, P. (Sept. 1986). "On the Positivity of the Wigner-Ville Spectrum". In: *Signal Processing* 11.2, pp. 187–189. DOI: 10/d6r789.

— (1988). "Maximum Signal Energy Concentration in a Time-Frequency Domain". In: *ICASSP-88., International Conference on Acoustics, Speech, and Signal Processing*. IEEE Computer Society, pp. 2176, 2177, 2178, 2179–2176, 2177, 2178, 2179. ISBN: 1520-6149.

— (Aug. 1998a). "Separability, Positivity, and Minimum Uncertainty in Time–Frequency Energy Distributions". In: *Journal of Mathematical Physics* 39.8, pp. 4016–4040. DOI: 10/c77rzn.

— (1998b). *Time-Frequency/Time-Scale Analysis*. Burlington: Elsevier. ISBN: 978-0-08-054303-1.

— (2015). "Time–Frequency Filtering Based on Spectrogram Zeros". In: *IEEE Signal Processing Letters* 22.11, pp. 2137–2141. DOI: 10/gn675p.

Flandrin, P., P. Borgnat, and P.-O. Amblard (2003). "From Stationarity to Self-similarity, and Back: Variations on the Lamperti Transformation". In: *Processes with Long-Range Correlations*. Ed. by R. Beig, B. .-. Englert, U. Frisch, P. Hänggi, K. Hepp, W. Hillebrandt, D. Imboden, R. L. Jaffe, R. Lipowsky, H. v. Löhneysen, I. Ojima, D. Sornette, S. Theisen, W. Weise, J. Wess, J. Zittartz, G. Rangarajan, and M. Ding. Vol. 621. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 88–117. ISBN: 978-3-540-40129-2 978-3-540-44832-7.

Gabor, D. (Nov. 1946). "Theory of Communication. Part 1: The Analysis of Information". In: *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering* 93.26, pp. 429–441. DOI: 10/2rz.

Gardner, W. A. (Nov. 1992). "A Unifying View of Coherence in Signal Processing". In: *Signal Processing* 29.2, pp. 113–140. DOI: 10/d68fnp.

Gerasimova, E., B. Audit, S. G. Roux, A. Khalil, O. Gileva, F. Argoul, O. Naimark, and A. Arneodo (May 2014). "Wavelet-Based Multifractal Analysis of Dynamic Infrared Thermograms to Assist in Early Breast Cancer Diagnosis". In: *Frontiers in Physiology* 5. DOI: 10/gjvnrh.

Gerasimova-Chechkina, E., B. Toner, Z. Marin, B. Audit, S. G. Roux, F. Argoul, A. Khalil, O. Gileva, O. Naimark, and A. Arneodo (Aug. 2016). "Comparative Multifractal Analysis of Dynamic Infrared Thermograms and X-Ray Mammograms Enlightens Changes in the Environment of Malignant Tumors". In: *Frontiers in Physiology* 7. DOI: 10/gjvnrj.

Gish, H. and D. Cochran (1988). "Generalized Coherence". In: *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2745–2748.

Goldberger, A. L., L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley (2000). "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals". In: *circulation* 101.23, e215–e220. DOI: 10/ggwwsx.

Goodman, N. R. (Mar. 1963). "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)". In: *The Annals of Mathematical Statistics* 34.1, pp. 152–177. DOI: 10/cgm7kq.

Goupillaud, P., A. Grossmann, and J. Morlet (Oct. 1984). "Cycle-Octave and Related Transforms in Seismic Signal Analysis". In: *Geoexploration* 23.1, pp. 85–102. DOI: 10/fq6vcz.

Grigg-Damberger, M. M. (2012). "The AASM Scoring Manual Four Years Later". In: *Journal of Clinical Sleep Medicine* 8.3, pp. 323–332. DOI: 10/f998fk.

Grinsted, A., J. C. Moore, and S. Jevrejeva (Nov. 2004). "Application of the Cross Wavelet Transform and Wavelet Coherence to Geophysical Time Series". In: *Nonlinear Processes in Geophysics* 11.5/6, pp. 561–566. DOI: 10/d6dhzv.

Grossmann, A. and J. Morlet (July 1984). "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape". In: *SIAM Journal on Mathematical Analysis* 15.4, pp. 723–736. DOI: 10/b9ddn3.

Guillemain, P. and R. Kronland-Martinet (1992). "Horizontal and Vertical Ridges Associated to Continuous Wavelet Transforms". In: *[1992] Proceedings of the IEEE-SP International Symposium on Time-Frequency and Time-Scale Analysis*. Victoria, BC, Canada: IEEE, pp. 63–66. ISBN: 978-0-7803-0805-3. DOI: 10/fr6ctv.

Guillemain, P. and R. Kronland-Martinet (1996). "Characterization of Acoustic Signals through Continuous Linear Time-Frequency Representations". In: *Proceedings of the IEEE* 84.4, pp. 561–585. DOI: 10/bksgmw.

Gulrajani, R. (Sept.-Oct. 1998). "The Forward and Inverse Problems of Electrocardiography". In: *IEEE Engineering in Medicine and Biology Magazine* 17.5, pp. 84–101, 122. DOI: 10/b5x3gv.

Gurley, K., T. Kijewski, and A. Kareem (Feb. 2003). "First- and Higher-Order Correlation Detection Using Wavelet Transforms". In: *Journal of Engineering Mechanics* 129.2, pp. 188–201. DOI: 10/crqrgt.

Halliday, D., J. Rosenberg, A. Amjad, P. Breeze, B. Conway, and S. Farmer (1995). "A Framework for the Analysis of Mixed Time Series/Point Process Data-Theory and Application to the Study of Physiological Tremor, Single Motor Unit Discharges and Electromyograms". In: *Progress in biophysics and molecular biology* 64.2, p. 237.

Harremoës, P. (2012). *Al-Kāshī's Constant*.

Hartl, M. (2010). *The Tau Manifesto*. https://tauday.com/tau-manifesto.

He, B. J. (2014). "Scale-Free Brain Activity: Past, Present, and Future". In: *Trends in cognitive sciences* 18.9, pp. 480–487. DOI: 10/gctmp2.

He, B. J., J. M. Zempel, A. Z. Snyder, and M. E. Raichle (2010). "The Temporal Structures and Functional Significance of Scale-Free Brain Activity". In: *Neuron* 66.3, pp. 353–369. DOI: 10/dwns22.

Helmholtz, H. von (1870). *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. 2007, Reprint der 3. Auflage von 1870. Edition classic. Saarbrücken: VDM, Müller. ISBN: 978-3-8364-0606-2.

Hildebrandt, G. (1976). "Einführungsreferat Chronobiologische Grundlagen Der Leistungsfähigkeit Und Chronohygiene". In: *Biologische Rhythmen Und Arbeit*. Springer, pp. 1–19.

Hinterberger, T., N. Walter, C. Doliwa, and T. Loew (2019). "The Brain's Resonance with Breathing—Decelerated Breathing Synchronizes Heart Rate and Slow Cortical Potentials". In: *Journal of breath research* 13.4, p. 046003. DOI: 10/gnzbnz.

Hirsch, J. A. and B. Bishop (Oct. 1981). "Respiratory Sinus Arrhythmia in Humans: How Breathing Pattern Modulates Heart Rate". In: *American Journal of Physiology-Heart and Circulatory Physiology* 241.4, H620–H629. DOI: 10/gjvpcd.

Hirschman, I. I. (Jan. 1957). "A Note on Entropy". In: *American Journal of Mathematics* 79.1, p. 152. DOI: 10/bxtjbn.

Hjorth, B. (1975). "An On-Line Transformation of EEG Scalp Potentials into Orthogonal Source Derivations". In: *Electroencephalography and clinical neurophysiology* 39.5, pp. 526–530. DOI: 10/cg7zzq.

Hlawatsch, F. and G. F. Boudreaux-Bartels (1992). "Linear and Quadratic Time-Frequency Signal Representations". In: *IEEE signal processing magazine* 9.2, pp. 21–67. DOI: 10/df8bhq.

Hoffmann, U. (2007). "Bayesian Machine Learning Applied in a Brain-Computer Interface for Disabled Users". PhD thesis. EPFL.

Holighaus, N., G. Koliander, Z. Průša, and L. D. Abreu (Aug. 2019). "Characterization of Analytic Wavelet Transforms and a New Phaseless Reconstruction Algorithm". In: *IEEE Transactions on Signal Processing* 67.15, pp. 3894–3908. DOI: 10/gkzc4s. arXiv: 1906.00738.

Holschneider, M. and P. Tchamitchian (1990). "Régularité Locale de La Fonction "Non-Differentiable" de Riemann". In: *Les Ondelettes En 1989*. Springer, pp. 102–124.

Hu, S., D. Yao, M. L. Bringas-Vega, Y. Qin, and P. A. Valdes-Sosa (July 2019). "The Statistics of EEG Unipolar References: Derivations and Properties". In: *Brain Topography* 32.4, pp. 696–703. DOI: 10/gf9tsw.

Ichimaru, Y. and G. Moody (Apr. 1999). "Development of the Polysomnographic Database on CD-ROM". In: *Psychiatry and Clinical Neurosciences* 53.2, pp. 175–177. DOI: 10/ddrg5n.

Ivanov, P. C., K. K. L. Liu, and R. P. Bartsch (Oct. 2016). "Focus on the Emerging New Fields of Network Physiology and Network Medicine". In: *New Journal of Physics* 18.10, p. 100201. DOI: 10.1088/1367-2630/18/10/100201.

Jiang, X., G.-B. Bian, and Z. Tian (Feb. 2019). "Removal of Artifacts from EEG Signals: A Review". In: *Sensors* 19.5, p. 987. DOI: 10/ggzm49.

Johnson, A. E., J. Behar, F. Andreotti, G. D. Clifford, and J. Oster (2014). "R-Peak Estimation Using Multimodal Lead Switching". In: *Computing in Cardiology 2014*. IEEE, pp. 281–284. ISBN: 1-4799-4347-9.

Jones, D. L. and T. W. Parks (1990). "A High Resolution Data-Adaptive Time-Frequency Representation". In: *IEEE Transactions on Acoustics, Speech, and Signal Processing* 38.12, pp. 2127–2135. DOI: 10/bq9xjv.

Jones, S. R. (Oct. 2016). "When Brain Rhythms Aren't 'Rhythmic': Implication for Their Mechanisms and Meaning". In: *Current Opinion in Neurobiology* 40, pp. 72–80. DOI: 10/f889sr.

Karlen, W., S. Raman, J. M. Ansermino, and G. A. Dumont (July 2013). "Multiparameter Respiratory Rate Estimation From the Photoplethysmogram". In: *IEEE Transactions on Biomedical Engineering* 60.7, pp. 1946–1953. DOI: 10/f43p7r.

Kayser, J. and C. E. Tenke (Feb. 2006a). "Principal Components Analysis of Laplacian Waveforms as a Generic Method for Identifying ERP Generator Patterns: I. Evaluation with Auditory Oddball Tasks". In: *Clinical Neurophysiology* 117.2, pp. 348–368. DOI: 10/b4qd59.

— (Feb. 2006b). "Principal Components Analysis of Laplacian Waveforms as a Generic Method for Identifying ERP Generator Patterns: II. Adequacy of Low-Density Estimates". In: *Clinical Neurophysiology* 117.2, pp. 369–380. DOI: 10/d8vnbg.

— (Sept. 2015). "On the Benefits of Using Surface Laplacian (Current Source Density) Methodology in Electrophysiology". In: *International Journal of Psychophysiology* 97.3, pp. 171–173. DOI: 10/gjkwbf.

Kemp, B. and H. A. C. Kamphuisen (Sept. 1986). "Simulation of Human Hypnograms Using a Markov Chain Model". In: *Sleep* 9.3, pp. 405–414. DOI: 10/gjvpc7.

Kendall, M. G., A. Stuart, and J. K. Ord (1987). *Kendall's Advanced Theory of Statistics*. Oxford University Press, Inc. ISBN: 0-19-520561-8.

Khalil, A., G. Joncas, F. Nekka, P. Kestener, and A. Arnéodo (2006). "Morphological Analysis of HI Features. II. Wavelet-based Multifractal Formalism". In: *The Astrophysical Journal Supplement Series* 165.2, p. 512. DOI: 10/bvxrvz.

Kishi, A., I. Yamaguchi, F. Togo, and Y. Yamamoto (Aug. 2018). "Markov Modeling of Sleep Stage Transitions and Ultradian REM Sleep Rhythm". In: *Physiological Measurement* 39.8, p. 084005. DOI: 10/gjvpcz.

Koliander, G., L. D. Abreu, A. Haimi, and J. L. Romero (2019). "Filtering the Continuous Wavelet Transform Using Hyperbolic Triangulations". In: *2019 13th International Conference on Sampling Theory and Applications (SampTA)*. IEEE, pp. 1–4. ISBN: 1-72813-741-1.

Kovesi, P. (Sept. 2015). *Good Colour Maps: How to Design Them*. arXiv: 1509.03700 [cs].

Kralemann, B., L. Cimponeriu, M. Rosenblum, A. Pikovsky, and R. Mrowka (June 2008). "Phase Dynamics of Coupled Oscillators Reconstructed from Data". In: *Physical Review E* 77.6, p. 066205. DOI: 10/bmxhb9.

Kralemann, B., M. Frühwirth, A. Pikovsky, M. Rosenblum, T. Kenner, J. Schaefer, and M. Moser (Dec. 2013). "In Vivo Cardiac Phase Response Curve Elucidates Human Respiratory Heart Rate Variability". In: *Nature Communications* 4.1, p. 2418. DOI: 10/gbdwc4.

Kranjec, J., S. Beguš, G. Geršak, and J. Drnovšek (Sept. 2014). "Non-Contact Heart Rate and Heart Rate Variability Measurements: A Review". In: *Biomedical Signal Processing and Control* 13, pp. 102–112. DOI: 10/f6cm3c.

Kronland-Martinet, R., J. Morlet, and A. Grossmann (Aug. 1987). "Analysis of Sound Patterns through Wavelet Transforms". In: *International Journal of Pattern Recognition and Artificial Intelligence* 01.02, pp. 273–302. DOI: 10/bfs8dp.

Kryger, M. H., T. Roth, and W. C. Dement (2005). *Principles and Practice of Sleep Medicine*. Philadelphia, PA: Elsevier/Saunders. ISBN: 978-0-7216-0797-9 978-1-4160-0320-5.

Kubo, R. (1966). "The Fluctuation-Dissipation Theorem". In: *Reports on progress in physics* 29.1, p. 255. DOI: 10/fq47nr.

Lancaster, G., D. Iatsenko, A. Pidde, V. Ticcinelli, and A. Stefanovska (July 2018). "Surrogate Data for Hypothesis Testing of Physical Systems". In: *Physics Reports* 748, pp. 1–60. DOI: 10/gg85xh.

Le, T.-P. and P. Argoul (Oct. 2004). "Continuous Wavelet Transform for Modal Identification Using Free Decay Response". In: *Journal of Sound and Vibration* 277.1-2, pp. 73–100. DOI: 10/dkdr47.

— (Dec. 2016). "Modal Identification Using the Frequency-Scale Domain Decomposition Technique of Ambient Vibration Responses". In: *Journal of Sound and Vibration* 384, pp. 325–338. DOI: 10/f8624r.

Lee, J.-M., D.-J. Kim, I.-Y. Kim, K.-S. Park, and S. I. Kim (2002). "Detrended Fluctuation Analysis of EEG in Sleep Apnea Using MIT/BIH Polysomnography Data". In: *Computers in biology and medicine* 32.1, pp. 37–47. DOI: 10/fg4kk9.

Li, F., W. Song, C. Li, and A. Yang (2019). "Non-Harmonic Analysis Based Instantaneous Heart Rate Estimation from Photoplethysmography". In: *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, pp. 1279–1283. ISBN: 1-4799-8131-1.

Lilly, J. and S. Olhede (Jan. 2009). "Higher-Order Properties of Analytic Wavelets". In: *IEEE Transactions on Signal Processing* 57.1, pp. 146–160. DOI: 10/fs9w7t.

Lilly, J. M. and S. C. Olhede (Aug. 2010). "On the Analytic Wavelet Transform". In: *IEEE Transactions on Information Theory* 56.8, pp. 4135–4156. DOI: 10/d69j62. arXiv: 0711.3834.

— (Nov. 2012). "Generalized Morse Wavelets as a Superfamily of Analytic Wavelets". In: *IEEE Transactions on Signal Processing* 60.11, pp. 6036–6041. DOI: 10/gjvnrq.

Lin, A., K. K. L. Liu, R. P. Bartsch, and P. C. Ivanov (Dec. 2020). "Dynamic Network Interactions among Distinct Brain Rhythms as a Hallmark of Physiologic State and Function". In: *Communications Biology* 3.1, p. 197. DOI: 10/ghj6wf.

Lin, C.-Y., L. Su, and H.-T. Wu (2018). "Wave-Shape Function Analysis". In: *Journal of Fourier Analysis and Applications* 24.2, pp. 451–505. DOI: 10/gc8dvx.

Lind, B. K., J. L. Goodwin, J. G. Hill, T. Ali, S. Redline, and S. F. Quan (2003). "Recruitment of Healthy Adults into a Study of Overnight Sleep Monitoring in the Home: Experience of the Sleep Heart Health Study". In: *Sleep and Breathing* 7.1, pp. 13–24. DOI: 10/drhrpz.

Liu, P. C. (1994). "Wavelet Spectrum Analysis and Ocean Wind Waves". In: *Wavelet Analysis and Its Applications*. Vol. 4. Elsevier, pp. 151–166. ISBN: 978-0-08-052087-2.

Lovett, E. G. and K. M. Ropella (Nov. 1997). "Time-Frequency Coherence Analysis of Atrial Fibrillation Termination during Procainamide Administration". In: *Annals of Biomedical Engineering* 25.6, pp. 975–984. DOI: 10/fw7rcv.

Mardia, K. V. and P. E. Jupp (2000). *Directional Statistics*. Wiley Series in Probability and Statistics. Chichester ; New York: J. Wiley. ISBN: 978-0-471-95333-3.

Matz, G. and F. Hlawatsch (2000). "Time-Frequency Coherence Analysis of Nonstationary Random Processes". In: *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing (Cat. No. 00TH8496)*. IEEE, pp. 554–558. ISBN: 0-7803-5988-7.

Meyer, Y., F. Sellan, and M. S. Taqqu (1999). "Wavelets, Generalized White Noise and Fractional Integration: The Synthesis of Fractional Brownian Motion". In: *Journal of Fourier Analysis and Applications* 5.5, pp. 465–494. DOI: 10/cbnv4m.

Minkowski, H. (1910). *Geometrie Der Zahlen*. Vol. 40. Teubner.

Mohrman, D. E. and L. J. Heller (2018). *Cardiovascular Physiology*. 9th edition. New York: McGraw-Hill, Education/Medical. ISBN: 978-1-260-02611-5.

Moody, G. B., R. G. Mark, M. A. Bump, J. S. Weinstein, A. D. Berman, J. E. Mietus, and A. L. Goldberger (1986). "Clinical Validation of the ECG-derived Respiration (EDR) Technique". In: *Computers in cardiology* 13.1, pp. 507–510.

Morse, P. M. (July 1929). "Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels". In: *Physical Review* 34.1, pp. 57–64. DOI: 10/d67h5k.

Moser, M., M. Lehofer, G. Hildebrandt, M. Voica, S. Egner, and T. Kenner (Feb. 1995). "Phase and Frequency Coordination of Cardiac and Respiratory Function". In: *Biological Rhythm Research* 26.1, pp. 100–111. DOI: 10/db4c4p.

Moser, M., M. Fruhwirth, and T. Kenner (2008). "The Symphony of Life [Chronobiological Investigations]". In: *IEEE Engineering in Medicine and Biology Magazine* 27.1, pp. 29–37. DOI: 10/fgh23h.

Moser, M., M. Frühwirth, R. Penter, and R. Winker (May 2006). "Why Life Oscillates – from a Topographical Towards a Functional Chronobiology". In: *Cancer Causes & Control* 17.4, pp. 591–599. DOI: 10/fdjspv.

Munia, T. T. K. and S. Aviyente (Dec. 2019). "Comparison of Wavelet and RID-Rihaczek Based Methods for Phase-Amplitude Coupling". In: *IEEE Signal Processing Letters* 26.12, pp. 1897–1901. DOI: 10/gp7gd5.

Muzy, J., E. Bacry, and A. Arneodo (Apr. 1994). "The Multifractal Formalism Revisited with Wavelets". In: *International Journal of Bifurcation and Chaos* 04.02, pp. 245–302. DOI: 10/fhhp5h.

Muzy, J.-F., E. Bacry, and A. Arneodo (1991). "Wavelets and Multifractal Formalism for Singular Signals: Application to Turbulence Data". In: *Physical review letters* 67.25, p. 3515. DOI: 10/cd4fvt.

Muzy, J.-F., R. Baïle, and E. Bacry (Apr. 2013). "Random Cascade Model in the Limit of Infinite Integral Scale as the Exponential of a Nonstationary 1 / f Noise: Application to Volatility Fluctuations in Stock Markets". In: *Physical Review E* 87.4, p. 042813. DOI: 10/f45znn.

Olbrich, E., J. C. Claussen, and P. Achermann (Oct. 2011). "The Multiple Time Scales of Sleep Dynamics as a Challenge for Modelling the Sleeping Brain". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 369.1952, pp. 3884–3901. DOI: 10/d53z78.

Olhede, S. and A. Walden (Nov. 2002). "Generalized Morse Wavelets". In: *IEEE Transactions on Signal Processing* 50.11, pp. 2661–2670. DOI: 10/dbm7ph.

Opie, L. H. (2004). *Heart Physiology: From Cell to Circulation*. 4th ed. Philadelphia: Lippincott Williams & Wilkins. ISBN: 978-0-7817-4278-8.

Orini, M., R. Bailon, L. Mainardi, A. Minchole, and P. Laguna (2009). "Continuous Quantification of Spectral Coherence Using Quadratic Time-Frequency Distributions: Error Analysis and Application". In: *Computers in Cardiology*. Vol. 36, pp. 681–684. ISBN: 978-1-4244-7282-6.

Orini, M., R. Bailon, L. T. Mainardi, P. Laguna, and P. Flandrin (Mar. 2012). "Characterization of Dynamic Interactions Between Cardiovascular Signals by Time-Frequency Coherence". In: *IEEE Transactions on Biomedical Engineering* 59.3, pp. 663–673. DOI: 10/b4rwmn.

Osselton, J. W. (Dec. 1969). "Bipolar, Unipolar and Average Reference Recording Methods II: Mainly Practical Considerations". In: *American Journal of EEG Technology* 9.4, pp. 117–133. DOI: 10/gnq74r.

Palais, B. (2001). "π Is Wrong!" In: *The Mathematical Intelligencer Springer-Verlag New York* 23.3, pp. 7–8.

Papandreou-Suppappola, A., F. Hlawatsch, and G. F. Boudreaux-Bartels (1998). "Quadratic Time-Frequency Representations with Scale Covariance and Generalized Time-Shift Covariance: A Unified Framework for the Affine, Hyperbolic, and Power Classes". In: *Digital Signal Processing* 8.1, pp. 3–48. DOI: 10/d2twxd.

Park, H.-D. and O. Blanke (2019). "Heartbeat-Evoked Cortical Responses: Underlying Mechanisms, Functional Roles, and Methodological Considerations". In: *Neuroimage* 197, pp. 502–511. DOI: 10/gmgkm8.

Penzel, T., J. W. Kantelhardt, R. P. Bartsch, M. Riedl, J. F. Kraemer, N. Wessel, C. Garcia, M. Glos, I. Fietze, and C. Schöbel (2016). "Modulations of Heart Rate, ECG, and Cardio-Respiratory Coupling Observed in Polysomnography". In: *Frontiers in physiology* 7, p. 460. DOI: 10/gpdw2n.

Percival, D. B. and A. T. Walden (1993). *Spectral Analysis for Physical Applications: Multitaper and Conventional Univariate Techniques*. Cambridge ; New York, NY, USA: Cambridge University Press. ISBN: 978-0-521-35532-2 978-0-521-43541-3.

Perrin, F., J. Pernier, O. Bertrand, and J. F. Echallier (1989). "Spherical Splines for Scalp Potential and Current Density Mapping". In: *Electroencephalography and clinical neurophysiology* 72.2, pp. 184–187. DOI: 10/frdn6x.

Picinbono, B. (Mar. 1997). "On Instantaneous Amplitude and Phase of Signals". In: *IEEE Transactions on Signal Processing* 45.3, pp. 552–560. DOI: 10/czznxx.

Pikovsky, A., M. Rosenblum, and J. Kurths (2001). *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge: Cambridge University Press. ISBN: 978-0-511-07752-4.

Pitton, J. W. (July 2000). "The Statistics of Time–Frequency Analysis". In: *Journal of the Franklin Institute* 337.4, pp. 379–388. DOI: 10/cpzg52.

Plemelj, J. (1964). *Problems in the Sense of Riemann and Klein*. 16. Interscience Publishers. ISBN: 978-0-470-69125-0.

Prerau, M. J., R. E. Brown, M. T. Bianchi, J. M. Ellenbogen, and P. L. Purdon (Jan. 2017). "Sleep Neurophysiological Dynamics Through the Lens of Multitaper Spectral Analysis". In: *Physiology* 32.1, pp. 60–92. DOI: 10/f9gvdh.

Purcell, S. M., D. S. Manoach, C. Demanuele, B. E. Cade, S. Mariani, R. Cox, G. Panagiotaropoulou, R. Saxena, J. Q. Pan, J. W. Smoller, S. Redline, and R. Stickgold (Aug. 2017). "Characterizing Sleep Spindles in 11,630 Individuals from the National Sleep Research Resource". In: *Nature Communications* 8.1, p. 15930. DOI: 10/gbjxzp.

Quan, S. F., B. V. Howard, C. Iber, J. P. Kiley, F. J. Nieto, G. T. O'Connor, D. M. Rapoport, S. Redline, J. Robbins, and J. M. Samet (1997). "The Sleep Heart Health Study: Design, Rationale, and Methods". In: *Sleep* 20.12, pp. 1077–1085.

Randall, R. B. (Dec. 2017). "A History of Cepstrum Analysis and Its Application to Mechanical Problems". In: *Mechanical Systems and Signal Processing* 97, pp. 3–19. DOI: 10/gdgv2m.

Rechtschaffen, A. (1968). "A Manual for Standardized Terminology, Techniques and Scoring System for Sleep Stages in Human Subjects". In: *Brain information service*.

Riedl, M., A. Müller, J. F. Kraemer, T. Penzel, J. Kurths, and N. Wessel (Apr. 2014). "Cardio-Respiratory Coordination Increases during Sleep Apnea". In: *PLoS ONE* 9.4. Ed. by J. Garcia-Ojalvo, e93866. DOI: 10/f57n83.

Rosenblum, M., A. Pikovsky, J. Kurths, C. Schäfer, and P. A. Tass (2001). "Phase Synchronization: From Theory to Data Analysis". In: *Handbook of Biological Physics*. Vol. 4. Elsevier, pp. 279–321. ISBN: 1383-8121.

Rouyer, T., J. Fromentin, N. Stenseth, and B. Cazelles (May 2008). "Analysing Multiple Time Series and Extending Significance Testing in Wavelet Analysis". In: *Marine Ecology Progress Series* 359, pp. 11–23. DOI: 10/bnc43m.

Saul, J. P., R. D. Berger, P. Albrecht, S. Stein, M. H. Chen, and R. Cohen (1991). "Transfer Function Analysis of the Circulation: Unique Insights into Cardiovascular Regulation". In: *American Journal of Physiology-Heart and Circulatory Physiology* 261.4, H1231–H1245. DOI: 10/gnm4t5.

Saul, J. P., R. D. Berger, M. H. Chen, and R. J. Cohen (Jan. 1989). "Transfer Function Analysis of Autonomic Regulation. II. Respiratory Sinus Arrhythmia". In: *American Journal of Physiology-Heart and Circulatory Physiology* 256.1, H153–H161. DOI: 10/gnm4tp.

Saxena, S. C., V. Kumar, and S. T. Hamde (Jan. 2002). "Feature Extraction from ECG Signals Using Wavelet Transforms for Disease Diagnostics". In: *International Journal of Systems Science* 33.13, pp. 1073–1085. DOI: 10/dw3b9n.

Schäfer, C., M. G. Rosenblum, J. Kurths, and H.-H. Abel (Mar. 1998). "Heartbeat Synchronized with Ventilation". In: *Nature* 392.6673, pp. 239–240. DOI: 10/bbwjzh.

Scholkmann, F. and U. Wolf (Apr. 2019). "The Pulse-Respiration Quotient: A Powerful but Untapped Parameter for Modern Studies About Human Physiology and Pathophysiology". In: *Frontiers in Physiology* 10, p. 371. DOI: 10/ghwqnz.

Seip, K. (1989). "Mean Value Theorems and Concentration Operators in Bargmann and Bergman Space". In: *Wavelets*. Springer, pp. 209–215.

Seip, K. (1993). "Wavelets in H2 (R) Sampling, Interpolation, and Phase Space Density". In: *Wavelets: A Tutorial in Theory and Applications*, pp. 529–540.

Sekimoto, K. (2010). *Stochastic Energetics*. Lecture Notes in Physics 799. Heidelberg ; New York: Springer. ISBN: 978-3-642-05410-5.

Selesnick, I. (June 2001). "Hilbert Transform Pairs of Wavelet Bases". In: *IEEE Signal Processing Letters* 8.6, pp. 170–173. DOI: 10/dws2g6.

— (May 2002). "The Design of Approximate Hilbert Transform Pairs of Wavelet Bases". In: *IEEE Transactions on Signal Processing* 50.5, pp. 1144–1152. DOI: 10/dh4qks.

Selesnick, I., R. Baraniuk, and N. Kingsbury (Nov. 2005). "The Dual-Tree Complex Wavelet Transform". In: *IEEE Signal Processing Magazine* 22.6, pp. 123–151. DOI: 10/fwhmtz.

Sethares, W. A. (2005). *Tuning, Timbre, Spectrum, Scale*. 2nd ed. London: Springer. ISBN: 978-1-85233-797-1.

Slepian, D. (May 1978). "Prolate Spheroidal Wave Functions, Fourier Analysis, and Uncertainty-V: The Discrete Case". In: *Bell System Technical Journal* 57.5, pp. 1371–1430. DOI: 10/gdm5d2.

Stanković, L. and S. Stanković (Feb. 1993). "Wigner Distribution of Noisy Signals". In: *IEEE Transactions on Signal Processing* 41.2, pp. 956–960. DOI: 10/brfrr8.

Stanković, L. (Mar. 2001). "A Measure of Some Time–Frequency Distributions Concentration". In: *Signal Processing* 81.3, pp. 621–631. DOI: 10/dmvq77.

Su, L. and H.-T. Wu (Feb. 2017). "Extract Fetal ECG from Single-Lead Abdominal ECG by De-Shape Short Time Fourier Transform and Nonlocal Median". In: *Frontiers in Applied Mathematics and Statistics* 3. DOI: 10/gj2z39.

Taillard, J., P. Sagaspe, C. Berthomier, M. Brandewinder, H. Amieva, J.-F. Dartigues, M. Rainfray, S. Harston, J.-A. Micoulaud-Franchi, and P. Philip (Mar. 2019). "Non-REM Sleep Characteristics Predict Early Cognitive Impairment in an Aging Population". In: *Frontiers in Neurology* 10, p. 197. DOI: 10/gmvk7k.

Tamura, T. (Feb. 2019). "Current Progress of Photoplethysmography and SPO2 for Health Monitoring". In: *Biomedical Engineering Letters* 9.1, pp. 21–36. DOI: 10/ggc54v.

Terzano, M. G., L. Parrino, A. Sherieri, R. Chervin, S. Chokroverty, C. Guilleminault, M. Hirshkowitz, M. Mahowald, H. Moldofsky, A. Rosa, R. Thomas, and A. Walters (Nov. 2001). "Atlas, Rules, and Recording Techniques for the Scoring of Cyclic Alternating Pattern (CAP) in Human Sleep". In: *Sleep Medicine* 2.6, pp. 537–553. DOI: 10/bb2mnw.

Thomas, R. J., J. E. Mietus, C.-K. Peng, and A. L. Goldberger (2005). "An Electrocardiogram-Based Technique to Assess Cardiopulmonary Coupling during Sleep". In: *Sleep* 28.9, pp. 1151–1161. DOI: 10/gf8nss.

Thomas, R. J., J. E. Mietus, C.-K. Peng, D. Guo, D. Gozal, H. Montgomery-Downs, D. J. Gottlieb, C.-Y. Wang, and A. L. Goldberger (2014). "Relationship between Delta Power and the Electrocardiogram-Derived Cardiopulmonary Spectrogram: Possible Implications for Assessing the Effectiveness of Sleep". In: *Sleep medicine* 15.1, pp. 125–131. DOI: 10/gpdxgh.

Thomson, D. (1982). "Spectrum Estimation and Harmonic Analysis". In: *Proceedings of the IEEE* 70.9, pp. 1055–1096. DOI: 10/fcfqm4.

Torrence, C. and G. P. Compo (1998). "A Practical Guide to Wavelet Analysis". In: *Bulletin of the American Meteorological Society* 79.1, pp. 61–78. DOI: 10/bhbwhf.

Torrésani, B. (1995). *Analyse continue par ondelettes*. Savoirs actuels. Paris: InterEditions CNRS éd. ISBN: 978-2-7296-0591-9 978-2-271-05364-0.

Urigüen, J. A. and B. Garcia-Zapirain (2015). "EEG Artifact Removal—State-of-the-Art and Guidelines". In: *Journal of neural engineering* 12.3, p. 031001. DOI: 10/gc6zcx.

Venugopal, V., S. G. Roux, E. Foufoula-Georgiou, and A. Arneodo (June 2006). "Revisiting Multifractality of High-Resolution Temporal Rainfall Using a Wavelet-Based Formalism". In: *Water Resources Research* 42.6. DOI: 10/dh43dk.

Vest, A. N., G. Da Poian, Q. Li, C. Liu, S. Nemati, A. J. Shah, and G. D. Clifford (Oct. 2018). "An Open Source Benchmarked Toolbox for Cardiovascular Waveform and Interval Analysis". In: *Physiological Measurement* 39.10, p. 105004. DOI: 10/gmdh98.

Vest, A. N., G. D. Poian, Q. Li, C. Liu, S. Nemati, A. Shah, and G. D. Clifford (Sept. 2019). *Cliffordlab/PhysioNet-Cardiovascular-Signal-Toolbox: PhysioNet-Cardiovascular-Signal-Toolbox 1.0.2*. Zenodo. DOI: 10.5281/ZENODO.1243111.

Ville, J. (1948). "Théorie et Application de La Notion de Signal Analytique". In: *Câbles et transmissions* 2.1, pp. 61–74.

Walden, A. T. (Dec. 2000). "A Unified View of Multitaper Multivariate Spectral Estimation". In: *Biometrika* 87.4, pp. 767–788. DOI: 10/d8dxbd.

Walden, A. T. and E. A. K. Cohen (Sept. 2012). "Statistical Properties for Coherence Estimators From Evolutionary Spectra". In: *IEEE Transactions on Signal Processing* 60.9, pp. 4586–4597. DOI: 10/gjvpmj.

White, G. (Mar. 2015). "Rush τ, Rush π!" In: *The Physics Teacher* 53.3, pp. 132–132. DOI: 10/gmq52z.

White, L. and B. Boashash (July 1990). "Cross Spectral Analysis of Nonstationary Processes". In: *IEEE Transactions on Information Theory* 36.4, pp. 830–835. DOI: 10/bhxr6c.

Wigner, E. (June 1932). "On the Quantum Correction For Thermodynamic Equilibrium". In: *Physical Review* 40.5, pp. 749–759. DOI: 10/bmw7jh.

Williams, W. J., M. L. Brown, and A. O. Hero III (Dec. 1991). "Uncertainty, Information, and Time-Frequency Distributions". In: *San Diego, '91, San Diego, CA*. Ed. by F. T. Luk. San Diego, CA, pp. 144–156. DOI: 10/dt8pp9.

Yaesoubi, M., E. A. Allen, R. L. Miller, and V. D. Calhoun (Oct. 2015). "Dynamic Coherence Analysis of Resting fMRI Data to Jointly Capture State-Based Phase, Frequency, and Time-Domain Information". In: *NeuroImage* 120, pp. 133–142. DOI: 10/f7sn8s.

Yan Xu, S. Haykin, and R. Racine (July 1999). "Multiple Window Time-Frequency Distribution and Coherence of EEG Using Slepian Sequences and Hermite Functions". In: *IEEE Transactions on Biomedical Engineering* 46.7, pp. 861–866. DOI: 10/dtsxt8.

Yao, D., Y. Qin, S. Hu, L. Dong, M. L. Bringas Vega, and P. A. Valdés Sosa (July 2019). "Which Reference Should We Use for EEG and ERP Practice?" In: *Brain Topography* 32.4, pp. 530–549. DOI: 10/ghhtvh.

Zhan, Y., D. Halliday, P. Jiang, X. Liu, and J. Feng (Sept. 2006). "Detecting Time-Dependent Coherence between Non-Stationary Electrophysiological Signals—A Combined Statistical and Time–Frequency Approach". In: *Journal of Neuroscience Methods* 156.1-2, pp. 322–332. DOI: 10/d37hjx.

Zhang, G.-Q., L. Cui, R. Mueller, S. Tao, M. Kim, M. Rueschman, S. Mariani, D. Mobley, and S. Redline (2018). "The National Sleep Research Resource: Towards a Sleep Data Commons". In: *Journal of the American Medical Informatics Association* 25.10, pp. 1351–1358. DOI: 10/gdntnm.

Bonus Part C.
Sonance: revisiting synchronization for multi-frequency signals

Contents

1. Introduction

Pulsatile and quasi-periodic behaviors are ubiquitous in non-linear out-of-equilibrium systems, in particular in living organisms. The synchronization of two rhythms at nearby characteristic frequencies has been well documented in physics, both in experimental systems and theoretical models. The nature and origin of synchronization in physiological systems has recently attracted much interest in a broad scientific community. One essential property that distinguishes living organisms from inert matter is their strong temporal variability. Interestingly, fluctuations arising from physiological networks and their rich spectral content could even facilitate the occurrence of synchronization. While synchronization is usually considered between two rhythms of single and often close frequencies, we propose to extend it to a multifrequency synchronization that results from all possible ratios a : b between frequencies of the compared physiological fluctuations. For exemple, the ratios a = 3, 4, 5, 6 and b = 1 between the heart and breathing fundamental frequencies are most likely for mammals. Thanks to a continuous complex wavelet transform analysis of physiological signals, we reveal the temporal evolution of their frequency ratios, even for arbitrarily rich spectral contents. Then we introduce a measure that quantifies the (ir)rationality of the frequency ratios, tightly related to the theory of continued fractions and rational numbers enumeration. We call it "sonance", or musical measure, in reference to concepts of consonance and dissonance in musical harmony. The temporal evolution of the frequency ratio content and its sonance are computed for different pairs of physiological signals, such as photoplethismograms, thermograms and polysomnograms, giving new insights into the fluctuations of physiological rhythms and inter-regulations between organs.

1.1. Generalizing synchronization from single modes of close frequencies to arbitrarily rich spectral content

The simplest example of synchronization is a 1 : 1 ratio between the characteristic frequencies of the considered rhythms. However, synchronizations of ratios 1 : n are not rare and ratios of integers m : n appears to be the most general form of synchronization. Therefore, synchronization has something to do with rational relation between frequencies. Our strategy is as follows. First we define, from the amplitude of two frequency spectra, the frequency ratio content between these spectra. Then, the non-stationarity of real signals is considered by extending the method to time-frequency representations. In particular, the use of analytic wavelets appears to be very appropriate. Finally, we define a new quantity that quantifies how rational the frequency ratios are, i.e. how synchronized the rhythms are.

1.2. Interpretation in terms of musical harmony

The musical term for synchronization in a multi-frequency sense is the consonance. Indeed, the 1 : 1 synchronization is called unison, 1 : 2 the octave, 2 : 3 the fifth, 3 : 4 the fourth, etc. The dissonances imply more complicated ratio or irrationals far from these simple ratios. The discrepancy between these musical terms and the rational number notation comes from the fact that a musician thinks of these relations in an additive way, i.e. in the log-frequency domain, where the base of the logarithm is one twelfth of octave: $\log(2)/12$. The reason happens to be related to a measure, that we call *sonance* because it places consonance and dissonance on a continuum. It quantifies how (ir)rational is a distribution of ratios, obtained from the frequencies present in two signals. In agreement with this musical intuition, we introduce a time-frequency representation of the signals and a log-frequency sampling to express easily the frequency ratio content as cross-correlation of the spectra.

1.3. Conventions: Fourier transform and analytic wavelet transform

Let us set $\tau = 2\pi$.

The Fourier transform of $x(t)$, denoted with a tilde, is defined here as:

$$\tilde{x}(f) = \mathcal{F}[x](f) = \int_{-\infty}^{+\infty} x(t) e^{-i\tau f t} \mathrm{d}t$$

so that its inverse is

$$x(t) = \mathcal{F}^{-1}[\tilde{x}](t) = \int_{-\infty}^{+\infty} \tilde{x}(f) e^{i\tau f t} \mathrm{d}f \quad .$$

$x(t)$ and $\tilde{x}(f)$ will be referred to as the time-domain and frequency-domain representations respectively.

Also, let us define the power spectral density:

$$PSD[x](f) = \lim_{T\to\infty} \frac{1}{T} \left| \int_0^T x(t) e^{-i\tau f t} \mathrm{d}t \right|^2 \quad ,$$

which can be understood as the squared modulus of the Fourier spectrum scaled by the signal duration $|\tilde{x}(f)|^2/T$.

Let us define the wavelet transform of a signal $s(t)$ in the form of a time-frequency representation:

$$\mathcal{W}_\psi[s](f,t) = \int s(t) \overline{\psi} \left(\frac{f}{f_0}(t'-t) \right) \frac{f}{f_0} \mathrm{d}t' \quad , \tag{1}$$

where $\overline{\psi}$ is the complex conjugate of the wavelet, f is the frequency parameter and t is the time translation parameter. The usual scale parameter of the wavelet transform here would be $a = \frac{f_0}{f}$ where f_0 is a reference frequency, characteristic of the wavelet. The norm of the wavelet which is conserved through scales is here chosen to be $p = 1$. The relation for other conventions (especially $p = 2$) is:

$$\mathcal{W}_\psi[s](f,t;p) = \left(\frac{f}{f_0} \right)^{\frac{1-p}{p}} \mathcal{W}_\psi[s](f,t) \quad . \tag{2}$$

In the frequency domain, the wavelet transform reads as a parametrized band-pass filtering of the signal:

$$\mathcal{W}_\psi[s](f,t) = \int \tilde{s}(f') \overline{\tilde{\psi}} \left(\frac{f'}{f} f_0 \right) e^{i\tau f' t} \mathrm{d}f' \quad . \tag{3}$$

This relation is extensively used in the following. In order to interpret it as a time-frequency representation, we consider analytic wavelets, i.e. with no negative frequency, so that the integral Eq. (3) is over $\mathbb{R}^+$. The analytic property of the wavelet allows a direct measurement of the phase of oscillating components, by completing the real signal by its Hilbert transform pair as its imaginary part. Moreover, we choose the wavelet so that its characteristic frequency f_0 (peak frequency in the Fourier domain) is also the unit frequency, $f_0 = 1$. By means of a slight shift of paradigm regarding the dimensionality of the wavelet variable, this allows to omit f_0 while ensuring a direct correspondence between frequencies of the Fourier and continuous wavelet transforms. Eventually, $\mathcal{W}_\psi[s](f,t)$ represents at any time a local frequency spectrum of the signal s, whose amplitude and phase for any time-frequency component is contained into the modulus and the complex argument of the transform.

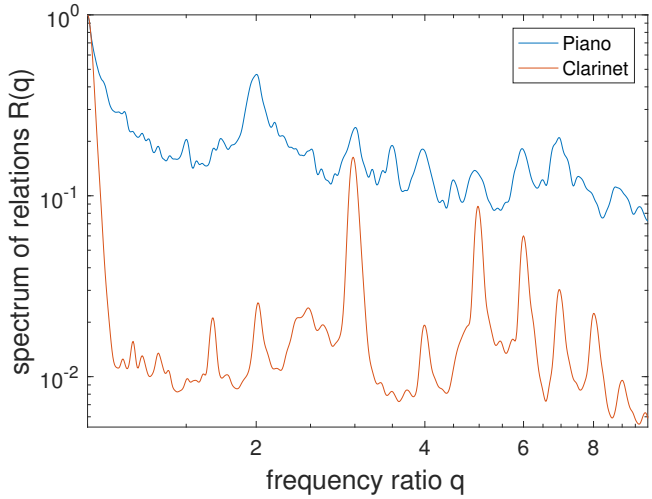


Figure 1: Spectra of relations $R[S,S](q)$ in the case $S_1 = S_2 = S$ for a piano sound (blue) and a clarinet sound (orange), both played staccato. Only $q > 1$ is represented because $R[S,S](q)$ is symmetric under $q \to q^{-1}$. The convention for the spectra is $\mu = 1, \nu = 0$ with the normalization by the norm 2. We observe clear differences in the ratios distribution ; for instance the clarinet as much more odd integers than even ones, due to acoustic impedance constraints. The spectrum is much less peaked in the piano case, possibly due to the presence of numerous harmonic frequencies.

The shape of the wavelet provides a time-frequency uncertainty relation for this representation: according to the uncertainty principle, the product of the widths Δt of ψ and Δf of $\tilde{\psi}$ is some constant, bounded by below by $\frac{1}{2\tau}$. For a wavelet dilated at frequency f, the time width is proportional to f^{-1} whereas the frequency width is proportional to f so that their product is the same constant. Therefore, the uncertainty of the time-frequency representation at any (f,t) is well described by the dimensionaless constants $f\Delta t$ and $\Delta f/f$ (relative frequency resolution).

As an illustration, the time-integrated squared modulus of the wavelet transform is a frequency smoothing of the squared Fourier spectrum of the signal by the wavelet squared spectrum:

$$\int_{-\infty}^{+\infty} |\mathcal{W}_\psi[s](f,t)|^2 \mathrm{d}t = \int_0^\infty |\tilde{s}(f')|^2 |\tilde{\psi}(f'/f)|^2 \mathrm{d}f' \quad , \tag{4}$$

from which we can deduce the following scaling between a time average of the wavelet transform and the power spectral density in the case of a stationary signal:

$$\left\langle |\mathcal{W}_\psi[s](f,t)|^2 \right\rangle_t \sim f\, PSD[s](f)\, \frac{1}{2} \int_0^\infty |\tilde{\psi}(f')|^2 \mathrm{d}f'/f' \quad . \tag{5}$$

2. The spectrum of relations or ratios distribution

2.1. Definition of the spectrum of relations

In this section, we introduce what we call a spectrum of relations between two positive spectra S_1, S_2, that we interpret as a distribution of the frequency ratios. We define it to be the multiplicative cross-correlation of these spectra:

$$R[S_1,S_2](q) = \int_0^\infty S_1(f) S_2(qf) \mathrm{d}f/f \quad , \tag{6}$$

so that the value of the spectrum of relations is high whenever peaks at frequencies related by the ratio q from one spectrum to the other are present. Note that the integral in carried over positive frequencies with the measure $\mathrm{d}\log f$, so that the change of variable $v = \log f$ would yield the usual additive cross-correlation integral of shift parameter $\log q$. The spectrum of relations has the following symmetry:

$$R[S_1, S_2](q) = R[S_2, S_1](q^{-1}) \quad . \tag{7}$$

Consider the specific case where the spectrum $S_1(f)$ is compared to $S_2(f) = S_1(\alpha f),\ \alpha > 0$, which is a good model for comparing two different tones of the same instrument. This yields the spectrum of relations $R[S_1, S_2](q) = R[S_1, S_1](\alpha q)$, which is just the spectrum of relations of S_1 dilated by α.

2.2. Possible definitions for the frequency spectra

The spectra $S_i(f)$ $(i = 1, 2)$ can be defined in different ways from frequency or time-frequency representations of the signals. First, we impose the spectra to be positive, which removes the information of the phase difference between components of different frequencies, that is considered irrelevant. This is important to make the spectra invariant with respect to time translation of the signals, as well as all quantities computed from these spectra. As a result, the positive frequency spectrum $S(f)$ is proportional to $|\tilde{s}(f)|^\mu$ (where $\mu = 2$ corresponds to the PSD) or $|\mathcal{W}_\psi[s](f,t)|^\mu$ in the non-stationary case (from the time-frequency interpretation of the analytic wavelet transform introduced above). However, the phase of the wavelet transform can be used in a specific way (frequency reassignment / synchrosqueezing technics), as evoked in the appendix, to refine the definition of the spectra bypassing the frequency (and/or time) uncertainty.

Next, $|\mathcal{W}_\psi[s](f,t)|^2$ scales as $|\tilde{s}(f)|^2 f$ (up to a constant) from Eq. (5). This is related to the choice of the norm $p = 1$ of the wavelet transform, coherent in a certain way with the measure $\mathrm{d}\log f$ in the spectrum of relations, but not consistent with the Fourier domain usually conceived as linear in f. To take into account the different possible choice of p, we introduce a factor f^ν, leading to S being proportional to $|\tilde{s}(f)\sqrt{f}|^\mu f^\nu$ or $|\mathcal{W}_\psi[s](f,t)|^\mu f^\nu = |\mathcal{W}_\psi[s](f,t;p = \frac{\mu}{\mu+\nu})|^\mu$.

But what should be ν? We could consider all these points of view:

- Amplitude density: $|\tilde{s}(f)|\mathrm{d}f = S(f)\mathrm{d}\log f \quad \Rightarrow \quad \mu = 1$ and $\nu = \frac{1}{2}$
- Energy or power density: $PSD[s](f)\mathrm{d}f = S(f)\mathrm{d}\log f \quad \Rightarrow \quad \mu = 2$ and $\nu = 0$
- Total power: $\int_0^\infty PSD[s](f)\mathrm{d}f = R[S,S](1) \quad \Rightarrow \quad \mu = 1$ and $\nu = 0$
- we could even consider a duality between squared Fourier spectrum and cross-correlation $C(\tau) = \int_{-\infty}^{+\infty} s(t)s(t+\tau)\mathrm{d}t$,

$$\int_0^\infty |\tilde{s}_1(f)|^2 |\tilde{s}_2(qf)|^2 \mathrm{d}f\sqrt{q} = \int_0^\infty C_1(\tau)C_2(\tau/q)\mathrm{d}\tau/\sqrt{q} \quad , \tag{8}$$

  to get a definition either from the frequency spectrum or from the cross-correlation:

$$R[S_1, S_2](q) = R[C_1, C_2](q^{-1}) = R[C_2, C_1](q) \quad \Rightarrow \quad \mu = 2 \text{ and } \nu = -\tfrac{1}{2} \quad . \tag{9}$$

Finally, it remains to normalize the frequency spectra S_1, S_2 or the spectrum of relation $R[S_1, S_2]$ so that the result is independent on an amplitude or a time scaling (change of amplitude or time unit) of the initial signals (and possibly of the wavelet). For instance, it could be:

- the norm 1 (probability density) normalization

$$\int_0^\infty R[S_1,S_2](q)\mathrm{d}q/q = 1 \quad \Leftrightarrow \quad \int_0^\infty S(f)\mathrm{d}f/f = \|S\|_1 = 1$$

- or normalizing by $\sqrt{R[S_1,S_1](1)R[S_2,S_2](1)}$, i.e. a norm 2

$$R[S,S](1) = 1 \quad \Leftrightarrow \quad \left(\int_0^\infty S(f)^2\mathrm{d}f/f\right)^{\frac{1}{2}} = \|S\|_2 = 1$$

Here we can observe that these correspond to standardized signals (unit energy) for $\mu = 2, \nu = 0$ in the first case, for $\mu = 1, \nu = 0$ in the second case.

3. How (ir)rational is a ratio distribution?

The spectrum of relations between two signals gives the distribution of the ratio of their respective frequencies. It remains to quantify how (ir)rational is this distribution. We propose here such measure, that we call the *sonance*, which maps the spectrum of relations to a single number.

3.1. Introduction as a comparison to ideally synchronized signals

The construction of the sonance measure can be viewed as a comparison with (a projection on) the spectrum of self-relations of an ideal periodic signal made of a fundamental (here the unit frequency) and its harmonic frequencies of decaying amplitudes:

$$|\tilde{s}_0(f)| = \sum_{n\geq 1} n^{-\alpha}\delta(f-n) = \sum_{n\geq 1} n^{-\alpha-1}\delta(\log f - \log n) \quad . \tag{10}$$

Note that the general form of the frequency spectrum of this signal does not change this polynomial decay of the amplitudes:

$$S_0(f) \propto |\tilde{s}_0(f)\sqrt{f}|^\mu f^\nu \tag{11}$$

$$\propto \sum_{n\geq 1} n^{-\sigma}\delta(\log f - \log n) \quad , \tag{12}$$

where $\sigma = \mu(\alpha+\frac{1}{2})+\nu$. Note that the frequency spectrum is here defined up to a normalization constant (which allows us to include cases where $\mu \neq 1$ such as $PSD[s_0](f) \propto |\tilde{s}_0(f)|^2$, even though the power of a Dirac delta is not well-defined).

When we compare this periodic signal to itself, we obtain a spectrum of self-relations $R_0(q;\sigma) = R[S_0,S_0](q)$ which represent all the ratios between the harmonic and fundamental frequencies that can be expected between two synchronized periodic signals.

$$R_0(q;\sigma) \propto \sum_{m,n\geq 1} (mn)^{-\sigma}\delta(\log q - \log \tfrac{m}{n}) \tag{13}$$

$$\propto \zeta(2\sigma) \sum_{\frac{a}{b}\in\mathbb{Q}^+} (ab)^{-\sigma}\delta(\log q - \log \tfrac{a}{b}) \quad , \tag{14}$$

where we have rewritten $m = ka,\ n = kb$ so that a and b are mutually prime and we have introduced the Riemann zeta function for the sum over k, $\zeta(\sigma) = \sum_{k\geq 1} k^{-\sigma}$. We can normalize it as a probability density function by the norm 1 $\zeta(\sigma)^2$, or by the norm 2 $\zeta(2\sigma)$.

Let us pick the first option so that we obtain a probability measure. We denote it

$$\mathrm{d}\oint_\sigma(q) = R_0(q;\sigma)\mathrm{d}q/q$$

in reference to the musical intuition that led us there and to its properties relevant to music that will be evoked. We use this measure to quantify how similar or dissimilar to $R_0(q;\sigma)$ a spectrum of relation $R_{1,2}(q) = R[S_1, S_2](q)$ is. We call this quantity the *sonance*:

$$\oint_\sigma[R_{1,2}] = \int_0^\infty R_{1,2}(q)\mathrm{d}\oint_\sigma(q) = \int_0^\infty R_{1,2}(q)R_0(q;\sigma)\mathrm{d}q/q \tag{15}$$

$$= \frac{\zeta(2\sigma)}{\zeta(\sigma)^2}\sum_{\frac{a}{b}\in\mathbb{Q}^+}(ab)^{-\sigma}R_{1,2}\left(\tfrac{a}{b}\right) = \frac{\sum_{q\in\mathbb{Q}^+}e^{-\sigma H(q)}R_{1,2}(q)}{\sum_{q\in\mathbb{Q}^+}e^{-\sigma H(q)}} \quad , \tag{16}$$

where we have denoted $\lambda[R] = \int R(q)\mathrm{d}\lambda(q)$ the λ-measure of a function R, and $H(\frac{a}{b}) = \log(ab)$.

Under this last form, we can appreciate what it does in statistical physics terms: the real spectrum of relations $R_{1,2}$ is averaged over a Boltzmann distribution of ratios given by the ideal spectrum of relations. $H(\frac{a}{b}) = \log ab$ in the Boltzmann weights of the measure assigns an energy to each ratio, whereas the inverse temperature parameter σ gives more or less importance to low energy states, i.e. simple ratios. Eventually, the normalization is a partition function over all the irreducible fractions $q = \frac{a}{b}$ with Hamiltonian $H(\frac{a}{b}) = \log ab$ and inverse temperature σ.

But for the moment, this expression is formal and the issue of convergence needs to be addressed.

3.2. Precise definition from a rational numbers enumeration

The formal expression

$$\sum_{\frac{a}{b}\in\mathbb{Q}^+}(ab)^{-\sigma} = \frac{\zeta(\sigma)^2}{\zeta(2\sigma)} \quad , \tag{17}$$

derived in the appendix, converges as long as the inverse temperature parameter is more than one, $\sigma > 1$ but it diverges otherwise. Even though the divergence in the numerator and in the denominator of Eq. (16) could compensate when taking a limit, different orders of summation could lead to different results because of the lack of absolute convergence.

To disambiguate the definition of the sonance when $\sigma \leq 1$, we need to specify an enumeration of the rational numbers to fix an order of summation in Eq. (16). A simple enumeration could consist in making a list of rational numbers of numerator and denominator smaller than a certain increasing threshold: $a, b < n$. However, the algorithm that would write this list would have to deal with a lot of redundancy. Another strategy could consist in using the prime number factorization to build them without redundancy, but this requires both a threshold on prime numbers and on their integer exponents.

Actually, all rational numbers are enumerated without redundancies in reduced form in two related binary trees: Stern-Brocot tree and Calkin-Wilf tree. Stern-Brocot tree grows vertically while conserving the ordering of the rational number from left to right. Each new stage is generated from the previous one by computing the "mediant" number between two "parents": $\frac{a}{b} < \frac{a+c}{b+d} < \frac{c}{d}$, starting from the stage 0: $\{\frac{0}{1}, \frac{1}{0}\}$. In Calkin-Wilf tree, each rational number generates two new rational numbers in the next growth stage: $\frac{a}{b} \to \{\frac{a}{a+b}; \frac{a+b}{b}\}$, starting from the stage 1: $\{\frac{1}{1}\}$. These trees have many interesting properties, among which we select the ones that serve for the purpose of this paper: in each stage they have the same rational numbers permuted in a different order ; there exist an algorithm (see numerical implementation) that

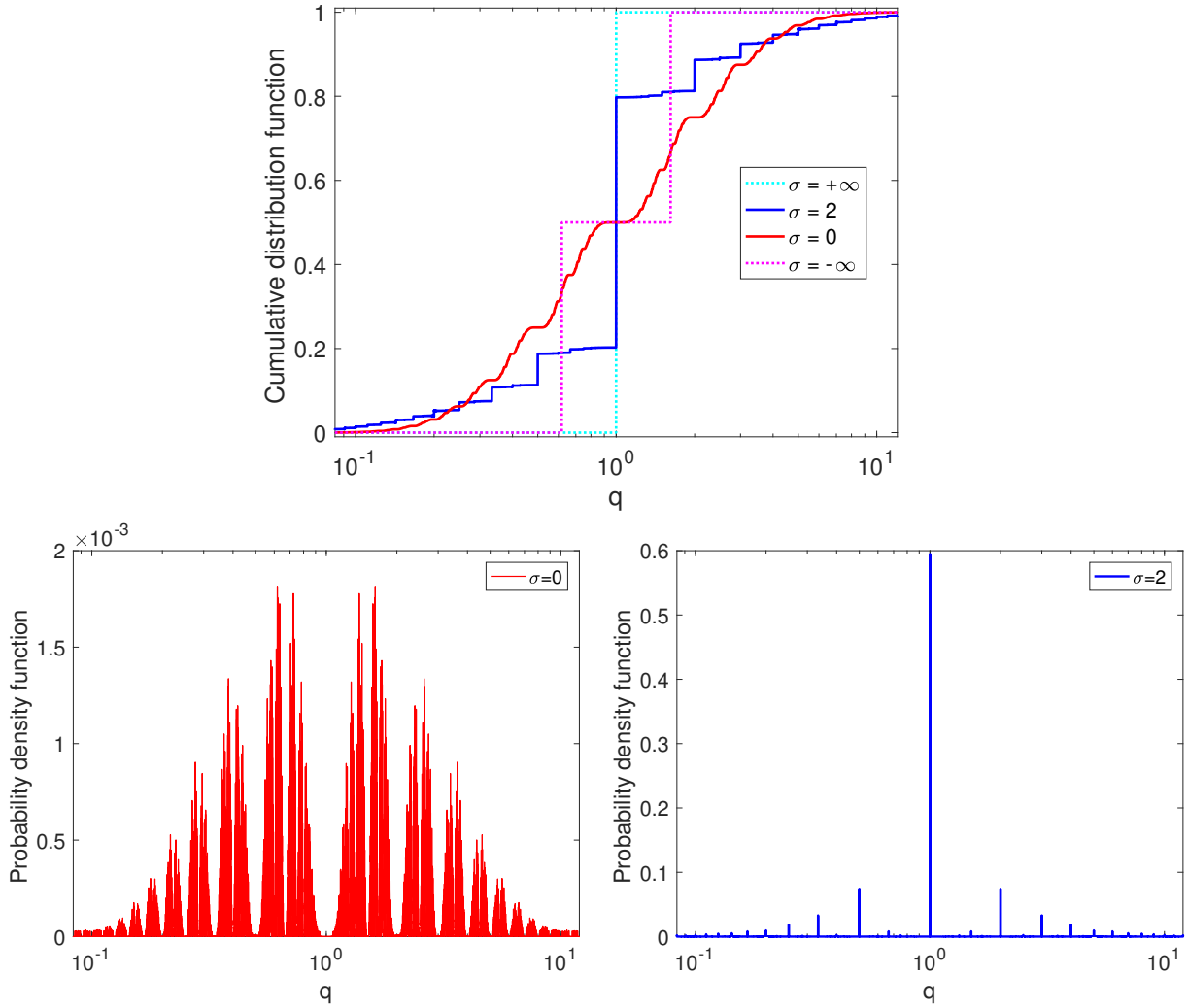


Figure 2: (a) Cumulative distribution function of the measure $\oint_\sigma(q)$ for the parameters $\sigma = -\infty, 0, 2, +\infty$. (b) Probability density function $\oint'_\sigma(q)$ (derivative with respect to $\log q$) for $\sigma = 0$, and (c) for $\sigma = 2$. Note that the probability distribution functions are singular, which makes this representation less suitable since it varies a lot with the numerical resolution, especially for $\sigma = 0$. The measures are better represented by their cumulative distribution functions. The case $\sigma \approx 1$ converges very slowly, therefore it is not shown.

enumerates at each iteration the next rational numbers in a stage, and stage by stage in both trees ; a trajectory in the Stern-Brocot tree is a graphical representation for the continued fraction of a real number, in particular the depth in the tree of a rational number is the sum of the coefficient of its continued fraction sequence. Let us call the depth of the stage in the tree the harmonic depth $h(q)$, finite for each rational number

Introducing a threshold, we define the set of rational numbers of harmonic depth at most h:

$$\mathbb{Q}_h = \{q \in \mathbb{Q}^+ \mid h(q) \leq h\} \quad , \tag{18}$$

which yields a precise definition of the sonance using this order of summation in the limit of a high harmonic depth:

$$\oint_\sigma [R_{1,2}] = \int_0^\infty R_{1,2}(q) \mathrm{d}\oint_\sigma (q) = \lim_{h\to\infty} \frac{\sum_{q\in\mathbb{Q}_h} e^{-\sigma H(q)} R_{1,2}(q)}{\sum_{q\in\mathbb{Q}_h} e^{-\sigma H(q)}} \quad . \tag{19}$$

3.3. Influence of σ: a continuum between rationality and irrationality

In the"low temperature limit", $\sigma \to +\infty$, the amplitude of all the harmonic frequencies vanishes and the sonance measure just evaluates the measured function to $q = 1$:

$$\mathrm{d}\oint_{+\infty} (q) = \delta(\log q)\mathrm{d}q/q = \delta(q-1)\mathrm{d}q \quad .$$

This means that in this limit, only the strength of the $1:1$ synchronization is measured.

Conversely, higher temperature i.e. lower σ would also measure a : b synchronization of higher and higher complexity.

In the "high temperature limit", $\sigma = 0$, there is an equiprobability between all the rational numbers, among which the simple ratios are extremely rare. This can be understood as giving all weight of the probability density to irrational numbers. The sonance measure defined with this enumeration is already known in number theory from more than a century ago. The cumulative distribution function of this singular probability measure restricted to the interval $[0, 1]$ is called *Minkowski's question mark function* $?(x)$ (in fact half of it):

$$\oint_0 (x) = \int_0^x \mathrm{d}\oint_0 (q) = \frac{?(x)}{2} \ , \quad 0 \leq x \leq 1 \quad . \tag{20}$$

First defined by Minkowski as a map from quadratic irrational to rational numbers based on their periodic continued fraction behavior, this function has been investigated more recently as a measure by Linas Vepstas ("Minkowski's measure") and most extensively by Giedrius Alkauskas during his PhD thesis (that also extended it to $\mathbb{R}^+$).

Although its density function is difficult to represent because of its singular nature, the cumulative distribution function has a fractal structure that contains stairs around each rational number, meaning that the probability density gives all its weight to irrational numbers. This is confirmed in the work of Giedrius Alkauskas where the measure can be defined from the same kind of limit but only over the rational numbers in the last stage of the tree (in other word, this is the limiting distribution of the rational numbers in the deep stages of the Stern-Brocot tree $\mathbb{Q}_h \setminus \mathbb{Q}_{h-1}$).

Transition and negative temperatures: Therefore, the sonance measures gives weight to rational numbers for high σ, and to irrational numbers for low σ. We can expect $\sigma = 1$ to be a transition. We observe from numerical simulations that it behaves as $\mathrm{d}\oint_\sigma (q) \sim_{\sigma\to 1^+} \frac{2}{\zeta(\sigma)}\mathrm{d}q/q$.

We also observe that the limit $\sigma \to -\infty$ is also a simple evaluation to the "most irrational numbers", the golden number $\varphi = \frac{\sqrt{5}+1}{2}$ and its inverse:

$$\mathrm{d}\mathscr{C}_{-\infty}(q) = \frac{1}{2}\left(\delta(\log(q\varphi)) + \delta(\log(q\varphi^{-1}))\right)\mathrm{d}q/q = \frac{1}{2}\left(\delta(q-\varphi^{-1}) + \delta(q-\varphi)\right)\mathrm{d}q \quad .$$

This last observation means that (and is equivalent to) the maximum product ab in the tree of rational numbers $\frac{a}{b}$ up to any stage h occurs for the best rational approximation of the golden numbers in this tree. This is proved using Calkin-Wilf tree: to maximize the product of the numerator and the denominator at each growth stage, we have to choose $\frac{a+b}{b} = 1 + \frac{a}{b}$ when $b > a$ and $\frac{a}{a+b} = \frac{1}{1+\frac{b}{a}}$ when $a > b$, i.e. we have to use them alternatively. The successive and alternative applications of these generators from the root $\frac{1}{1}$ yields, in the limit, the continued fraction of the golden numbers $\varphi = 1 + \frac{1}{1+\frac{1}{1+\frac{1}{1+\ldots}}}$ and $\varphi^{-1} = \frac{1}{1+\frac{1}{1+\frac{1}{1+\ldots}}}$.

4. Numerical implementation from the Mellin domain

4.1. From a linear to a geometric frequency sampling

The numerical computation of the spectrum of relations $R_{1,2}(q)$ defined in Eq. (6), i.e. of a multiplicative cross-correlation, is complicated by the fact that, contrary to the usual "additive" one, the sampling frequency of the spectrum $S_2(qf)$ is dilated by q and does not match anymore the sampling of $S_1(f)$. There is an incompatibility between the linear sampling in time and frequency domains and the multiplicative behavior in the integral that would require a geometric sampling in the frequency domain. Actually, the continuous wavelet transform does provide this possibility because the frequency (scaling parameter) sampling is arbitrary in its numerical implementation:

$$\mathcal{W}_\psi[s](f,t) = \mathcal{F}^{-1}[\mathcal{F}[s]\cdot\tilde{\psi}_f](t) \tag{21}$$

where $\tilde{\psi}_f(f') = \tilde{\psi}(f'/f)$ is the dilated analytic wavelet of characteristic frequency f, and the fast Fourier transform algorithm (FFT) is used instead of the Fourier transform. The geometric sampling is even natural since the frequency resolution is proportional to the frequency. Note that the power spectral density can be estimated from the continuous wavelet transform by averaging its squared modulus over time, as in Eq. (5), hence keeping the geometric frequency sampling. This geometric frequency sampling can be viewed as a linear sampling in the log-frequencies. Note that there is no negative frequency because the wavelet is analytic.

4.2. Mellin transform and convolution theorem

The Mellin transform of a function $S(f)$ is defined by

$$\mathcal{M}[S](z) = \int_0^\infty S(f) f^{z-1}\mathrm{d}f \quad . \tag{22}$$

The spectrum of relations: Thanks to the geometric frequency sampling, a FFT in the log-frequency domain is used for a fast computation of the multiplicative cross-correlation:

$$R[S_1,S_2](q) = \mathcal{F}^{-1}\left[\overline{\mathcal{F}[S_1\circ\exp]}\cdot\mathcal{F}[S_2\circ\exp]\right](\log q) \quad , \tag{23}$$

where the change of variable is here expressed with a composition of functions denoted by $g\circ h(x) = g(h(x))$. When using the FFT, we should care avoiding artifacts from periodic

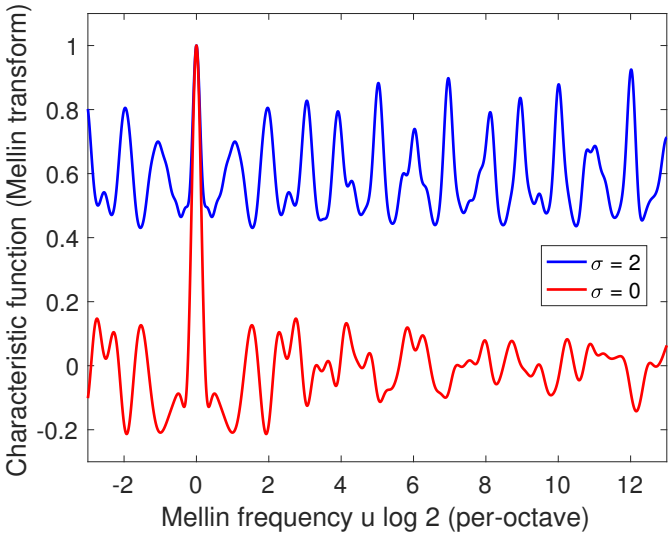


Figure 3: Mellin transform of the measures $\widehat{\mathscr{C}'}_\sigma(u)$, which correspond to the characteristic functions of the probability distribution, for the case $\sigma = 0, 2$.

boundary conditions in the the log-frequency domain. This can be achieved by windowing or zero-padding the spectrum in the log-frequency direction so its length is at least doubled.

Taking the Fourier transform with respect to the log-frequency corresponds to a Mellin transform:

$$\mathcal{F}[S \circ \exp](u) = \int_0^\infty S(f) f^{-i\tau u} \mathrm{d}f/f = \mathcal{M}[S](-i\tau u) \quad , \tag{24}$$

for which the convolution (and cross-correlation) theorem works in the multiplicative case:

$$\mathcal{M}\left[R[S_1, S_2]\right] = \overline{\mathcal{M}[S_1]}\mathcal{M}[S_2] \quad . \tag{25}$$

Since we are especially interested in the frequency-like parameter u (let us call it Mellin frequency), we simply denote it

$$\hat{S}(u) \equiv \mathcal{F}[S \circ \exp](u) = \mathcal{M}[S](-i\tau u) \quad , \tag{26}$$

similarly to the Fourier transform notation.

By replacing the frequency spectra by the general forms in the stationary and non-stationary case proposed above, we get the following expression in the Mellin domain:

$$S(f) = |\tilde{s}(f)\sqrt{f}|^\mu f^\nu \quad \Leftrightarrow \quad \hat{S}(u) = \mathcal{M}[|\tilde{s}(f)|^\mu](\mu/2 + \nu - i\tau u) \tag{27}$$

$$S(f,t) = |\mathcal{W}_\psi[s](f,t)|^\mu f^\nu \quad \Leftrightarrow \quad \hat{S}(u,t) = \mathcal{M}\left[|\mathcal{W}_\psi[s](\cdot,t)|^\mu\right](\nu - i\tau u) \quad . \tag{28}$$

The Mellin transform of a power of a certain function $|X(f)|^\mu$ can only be computed explicitly in the case of μ even (in order to avoid dealing with the absolute values), see a calculation for the case $\mu = 2$ in the appendix.

The measure: The Mellin transform can also serve to accelerate a lot the computation of the sonance measure. Indeed, its definition currently require to interpolate the spectrum of relation

$R_{1,2}(q)$ at the enumerated $q = \frac{a}{b}$ and sum the values. Instead, we can use the Mellin transform of the measure:

$$\widehat{\oint}'_{\sigma}(u) \equiv \int_0^{\infty} q^{-i\tau u} \mathrm{d}\oint_{\sigma}(q) = \int_0^{\infty} \cos(\tau u \log q) \mathrm{d}\oint_{\sigma}(q) \quad , \tag{29}$$

where the dash denotes a derivative with respect to $\log q$. Even though the probability density function $\oint'_{\sigma}(q) = \mathrm{d}\oint_{\sigma}(q)/\mathrm{d}\log q$ is not well defined, it has a well defined Mellin transform that can be interpreted as the associated characteristic function. The symmetry $\mathrm{d}\oint_{\sigma}(q^{-1}) = \mathrm{d}\oint_{\sigma}(q)$ implies that $\widehat{\oint}'_{\sigma}(u)$ is real and can be written as the above cosine transform. Note that the cases $\sigma > 1$ can be computed analytically (see appendix):

$$\widehat{\oint}'_{\sigma}(u) = \frac{|\zeta(\sigma + i\tau u)|^2}{\zeta(\sigma)^2} \quad , \tag{30}$$

as well as the limit cases:

$$\widehat{\oint}'_{-\infty}(u) = \cos(\tau u \log \varphi) \qquad ; \qquad \widehat{\oint}'_{1}(u) = \delta_{u,0} \qquad ; \qquad \widehat{\oint}'_{+\infty}(u) = 1 \quad . \tag{31}$$

In the other cases, the Mellin transform of the measure is estimated as:

$$\widehat{\oint}'_{\sigma}(u) = \frac{\sum_{q\in\mathbb{Q}_h} e^{-\sigma H(q)} \cos(\tau u \log q)}{\sum_{q\in\mathbb{Q}_h} e^{-\sigma H(q)}}, \quad H(q = \tfrac{a}{b}) = \log ab \tag{32}$$

$\mathbb{Q}_h$ defined as in (18) for a high enough harmonic depth h, except for σ close to 1 because the convergence is very slow.

The sonance: As a result, the sonance is expressed in a convenient way with the Plancherel-Parseval theorem from the Mellin domain:

$$\oint_{\sigma}[R_{1,2}] = \int_0^{\infty} R_{1,2}(q) \mathrm{d}\oint_{\sigma}(q) = \int_{-\infty}^{+\infty} \hat{R}_{1,2}(u) \widehat{\oint}'_{\sigma}(u) \mathrm{d}u \quad , \tag{33}$$

so that it can be computed directly from the spectra in the Mellin domain:

$$\oint_{\sigma}[S_1, S_2] = \int_{-\infty}^{+\infty} \overline{\hat{S}_1}(u) \hat{S}_2(u) \widehat{\oint}'_{\sigma}(u) \mathrm{d}u \quad . \tag{34}$$

We use equivalently $\oint_{\sigma}[R_{1,2}]$ or $\oint_{\sigma}[S_1, S_2]$ as a short notation for $\oint_{\sigma}[R[S_1, S_2]]$.

Finally, this result is generalized to the sonances of $R_{1,2}(qq')$, that is the spectrum of relations between $S_1(f)$ and $S_2(qf)$ where S_2 is scaled by any coefficient $q > 0$:

$$\oint_{\sigma}[R_{1,2}](q) \equiv \int_0^{\infty} R_{1,2}(qq') \mathrm{d}\oint_{\sigma}(q') = \int_{-\infty}^{+\infty} \hat{R}_{1,2}(u) q^{i\tau u} \widehat{\oint}'_{\sigma}(u) \mathrm{d}u \tag{35}$$

$$\oint_{\sigma}[S_1, S_2](q) \equiv \int_0^{\infty} \int_0^{\infty} S_1(f) S_2(q'qf) \mathrm{d}f/f \, \mathrm{d}\oint_{\sigma}(q') = \int_{-\infty}^{+\infty} \overline{\hat{S}_1}(u) \hat{S}_2(u) q^{i\tau u} \widehat{\oint}'_{\sigma}(u) \mathrm{d}u \tag{36}$$

where the notation of the sonance has been extended to incorporate this additional scaling parameter. The particular value at $q = 1$ of the sonance function is the original sonance value $\oint_{\sigma}[S_1, S_2](1) = \oint_{\sigma}[S_1, S_2]$.

Acoustically, this can be interpreted as the comparison of two tones of fundamental frequency separated by the ratio q, each with the timbre of a certain instrument, modeled by the spectra S_1, S_2. In this model, an invariance of the spectral shape for lower or higher tones is assumed.

Sonances and spectrum of relation, a single object: In the limit $\sigma \to +\infty$, we notice that the measure is just one so that

$$\oint_{+\infty}[R_{1,2}](q) = R_{1,2}(q) \quad , \tag{37}$$

i.e. the spectrum of relation is the sonances function in the low temperature limit. Also note that the sonances for $\sigma \geq 2$ are very close to $R(q)$ (cf. Fig. 4) because the ratio $q = 1$ has a great weight. Conversely, all sonances functions are Mellin-filtering of the spectrum of relations.

4.3. Sonance in Mellin domain in terms of musical scales

The Mellin transform of the sonance measure $\widehat{\oint}'_\sigma(u)$ has a strong musical significance that justifies the choice of the treble key symbol. Indeed, it measures the periodicity in $\log q$ of the sonance of two pure tones separated by any ratio q, for which the spectrum of relations is just $R_{1,2}(q') = \delta(\log q' - \log q)$. When the tones have an ideal timbre as in Eq. (12), this function is just multiplied by a positive quantity, which make the sign of $\widehat{\oint}'_\sigma(u)$ very important. When expressing the Mellin frequency in per octave unit, we obtain a number of oscillations $n = u \log 2$ per octave. When choosing a musical scale, i.e. a more or less regularly spaced set of frequencies, we would like the musical intervals (i.e. the frequency ratios) in this set to be as consonant as possible. Given that the sonance is high for dissonant musical intervals and low for consonant ones for $\sigma = 0$ (in fact $\sigma < 1$), the negative peaks in $\widehat{\oint}'_0(\frac{n}{\log 2})$ is of great help for this task since it indicates the cardinality of musical scales that provides consonance. Negative peaks are observed at $n = 5, 7, 12$ (per octave), in agreement with the widespread use of the pentatonic, heptatonic and dodecaphonic musical scales all over the world ; they confirm the tight link with musical harmony. More explanation are given in another manuscript.

4.4. Summary on the use of the fast Fourier transform

Given that the wavelet transform is computed in practice from FFT+IFFT, that the wavelet is defined in the frequency domain and that the Mellin transform of the measure is estimated just once as Eq.(32), the computation of the sonance in time requires:

- 1 FFT+wavelet Fourier filter+IFFT on the linearly sampled signals per frequency value,
- 1 FFT+sonance Mellin filter+IFFT on the geometrically sampled spectra per time step.

4.5. Sonance and inequalities: a suggestion for the normalization?

Consider positive non-normalized frequency spectra $S_1(f)$ and $S_2(qf)$.

The sonance of this pair can be written in the Mellin domain:

$$\oint_\sigma[S_1, S_2](q) = \int_{-\infty}^{+\infty} \overline{\hat{S}}_1(u)\hat{S}_2(u)\widehat{\oint}'_\sigma(u)q^{i\tau u}\mathrm{d}u \tag{38}$$

As long as $\sigma \geq 1$, we can apply the Cauchy-Schwarz inequality as follows:

$$\left(\oint_\sigma[S_1, S_2](q)\right)^2 \leq \int_{-\infty}^{+\infty} \left|\overline{\hat{S}}_1(u)\sqrt{\widehat{\oint}'_\sigma(u)}\right|^2 \mathrm{d}u \int_{-\infty}^{+\infty} \left|\hat{S}_2(u)q^{i\tau u}\sqrt{\widehat{\oint}'_\sigma(u)}\right|^2 \mathrm{d}u \tag{39}$$

$$\oint_\sigma[S_1, S_2](q) \leq \sqrt{\oint_\sigma[S_1, S_1]\oint_\sigma[S_2, S_2]} \quad , \tag{40}$$

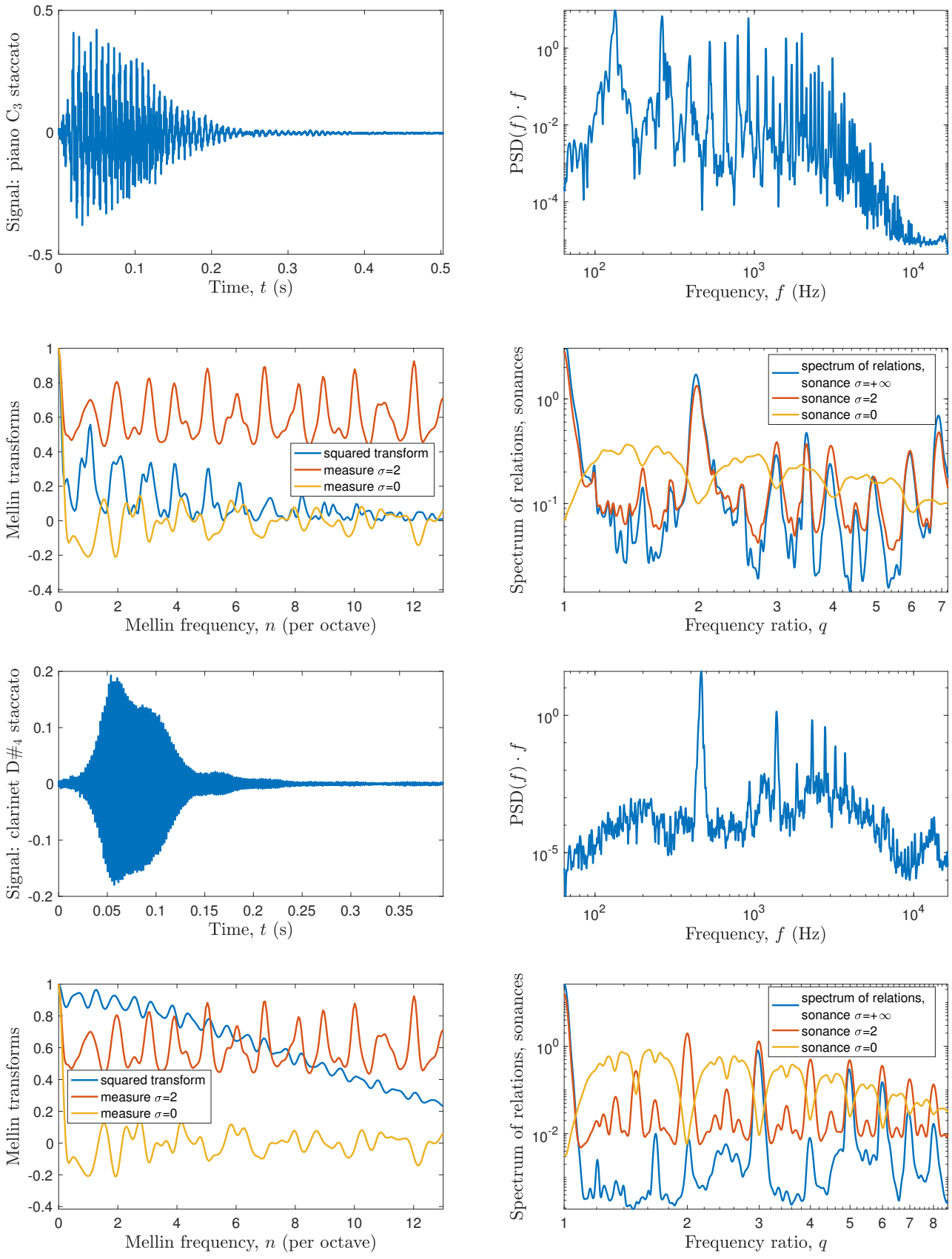


Figure 4: Stationary $\mathscr{C}$-analysis for piano and clarinet sounds: (a) Signal. (b) Frequency spectrum normalized as a probability density function in the convention $\mu = 2, \nu = 0$. (c) Square of the Mellin transform of the frequency spectrum (blue line) and Mellin transforms $\widehat{\mathscr{C}'}_\sigma$ of the measures for $\sigma = 2$ (red) and $\sigma = 0$ (yellow), with respect to the variable $n = u \log 2$ in per octave unit. (d) Spectrum of relations, i.e. distribution of frequency ratios (blue line), sonances for all ratios for $\sigma = 2$ (red) and $\sigma = 0$ (yellow). Note that the spectrum of relations is also the sonances for all ratios for $\sigma = +\infty$.

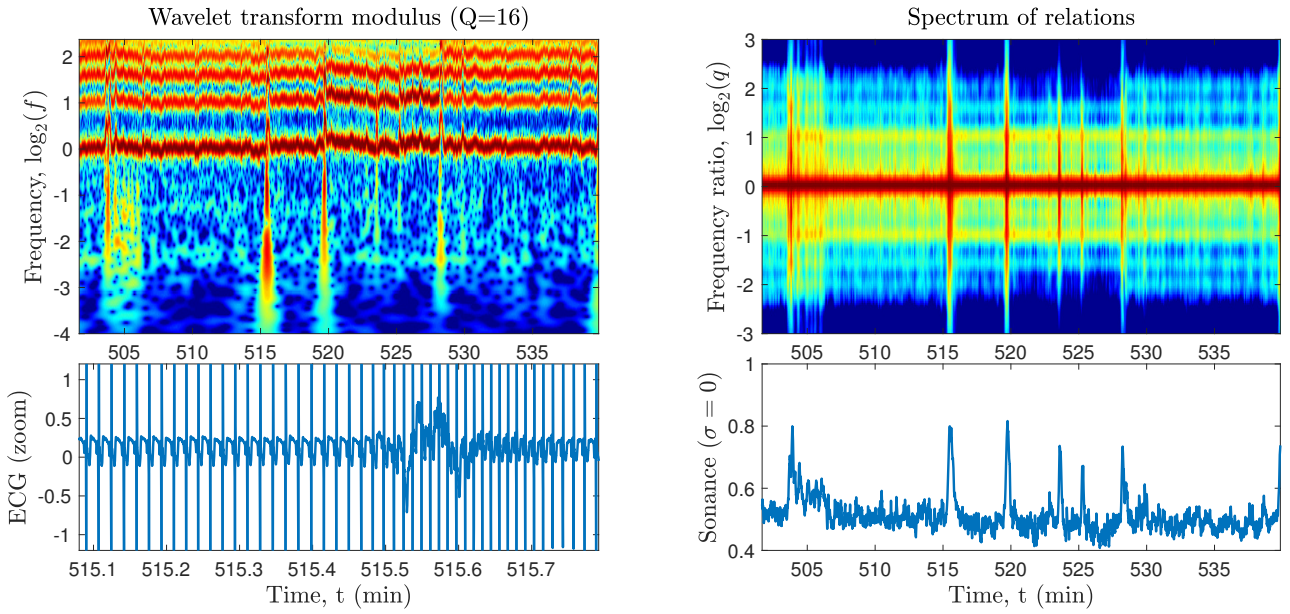


Figure 5: Non-stationary $\oint$-analysis of an electrocardiogram (ECG): (a) Wavelet transform modulus of the ECG with quality factor $Q = \frac{f}{\Delta f} = 16$. (b) Spectrum of relations normalized with the norm 2 of the frequency spectra in the convention $\mu = 1, \nu = 0$. (c) Zoom on the ECG signal around a perturbation. (d) Sonance value ($\sigma = 0$) in time. We observe clear sonance peaks at perturbations.

since $\widehat{\oint}'_\sigma(u) = \frac{|\zeta(\sigma+i\tau u)|^2}{\zeta(\sigma)^2} \geq 0,\ \forall \sigma \geq 1$. Note that the Cauchy-Schwarz inequality would always apply for a measure defined by $\frac{|\zeta(\sigma+i\tau u)|^2}{\zeta(\sigma)^2}$ for all σ, but $\widehat{\oint}'_\sigma(u)$ is different from it as soon as $\sigma < 1$.

We could use this inequality to compare the sonance of a pair of signals and the sonances of each signal with itself,

$$\frac{\oint_\sigma [S_1, S_2](q)}{\sqrt{\oint_\sigma [S_1, S_1] \oint_\sigma [S_2, S_2]}} \quad , \tag{41}$$

like a correlation coefficient. Even though this quantity can be greater than one in the important case of $\sigma = 0$, this can remain a way of comparing signals with very different sonances.

The Young's convolution inequality also applies very well to the definition of the sonance from the log-frequency domain:

$$\oint_\sigma [S_1, S_2](q) = \int_0^\infty \int_0^\infty S_1(f) S_2(qf) \mathrm{d}f/f \, \mathrm{d}\oint_\sigma(q) \tag{42}$$

$$\leq \left(\int_0^\infty S_1(f)^2 \mathrm{d}f/f \int_0^\infty S_2(f)^2 \mathrm{d}f/f \right)^{\frac{1}{2}} \int_0^\infty \mathrm{d}\oint_\sigma(q) \quad . \tag{43}$$

This inequality applies for any σ, which strongly suggests to normalize the spectra by their norm 2 (and to use the musical measure as a probability measure) so that:

$$\oint_\sigma [S_1, S_2](q) \in [0, 1] \tag{44}$$

$$-\log \oint_\sigma [S_1, S_2](q) \in \mathbb{R}^+ \quad . \tag{45}$$

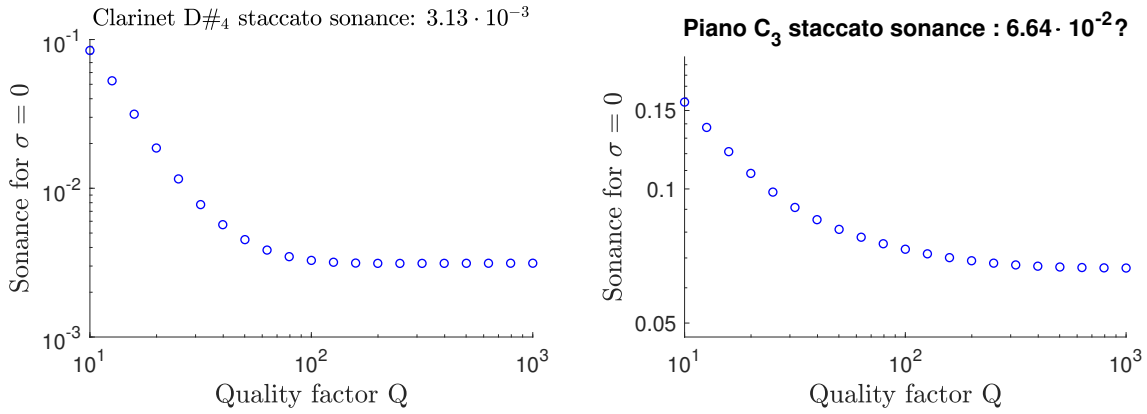


Figure 6: Sonance ($\sigma = 0$ and ratio $q = 1$) for the previous piano and clarinet sounds as a function of the quality factor Q (that determines the frequency resolution as $\Delta \log f = Q^{-1}$). The sonance has a clear saturation value for the clarinet sound, which may also be the case for the piano sound at very high Q.

Depending on ones tastes, we can choose to reverse the sonance continuum as in Eq. (45) so that a dissonance between S_1 and S_2 goes to zero whereas a consonance goes to infinity for $\sigma = 0$. The norm 2 normalization favors the convention $\mu = 1, \nu = 0$ since it corresponds to the normalization of the signal by its energy.

4.6. Generalization to multi-dimensional signals

The generalization of the definitions Eq. (6,16,35) is straightforward: for instance let the spectrum be

$$\vec{S}(f) = \begin{pmatrix} S_1(f) \\ S_2(f) \end{pmatrix} \quad , \tag{46}$$

then, its spectrum of relations writes

$$\mathbf{R}[\vec{S}](q) = \int_0^\infty \vec{S}(f)\vec{S}(qf)^T \mathrm{d}f/f = \begin{pmatrix} R[S_1, S_1](q) & R[S_1, S_2](q) \\ R[S_2, S_1](q) & R[S_2, S_2](q) \end{pmatrix}, \tag{47}$$

and satisfies the symmetry $\mathbf{R}[\vec{S}](q^{-1}) = \mathbf{R}[\vec{S}](q)^T$ (where $\mathbf{R}^T$ is the transpose of $\mathbf{R}$). The sonance

$$\oint_\sigma [\mathbf{R}] = \int_0^\infty \mathbf{R}(q) \mathrm{d} \oint_\sigma (q) = \begin{pmatrix} \oint_\sigma [S_1, S_1] & \oint_\sigma [S_1, S_2] \\ \oint_\sigma [S_2, S_1] & \oint_\sigma [S_2, S_2] \end{pmatrix} \quad , \tag{48}$$

is a symmetric matrix, and the sonance function inherits the symmetry $\oint_\sigma [\mathbf{R}](q^{-1}) = \oint_\sigma [\mathbf{R}](q)^T$.

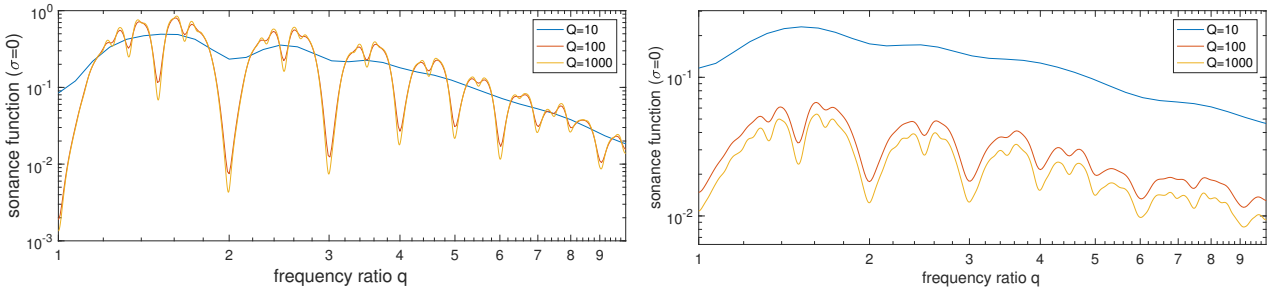


Figure 7: Sonance function ($\sigma = 0$) for the clarinet sound for different quality factor $Q = 10, 100, 1000$ in blue, orange and yellow respectively. The convention in (a) is $\mu = 2$, $\nu = 0$ and the normalization uses the norm 1, whereas (b) shows the same result for $\mu = 1$, $\nu = 0$ and the norm 2. The smoothing of the sonance function as in Eq. (53) holds exactly in (a) only, but the behavior is qualitatively the same in (b).

5. Sonance dependences

5.1. Wavelet influence

Suppose a frequency spectrum $S(f)$ is a smoothed version of $S_0(f)$ with the positive kernel $\phi(q)$:

$$S(f) = \int_0^\infty S_0(f')\phi(f/f')\mathrm{d}f'/f' \tag{49}$$

$$\hat{S}(u) = \hat{S}_0(u)\hat{\phi}(u) \tag{50}$$

$$\hat{R}(u) = |\hat{S}_0(u)|^2|\hat{\phi}(u)|^2 \equiv \hat{R}_0(u)\hat{\Phi}(u) \tag{51}$$

$$\widehat{\mathscr{C}}_\sigma[S,S](u) = \widehat{\mathscr{C}}_\sigma[S_0,S_0](u)|\hat{\phi}(u)|^2 \tag{52}$$

$$\mathscr{C}_\sigma[S,S](q) = \int_0^\infty \mathscr{C}_\sigma[S_0,S_0](q')\Phi(q'/q)\mathrm{d}q'/q' \quad \text{where} \quad \Phi(q) = \int_0^\infty \phi(q')\phi(q/q')\mathrm{d}q'/q' \tag{53}$$

$$\mathscr{C}_\sigma[S,S] = \int_0^\infty \mathscr{C}_\sigma[S_0,S_0](q)\Phi(q)\mathrm{d}q/q \quad , \tag{54}$$

then, the sonance of $S(f)$ is a smoothed version of the sonance of $S_0(f)$ with the positive kernel $\Phi(q)$. This calculation holds for the normalization with the norm 1 of the functions S, S_0, ϕ. In particular, it is adapted to the convention $\mu = 2, \nu = 0$ with a stationary signal and a wavelet such that $\phi = |\tilde{\psi}|^2$ (cf. Eq. (53)). It does not hold for other norms p, but we expect a similar behavior (cf. Fig. 7). Note that the sonances of a spectrum with itself is symmetric about $\log q = 0$.

Denoting the frequency resolution of the wavelet by the inverse quality factor $Q^{-1} = \frac{\Delta f}{f}$, we expect the sonance to saturate as in the Fig. 6 when the quality factor is high enough for the wavelet to fit the intrinsic frequency width of the underlying spectrum S_0.

A. Refined frequency spectrum using the phase of the wavelet transform

The previous approach uses the fact that the wavelet has a unit characteristic frequency to directly interpret the scaling parameter (or its inverse) as the frequency variable of the spectrum. Note that for a well-defined frequency mode in the signal, the scaling parameter that maximizes the amplitude yields the unique characteristic frequency of this mode. The extraction of such maxima lines is used to study singularities (WTMM) or rhythms evolution (amplitude ridges, ref. Lilly). It is known that the phase alone can serve to extract the frequency of the modes when the wavelet used is analytic (phase ridges, Lilly). Instead of extracting some time-frequency lines, the phase in the full time-frequency plane can be used to obtain a spectrum with more contrast.

This leads to the refined definition of the frequency as the rate of variation of the phase of the wavelet transform:

$$F(f,t) = \frac{1}{\gamma}\Im\left\{\frac{\partial}{\partial t}\log \mathcal{W}_\psi[s](f,t)\right\}$$

This is made possible by the fact that the wavelet is analytic and the transform is continuous, but it is numerically expensive.

Then, the histogram of the phase-based frequencies F (for all f at a fixed time t) would provide another definition of a frequency spectrum. In the case of a histogram weighted by the amplitude $|\mathcal{W}_\psi[s](f,t)|$, the result is known as the synchrosqueezed (Daubechies) or frequency-reassigned scalogram (?). The weight can otherwise take the general form described before $S(f,t) = |\mathcal{W}_\psi[s](f,t)|^\mu f^\nu$:

$$\hat{S}(f,t) = \int_0^\infty S(f',t)\delta(\log f - \log F(f',t))\mathrm{d}f'/f'$$

The spectrum of relations is then obtained or computed directly as the histogram of frequency ratio over all pairs of frequency parameters:

$$R[\hat{S}_1,\hat{S}_2](q,t) = \int_0^\infty S_1(f',t)S_2(f'',t)\delta(\log F_1(f',t) - \log qF_2(f'',t))\mathrm{d}f'/f'\mathrm{d}f''/f''$$

These operation can be conceived as a refinement of the spectra in the sense that the effect of the frequency uncertainty, i.e. the frequency width of the wavelet is reduced ("squeezed") in this procedure. A similar procedure could be applied for a time reassignment to reduce the time width of the wavelet.

Also note that this use of the phase echos various methods in the analysis of synchronization that is based on the estimation of an unwrapped phase of the beating phenomenon.

As a last remark, since the frequency uncertainty is neglected, the use of this squeezing can be questionable.

B. Zeta function and sum over rational numbers

For $\sigma > 1$:

$$|\zeta(\sigma + i\tau u)|^2 = \left(\sum_{m=1}^{+\infty} m^{-\sigma+i\tau u}\right)\left(\sum_{n=1}^{+\infty} n^{-\sigma-i\tau u}\right) \tag{55}$$

$$= \sum_{m,n=1}^{+\infty} (mn)^{-\sigma}\left(\frac{m}{n}\right)^{i\tau u} \tag{56}$$

$$= \sum_{\frac{a}{b}\in\mathbb{Q}^+}\sum_{k=1}^{\infty}(abk^2)^{-\sigma}\left(\frac{a}{b}\right)^{i\tau u} \tag{57}$$

$$= \zeta(2\sigma)\sum_{\frac{a}{b}\in\mathbb{Q}^+}(ab)^{-\sigma}\left(\frac{a}{b}\right)^{i\tau u} \tag{58}$$

$$\Rightarrow \qquad \frac{|\zeta(\sigma + i\tau u)|^2}{\zeta(2\sigma)} = \sum_{\frac{a}{b}\in\mathbb{Q}^+}(ab)^{-\sigma}\left(\frac{a}{b}\right)^{i\tau u} \quad . \tag{59}$$

C. Young's convolution inequality

In general, the Young's convolution inequality reads:

$$\int\int f(x)g(x-y)h(y)\mathrm{d}x\mathrm{d}y \leq C_{p,q,r}\left(\int f(x)^p\mathrm{d}x\right)^{\frac{1}{p}}\left(\int g(x)^q\mathrm{d}x\right)^{\frac{1}{q}}\left(\int h(x)^r\mathrm{d}x\right)^{\frac{1}{r}} \tag{60}$$

$$\forall p,q,r \geq 1, \quad \tfrac{1}{p}+\tfrac{1}{q}+\tfrac{1}{r} = 2, \quad C_{p,q,r} = A_pA_qA_r \text{ and } A_p = p^{\frac{2-p}{2p}}(p-1)^{\frac{p-1}{2p}} \quad .$$

Note that $\oint'_\sigma(q)^r \propto \oint'_{r\sigma}(q)$ and $C_{p,q,1} = 1$. Does $\int_0^\infty \oint'_\sigma(q)^r\mathrm{d}q/q$ converge? We have at least the slightly more general inequality:

$$\oint_\sigma [S_1, S_2](q) = \int_0^\infty\int_0^\infty S_1(f)S_2(qf)\mathrm{d}f/f\,\mathrm{d}\oint_\sigma(q) \tag{61}$$

$$\leq \left(\int_0^\infty S_1(f)^p\mathrm{d}f/f\right)^{\frac{1}{p}}\left(\int_0^\infty S_2(f)^q\mathrm{d}f/f\right)^{\frac{1}{q}}\int_0^\infty \mathrm{d}\oint_\sigma(q) \tag{62}$$

$$\forall p,q \geq 1, \quad \tfrac{1}{p}+\tfrac{1}{q} = 1 \quad .$$

D. Long calculations

D.1. Fourier-Mellin relation

In the following, we will need the following formula for the (inverse) Fourier transform of $\theta(f)f^{z-1}$ (where θ is the Heaviside step function) which is also the Mellin transform of $e^{\pm i\tau ft}$:

$$\int_0^\infty e^{\pm i\tau ft}f^z\mathrm{d}f/f = (\mp i\tau t)^{-z}\Gamma(z), \quad 0 < \Re(z) < 1 \quad . \tag{63}$$

From this, we can for instance relate the Mellin transform of the wavelet in time and frequency

domains:

$$\mathcal{M}[\overline{\psi}(\pm)](1-z) = \int_0^\infty \overline{\psi}(\pm v) v^{-z} \mathrm{d}v \tag{64}$$

$$= \int_0^\infty \int_{-\infty}^{+\infty} \overline{\tilde{\psi}}(u) e^{\mp i\tau uv} \mathrm{d}u v^{-z} \mathrm{d}v \tag{65}$$

$$= \int_{-\infty}^{+\infty} \overline{\tilde{\psi}}(u) \int_0^\infty e^{\mp i\tau uv} v^{-z} \mathrm{d}v \mathrm{d}u \tag{66}$$

$$= \int_{-\infty}^{+\infty} \overline{\tilde{\psi}}(u) (\pm i\tau u)^{z-1} \Gamma(1-z) \mathrm{d}u, \quad 0 < \Re(z) < 1 \tag{67}$$

$$= \Gamma(1-z)(\pm i\tau)^{z-1} \int_{-\infty}^{+\infty} \overline{\tilde{\psi}}(u) \frac{\mathrm{d}u}{u^{1-z}}, \quad 0 < \Re(z) < 1 \tag{68}$$

$$= \Gamma(1-z) \left((\pm i\tau)^{z-1} \int_0^\infty \overline{\tilde{\psi}}(u) \frac{\mathrm{d}u}{u^{1-z}} + (\mp i\tau)^{z-1} \int_0^\infty \overline{\tilde{\psi}}(-u) \frac{\mathrm{d}u}{u^{1-z}} \right), \quad 0 < \Re(z) < 1 \tag{69}$$

$$= \Gamma(1-z) \left((\pm i\tau)^{z-1} \mathcal{M}[\overline{\tilde{\psi}}(+)](z) + (\mp i\tau)^{z-1} \mathcal{M}[\overline{\tilde{\psi}}(-)](z) \right), \quad 0 < \Re(z) < 1 \quad . \tag{70}$$

And the converse:

$$\mathcal{M}[\overline{\tilde{\psi}}(\pm)](z) = \int_0^\infty \int_{-\infty}^{+\infty} \overline{\psi}(v) e^{\pm i\tau uv} \mathrm{d}v \, u^z \mathrm{d}u/u \tag{71}$$

$$= \Gamma(z)(\mp i\tau)^{-z} \int_{-\infty}^{+\infty} \overline{\psi}(v) v^{-z} \mathrm{d}v, \quad 0 < \Re(-z) < 1 \tag{72}$$

$$= \Gamma(z) \left((\mp i\tau)^{-z} \mathcal{M}[\overline{\psi}(+)](1-z) + (\pm i\tau)^{-z} \mathcal{M}[\overline{\psi}(-)](1-z) \right), \quad 0 < \Re(z) < 1 \quad . \tag{73}$$

Consistency?

$$\frac{\mathcal{M}[\overline{\psi}(\pm)](1-z)}{\Gamma(1-z)} = (\pm i\tau)^{z-1} \mathcal{M}[\overline{\tilde{\psi}}(+)](z) + (\mp i\tau)^{z-1} \mathcal{M}[\overline{\tilde{\psi}}(-)](z) \tag{74}$$

$$= (\pm i\tau)^{z-1} \Gamma(z) \left((-i\tau)^{-z} \mathcal{M}[\overline{\psi}(+)](1-z) + (i\tau)^{-z} \mathcal{M}[\overline{\psi}(-)](1-z) \right) \tag{75}$$

$$+ (\mp i\tau)^{z-1} \Gamma(z) \left((i\tau)^{-z} \mathcal{M}[\overline{\psi}(+)](1-z) + (-i\tau)^{-z} \mathcal{M}[\overline{\psi}(-)](1-z) \right) \tag{76}$$

$$\frac{\mathcal{M}[\overline{\psi}(\pm)](1-z)\tau}{\Gamma(1-z)\Gamma(z)} = (\pm i)^{z-1} \left((-i)^{-z} \mathcal{M}[\overline{\psi}(+)](1-z) + (+i)^{-z} \mathcal{M}[\overline{\psi}(-)](1-z) \right) \tag{77}$$

$$+ (\mp i)^{z-1} \left((+i)^{-z} \mathcal{M}[\overline{\psi}(+)](1-z) + (-i)^{-z} \mathcal{M}[\overline{\psi}(-)](1-z) \right) \tag{78}$$

$$= 2\sin(\pi z) \mathcal{M}[\overline{\psi}(\pm)](1-z) \tag{79}$$

In other words, does this holds?

$$\Gamma(1-z)\Gamma(z)\sin(\pi z) = \pi \tag{80}$$

True! This is called Euler's reflection formula.

For an analytic wavelet, $\mathcal{M}[\overline{\tilde{\psi}}(-)] = 0$ and when $0 < \Re(z) < 1$, we have the relations:

$$\mathcal{M}[\overline{\tilde{\psi}}(+)](z) = \Gamma(z) \left((-i\tau)^{-z} \mathcal{M}[\overline{\psi}(+)](1-z) + (i\tau)^{-z} \mathcal{M}[\overline{\psi}(-)](1-z) \right) \quad , \tag{81}$$

$$\mathcal{M}[\overline{\psi}(\pm)](1-z) = \Gamma(1-z)(\pm i\tau)^{z-1} \mathcal{M}[\overline{\tilde{\psi}}(+)](z) \quad . \tag{82}$$

For a causal wavelet, $\mathcal{M}[\overline{\psi}(+)] = 0$ and when $0 < \Re(z) < 1$, we have the relations:

$$\mathcal{M}[\overline{\psi}(-)](1-z) = \Gamma(1-z)\left((\pm i\tau)^{z-1}\mathcal{M}[\overline{\tilde{\psi}}(+)](z) + (\mp i\tau)^{z-1}\mathcal{M}[\overline{\tilde{\psi}}(-)](z)\right) \quad , \tag{83}$$

$$\mathcal{M}[\overline{\tilde{\psi}}(\pm)](z) = \Gamma(z)(\pm i\tau)^{-z}\mathcal{M}[\overline{\psi}(-)](1-z) \quad . \tag{84}$$

We notice here that we can not have a wavelet both analytic and causal.

D.2. Explicit Mellin transforms for $\mu = 2$?

The Mellin transform of a power of a certain function $|Q(f)|^\mu$ can only be computed analytically in the case $\mu = 2$ by using the Fourier transform:

$$\mathcal{M}[|Q(f)|^2](-i\tau u) = \int_{-\infty}^{+\infty} \overline{\mathcal{M}[Q](-i\tau u')}\mathcal{M}[Q](-i\tau(u'+u))\mathrm{d}u' \tag{85}$$

$$= \int_{-\infty}^{+\infty} \mathcal{M}[\overline{Q}](i\tau u')\mathcal{M}[Q](-i\tau(u'+u))\mathrm{d}u' \quad , \tag{86}$$

therefore:

$$\mathcal{M}[|\tilde{s}(f)|^2 f^{1+\nu}](-i\tau u) = \int_{-\infty}^{+\infty} \mathcal{M}[\overline{\tilde{s}}](\tfrac{1+\nu}{2} + i\tau u')\mathcal{M}[\tilde{s}](\tfrac{1+\nu}{2} - i\tau(u'+u))\mathrm{d}u' \quad , \tag{87}$$

$$\mathcal{M}[|\mathcal{W}_\psi[s](f,t)|^2 f^\nu](-i\tau u) = \int_{-\infty}^{+\infty} \mathcal{M}[\overline{\mathcal{W}_\psi[s](\cdot,t)}](\tfrac{\nu}{2} + i\tau u')\mathcal{M}[\mathcal{W}_\psi[s](\cdot,t)](\tfrac{\nu}{2} - i\tau(u'+u))\mathrm{d}u' \quad . \tag{88}$$

By using Eq (63), we can write in the time domain:

$$\mathcal{M}[|\tilde{s}(f)|^2 f^{1+\nu}](-i\tau u) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \overline{s(t)}s(t')\frac{\Gamma(1+\nu-i\tau u)}{(i\tau(t'-t))^{1+\nu-i\tau u}}\mathrm{d}t\mathrm{d}t', \quad -1 < \nu < 0 \quad , \tag{89}$$

$$= \Gamma(1+\nu-i\tau u)\int_{-\infty}^{+\infty} \overline{s(t)}\int_{-\infty}^{+\infty} s(t+\tau)\frac{\mathrm{d}\tau}{(i\tau\tau)^{1+\nu-i\tau u}}\mathrm{d}t, \quad -1 < \nu < 0 \quad , \tag{90}$$

which is also, interestingly, a Mellin transform from the time domain. Note that the condition on the real part of z, here $1+\nu$, gives the range $-1 < \nu < 0$, which means that this formula only holds for the case $\mu = 2, \nu = -\frac{1}{2}$ previously proposed. The translation invariance is clear under this form.

For the non-stationary case (with an analytic wavelet):

$$\mathcal{M}[\mathcal{W}_\psi[s](\cdot,t)](z) = \int_0^\infty\int_0^\infty \tilde{s}(f')\overline{\tilde{\psi}}(f'/f)e^{i\tau f' t}\mathrm{d}f' f^z \mathrm{d}f/f \tag{91}$$

$$= \int_0^\infty \tilde{s}(f')e^{i\tau f' t} f'^z \mathrm{d}f' \int_0^\infty \overline{\tilde{\psi}}(v)v^{-z}\mathrm{d}v/v \tag{92}$$

$$= \mathcal{M}[\tilde{s}(f)e^{i\tau f t}](1+z)\mathcal{M}[\overline{\tilde{\psi}}](-z) \tag{93}$$

$$= \int_{-\infty}^{+\infty} s(t')\int_0^\infty e^{i\tau f'(t-t')} f'^z \mathrm{d}f'\mathrm{d}t'\mathcal{M}[\overline{\tilde{\psi}}](-z), \quad 0 < \Re(1+z) < 1 \tag{94}$$

$$= \int_{-\infty}^{+\infty} s(t')\frac{\Gamma(1+z)}{(i\tau(t'-t))^{1+z}}\mathrm{d}t'\mathcal{M}[\overline{\tilde{\psi}}](-z), \quad -1 < \Re(z) < 0 \tag{95}$$

$$= \Gamma(1+z)\int_{-\infty}^{+\infty} s(t+\tau)\frac{\mathrm{d}\tau}{(i\tau\tau)^{1+z}}\mathcal{M}[\overline{\tilde{\psi}}](-z), \quad -1 < \Re(z) < 0 \quad . \tag{96}$$

An important comment here is the following: the Mellin transform factorizes the wavelet transform as the product of transforms of the signal (centered at time t) and the wavelet. This transform combines the Fourier and Mellin successive transforms, hence we could call it Fourier-Mellin transform and write it with the symbol $\mathcal{MF}$:

$$\mathcal{M}[\mathcal{W}_\psi[s](\cdot,t)](z) = \mathcal{MF}[s(t+\cdot)](1+z)\overline{\mathcal{MF}[\psi](-\overline{z})}, \quad -1 < \Re(z) < 0 \tag{97}$$

$$= \mathcal{MF}[s(t+\cdot)](1+z)\mathcal{MF}[\overline{\psi}(-)](-z), \quad -1 < \Re(z) < 0 \quad . \tag{98}$$

As computed previously, we have for $-1 < \Re(z) < 0$:

$$\mathcal{MF}[s(t+\cdot)](1+z) = \int_0^\infty \int_{-\infty}^{+\infty} s(t+\tau)e^{-i\tau f\tau}\mathrm{d}\tau\, f^{1+z}\mathrm{d}f/f \tag{99}$$

$$= \Gamma(1+z)\int_{-\infty}^{+\infty} s(t+\tau)\frac{\mathrm{d}\tau}{(i\tau\tau)^{1+z}}, \quad 0 < \Re(1+z) < 1 \tag{100}$$

$$= \Gamma(1+z)\left((i\tau)^{-1-z}\mathcal{M}[s(t+\cdot)](-z) + (-i\tau)^{-1-z}\mathcal{M}[s(t-\cdot)](-z)\right) \quad . \tag{101}$$

As before, we can try to rewrite in the time domain:

$$\mathcal{M}[|\mathcal{W}_\psi[s](f,t)|^2 f^\nu](-i\tau u) = \int_0^\infty \int_0^\infty \overline{\tilde{s}}(f_1)\tilde{s}(f_2)\tilde{K}(f_1,f_2,\nu-i\tau u)e^{i\tau(f_2-f_1)t}\mathrm{d}f_1\mathrm{d}f_2 \tag{102}$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \overline{s}(t+\tau_1)s(t+\tau_2)K(\tau_1,\tau_2,\nu-i\tau u)\mathrm{d}\tau_1\mathrm{d}\tau_2 \tag{103}$$

$$\text{where } K(\tau_1,\tau_2,z) = \int_0^\infty \psi(f\tau_1)\overline{\psi}(f\tau_2)f^{1+z}\mathrm{d}f \quad , \tag{104}$$

$$\tilde{K}(f_1,f_2,z) = \int_0^\infty \tilde{\psi}(f_1/f)\overline{\tilde{\psi}}(f_2/f)f^z\mathrm{d}f/f \quad . \tag{105}$$

D.3. Analytic wavelets in the Mellin domain

Define the log-normal wavelet in the Fourier domain, parametrized by the quality factor Q:

$$\tilde{\psi}_Q(f'/f) = e^{-\frac{1}{2}(Q\log f'/f)^2} \quad , \tag{106}$$

then it take the form

$$\Psi_Q(u) = \mathcal{M}[\tilde{\psi}_Q](-i\tau u) = \frac{\sqrt{\tau}}{Q}e^{-\frac{1}{2}(\frac{\tau u}{Q})^2} \tag{107}$$

$$\mathcal{M}[\tilde{\psi}_Q](z) = \frac{\sqrt{\tau}}{Q}e^{\frac{1}{2}(\frac{z}{Q})^2} \quad . \tag{108}$$

We see that the wavelet acts as a low-pass filter of cut-off Q in the Mellin domain.

When we choose the more general Morse wavelet parametrized by (n,γ), we obtain the following Mellin domain filter:

$$\tilde{\phi}_{n,\gamma}(f'/f) = (f'/f)^n e^{-(f'/f)^\gamma} \tag{109}$$

$$\Phi_{n,\gamma}(u) = \mathcal{M}[\tilde{\phi}_{n,\gamma}](-i\tau u) = \frac{1}{\gamma}\Gamma\left(\frac{n-i\tau u}{\gamma}\right) \tag{110}$$

$$\mathcal{M}[\tilde{\phi}_{n,\gamma}](z) = \frac{1}{\gamma}\Gamma\left(\frac{n+z}{\gamma}\right) \quad . \tag{111}$$

If we impose the characteristic frequency of the Morse wavelet to be the unit one, and imposing $\tilde{\psi}_{n,\gamma}(1) = 1$ as the log-normal wavelet, the rescaled version $\tilde{\psi}_{n,\gamma}(f) = \left(e^{\frac{1-f^\gamma}{\gamma}} f\right)^n$ yields:

$$\Psi_{n,\gamma}(u) = \mathcal{M}[\tilde{\psi}_{n,\gamma}](-i\tau u) = \frac{e^{\frac{n}{\gamma}}}{\gamma} \left(\frac{n}{\gamma}\right)^{\frac{n-i\tau u}{\gamma}} \Gamma\left(\frac{n-i\tau u}{\gamma}\right) \tag{112}$$

$$\mathcal{M}[\tilde{\psi}_{n,\gamma}](z) = \frac{e^{\frac{n}{\gamma}}}{\gamma} \left(\frac{n}{\gamma}\right)^{\frac{n+z}{\gamma}} \Gamma\left(\frac{n+z}{\gamma}\right) \quad . \tag{113}$$

Note that $\tilde{\psi}_{n,\gamma} \longrightarrow \tilde{\psi}_Q$ in the scaling limit $n \longrightarrow \infty, \gamma \longrightarrow 0, n\gamma = Q^2$.

From these expressions, we can compute the kernel K of the previous section:

$$\tilde{K}(f_1, f_2, z) = \int_0^\infty \tilde{\psi}(f_1/f)\overline{\tilde{\psi}}(f_2/f) f^z \mathrm{d}f/f \tag{114}$$

$$= \frac{\sqrt{\pi}}{Q} e^{\frac{1}{4}\left(-(Q\log f_1/f_2)^2 + \frac{z^2}{Q^2} + 2z\log(f_1 f_2)\right)} \quad \text{for } \tilde{\psi}_Q \tag{115}$$