LE CONCEPT DE FRÉQUENCE : REPÈRES HISTORIQUES ET MÉTHODOLOGIQUES POUR LA RECHERCHE MUSICALE

Vincent Lostanlen
Nantes Université,
École Centrale Nantes, CNRS,
LS2N, UMR 6004,
F-44000 Nantes, France

Florian Hecker
Akademie der Bildenden Künste München
80799 Munich, Allemagne

RÉSUMÉ

La synthèse FM, la synthèse modale, le cepstre, les sons de Shepard et la décomposition modale empirique sont des outils mathématiques en informatique musicale qui rompent radicalement avec la notion ordinaire de fréquence, comprise comme l'inverse d'une période entre évènements consécutifs identiques. De fait, ces outils traitent respectivement la fréquence comme une quantité irrationnelle, à valeurs complexes, multivaluée et variable dans le temps. Dans cet article, nous réalisons une brève étude historique du concept de fréquence de Mersenne à Helmholtz, puis relions cette étude à des travaux de recherche–création musicale contemporaine. Empruntant à l'épistémologie historique de Gaston Bachelard, nous défendons une pratique "surrationaliste" de la fréquence qui critique l'interprétation essentialiste de la théorie des sons partiels. Nous présentons un modèle acoustique non linéaire, la diffusion en ondelettes temps–fréquence (JTFS pour *joint time–frequency scattering transform*), donnant l'opportunité d'un nouvel effet audionumérique, le retournement de spin, qui relativise l'orientation du plan temps–fréquence. Nous discutons l'usage créatif du renversement de spin à partir d'une installation sonore multicanal, *Syn 21845 (HEFT Sequenz)* (2022) et d'une œuvre sur format CD, *Scattering Impression Ear Brain Dance Mix* (2024).

1. INTRODUCTION

Le concept de fréquence ne doit pas être réduit à l'inverse d'une durée entre évènements consécutifs. Pourtant, quand nous tâchons d'imaginer le rôle de la fréquence dans un système physique, des images naïves font obstacle à notre compréhension. Vue depuis le rivage, la mer dessine un mouvement de flux et de reflux, quasi périodique, dont les crêtes et les creux marquent l'opposition de phase. Cette image naïve de l'onde comme d'un va-et-vient, acquise peut-être dès le confort du berceau, renforcée ensuite par la démarche bipède, tend à valoriser le mouvement pendulaire dans nos explications mécaniques.

Or, l'informatique musicale a rompu avec cette image ancienne, en vagues successives, des signaux pourvus d'une fréquence. L'invention de la synthèse par modulation

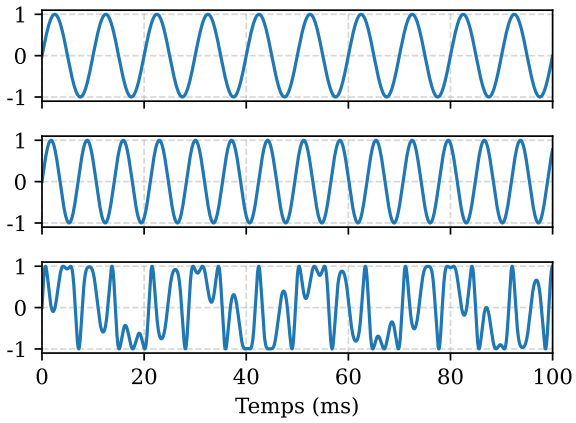

Figure 1. Haut : onde porteuse de fréquence $f_c = 100\,\mathrm{Hz}$. Milieu : onde modulante de fréquence $f_m = \sqrt{2} f_c \approx 141.4\,\mathrm{Hz}$. Base : signal FM (voir Équation 1) avec $\beta = 2$.

de fréquence (synthèse FM, (18)) est emblématique :

$$x(t) = \sin(2\pi f_\mathrm{c} t + \beta \sin(2\pi f_\mathrm{m} t)). \tag{1}$$

Avec une fréquence de modulation supérieure ou égale à la fréquence porteuse ($f_\mathrm{m} \geq f_\mathrm{c}$), on n'entend pas nécessairement de hauteur tonale qui corresponde à f_m. Qui plus est, si le rapport $f_\mathrm{m}/f_\mathrm{c}$ est irrationnel, la pseudo-période $1/f_\mathrm{m}$ n'est pas discernable sur la forme d'onde : voir Figure 1. De fait, on obtient un spectre inharmonique pour x. John Chowning a choisi un rapport $f_\mathrm{m}/f_\mathrm{c} = \sqrt{2} \approx 1.414$ dans sa pièce *Turenas* (1972), puis un rapport $f_\mathrm{m}/f_\mathrm{c} = (\sqrt{5}+1)/2 \approx 1.618$ dans *Stria* (1977), *Phonē* (1981), et *Voices* (2005) (19). Empruntant à l'histoire des sciences, nous pouvons réécouter ces pièces comme autant de jalons d'un parcours de pratique critique, dans lequel l'objet livré à la critique est l'oscillateur commandé en tension (VCO pour *voltage-controlled oscillator*).

L'exemple, bien connu, de la synthèse FM nous invite à un effort d'historicisation du concept de fréquence, effort qui ne doit pas renier sa visée polémique ni son caractère essentiellement provisoire. Dans cet article, nous voulons poursuivre cet effort grâce aux outils mathématiques de l'analyse temps–fréquence et ceux, informatiques, du calcul différentiable. Simultanément, nous situons

notre étude du concept de fréquence dans le contexte contemporain de la création numérique, en présentant deux œuvres nouvelles : une installation sonore—*Syn 21845 (HEFT Sequenz)*, 2022—et une œuvre stéréo—*Scattering Impression Ear Brain Dance Mix*, 2024.

Dans leur ensemble, nos contributions s'engagent en faveur d'une pratique de recherche–création que nous avons baptisée "relativisme spectrotemporel critique" (RSC). D'emblée, nous devons faire une clarification terminologique. Le RSC ne constitue pas une défense du "relativisme" en général, un vocable imprécis qui recouvre des thèses ontologiques (Quine), linguistiques (Whorf) ou encore conceptuelles (Putnam) (26). Étranger à ces philosophies, le R dans RSC s'attache plutôt à la théorie einsteinienne de la relativité, et plus encore au mot allemand *Relativierung*, que Maurice Solovine a traduit par "relativation" et que Gaston Bachelard a repris à son compte dans *La valeur inductive de la relativité* (6). Par ailleurs, dans *La Philosophie du non* (1940), Bachelard qualifie de "critique" la "description des substances obtenues par synthèse" (7). Dès lors, le C dans RSC ne revendique pas tant une parenté avec la "théorie critique" de l'École de Francfort (Adorno, Horkheimer) qu'avec la maxime bachelardienne : "il faut multiplier les réalisations" (p. 56). Nos "réalisations", c'est-à-dire nos œuvres, scandent un parcours critique : l'extension du concept de fréquence n'y est dégagée ni axiomatiquement ni expérimentalement, mais pas une série conjointe d'inventions, de calculs et d'écoutes.

2. ASPECTS HISTORIQUES

2.1. Astronomie babylonienne

Aux IVe et IIIe siècles avant notre ère, la tabulation babylonienne des cycles lunaires contient une série temporelle, dite "colonne Φ", qui forme une ligne brisée en zigzag entre deux valeurs extrêmes (16). La colonne Φ procède d'un algorithme récursif en temps discret. Or, si une historiographie présentiste a permis d'en calculer la période (non entière) (47, p. 112), le contexte scientifique de l'époque plaide plutôt pour une conception en va-et-vient, d'après l'image poétique déjà évoquée.

2.2. Révolution copernicienne

Copernic, Galilée et Kepler ont profondément remodelé notre conception de l'espace, remplaçant la hiérarchie aristotélicienne des objets célestes dans le cosmos par un univers homogène et infini, sujet à des lois uniformes et à la géométrie euclidienne (31). Mais on ne peut pas en dire de même de notre conception du temps et de la fréquence : l'orbite elliptique des planètes, telle que les lois de Kepler la décrivent, est fonctionnellement équivalente à un mouvement alternatif de même période. (Faire du vélo suffit à s'en convaincre.) Notre hypothèse est que l'image ancienne de l'onde comme va-et-vient a survécu à la révolution copernicienne et est restée dominante en Europe jusqu'au début du XVIIe siècle. C'est cette même image que convoque, en 1636, Marin Mersenne dans la préface de son *Harmonie universelle*, comparant une cloche vibrante à un pendule simple allant "deçà-delà" :

> *[...] lorsque les parties d'une cloche frémissent, et vont souvent deçà-delà avant de se reposer, ce mouvement arrive à cause de la trop grande impression que chaque partie s'imprime et se donne à soi-même pour se remettre dans son lieu naturel, comme la boule qui tombe ou qui est suspendue à une corde s'ébranle trop fort elle-même pour demeurer en son centre dès son premier retour.* (45)

Si notre hypothèse est juste, la formation du concept de fréquence à l'époque moderne n'est pas annoncée, même en germe, par l'essor du télescope. Mersenne, bien que lecteur de Galilée et des astronomes de son temps, reste fidèle à l'image archaïque du mouvement en va-et-vient quand il énonce les lois d'échelle de la fréquence fondamentale ξ d'une corde vibrante : $L \propto \xi^{-1}$ (longueur), $T \propto \xi^2$ (tension), $\mu \propto \xi^{-2}$ (masse linéique).

2.3. Calcul infinitésimal

C'est plus tard au XVIIe siècle, avec Newton et Leibniz, que le concept de fréquence acquiert la forme mathématisée que nous lui connaissons aujourd'hui (25). Dans un oscillateur harmonique, le déplacement relatif x est associé à une force centrale $F = -kx = m\ddot{x}$. D'où l'équation différentielle : $\ddot{x} + \omega_0^2 x = 0$, dans laquelle $\omega_0 = \sqrt{k/m}$ est la fréquence d'oscillation en radians par seconde (12).

Outre le calcul infinitésimal, une autre invention importante pour l'analyse des phénomènes périodiques est l'exponentielle complexe : $x = |x|e^{\mathrm{i}\varphi}$. Triomphe du rationalisme classique, la formule d'Euler : $e^{\mathrm{i}\varphi} = \cos\varphi + \mathrm{i}\sin\varphi$, indique un changement de variables *réversible* entre coordonnées cartésiennes et coordonnées polaires. La phase n'y est plus tenue comme le résultat de l'observation instrumentale d'une série temporelle mais comme une variable mathématique à part entière, en égale dignité avec les parties réelle et imaginaire du nombre :

$$\cos\varphi = \frac{e^{\mathrm{i}\varphi} + e^{-\mathrm{i}\varphi}}{2} \qquad \sin\varphi = \frac{e^{\mathrm{i}\varphi} - e^{-\mathrm{i}\varphi}}{2\mathrm{i}}. \quad (2)$$

Dans le cas de l'oscillateur harmonique, poser $\boldsymbol{x}(t) = ae^{\mathrm{i}(\omega t + \varphi_0)}$ mène à $\omega = \pm\omega_0$. Autrement dit, les solutions sinusoïdales sont *caractérisées* par le paramètre physique ω_0^2, homogène au carré d'une fréquence. Ce paramètre est observable sur chaque intervalle de temps infinitésimal $[t, t + \mathrm{d}t]$ comme l'opposé du rapport entre accéleration $\ddot{x}$ et déplacement x. À ce stade de l'enquête, l'image indistincte du va-et-vient est clarifiée et réduite à un seul term : φ_0, la phase à l'origine, rien de plus qu'une *condition initiale* pour le système en oscillation libre. Depuis lors, l'expression $e^{\mathrm{i}\omega t}$ sert d'archétype mathématique à tout mouvement périodique.

2.4. Séries de Fourier

Les ouvrages de Brook Taylor, *Methodus incrementorum directa et inversa* (1715), et de Joseph Fourier, *Mémoire sur la théorie analytique de la chaleur* (1822), servent de bornes historiques approximatives pour ce qu'Alain Herreman a nommé la "controverse des cordes vibrantes" (27) ; c'est-à-dire, la description du mouvement continu d'une corde pincée selon une condition initiale arbitraire. Cette controverse se heurte au principe de superposition : pour deux solutions x_1 and x_2 d'une même équation des ondes, $(c_1x_1 + c_2x_2)$ est solution pour toutes constantes c_1, c_2. Après Euler, d'Alembert et Bernoulli, c'est Joseph Fourier qui accomplit un pas décisif en représentant un signal de période $T = 1/\xi$ par une série trigonométrique

$$\boldsymbol{x}(t) = \sum_{p=1}^{+\infty} a_p \cos(2\pi p\xi t + \varphi_p) \tag{3}$$

dont les coefficients sont donnés par la formule

$$a_p e^{\mathrm{i}\varphi_p} = 2\xi \int_0^{1/\xi} \boldsymbol{x}(t) e^{-2\pi \mathrm{i} p\xi t}\, \mathrm{d}t. \tag{4}$$

(Nous en avons légèrement modernisé les notations (35)). Les deux équations ci-dessus allient la simplicité des composantes sinusoïdale avec la généralité de description du développement en série infinie. Or Fourier, en omettant de spécifier des hypothèses d'intégrabilité pour $\boldsymbol{x}$ (qu'il décrit sans détail comme "l'ordonnée d'une ligne courbe tracée arbitrairement"), exagère la portée de sa formule. Mais la victoire est remportée, et la controverse des cordes vibrantes change de sens : plutôt que de se demander si l'explication physique de Fourier est vraie ou fausse, on se met en quête d'une caractérisation des signaux $\boldsymbol{x}$ dont la série de Fourier converge. Cette quête culmine au XXe siècle avec la théorie des distributions tempérées (53) : nous renvoyons à (29) pour un aperçu historique.

2.5. Psychophysique

L'influence des séries de Fourier en acoustique est reconnue explicitemnt par Ohm en 1843 (49, p. 519) :

> *[...] gebrauche ich das durch seine vielfachen und wichtigen Anwendongen berühmt gewordene Theorem von Fourier [...].*

Ohm propose de modéliser l'oreille humaine comme un banc de filtres qui, à la manière de Fourier, mesure une fréquence fondamentale ξ et des coefficients d'amplitude a_p pour les sons partiels aux fréquences $p\xi$ avec p entier ; sans, toutefois, retenir leurs phases φ_p (46). Ce modèle est toujours en vigueur aujourd'hui, à tel point qu'il est rare de trouver un ouvrage à destination des ingénieur-e-s du son qui ne se réfère, tôt ou tard, aux "fréquences" d'un son, au pluriel, comme des éléments essentiels et mutuellement indépendants. À partir d'Ohm, les harmoniques $p\xi$ et leurs amplitudes a_p sont tenues pour des qualités premières, comme les côtés d'un polygone.

Pourtant, l'interprétation psychophysique de la théorie de Fourier a essuyé des critiques dès 1843 et l'article de Seebeck *Über die Sirene*, qui construit un contre-exemple : un son dont les partiels d'ordre élevé renforcent l'intensité perçue de la fondamentale (56). Le différend entre Ohm et Seebeck a conduit Helmholtz à étudier de près l'existence des sons résultants, pour *in fine* donner raison à la "loi d'Ohm" ; à savoir, la croyance selon laquelle l'oreille opère une analyse de Fourier. Mais comme en mathématiques, la réhabilitation psychoacoustique de Fourier (via Ohm) par Helmholtz s'est faite au prix d'une complexification des concepts de base, à commencer par celui de "son pur". De fait, celui-ci n'est plus donné par simple résonance sympathique, comme chez Mersenne : au contraire, il met en jeu des processus cognitifs de haut niveau, que Helmholtz nomme "théorie des signes" *(Zeichentheorie)* et "inférence inconsciente". Le principe de superposition, hérité du siècle précédent, s'en voit révoqué. La fréquence n'est plus une qualité première mais une qualité seconde (9). Par ce geste radical, Helmholtz inaugure un partage programmatique entre l'oreille coporelle *(das köperliches Ohr)* et l'oreille mentale *(das geistiges Ohr)* (15).

3. ASPECTS MÉTHODDOLOGIQUES

3.1. Relativation *(Relativierung)*

De notre brève étude historique, il ressort que le concept de fréquence n'a rien d'évident : bien plutôt, celui-ci est construit par l'enquête scientifique, et sa formation est scandée par des moments de crise. Nous avons situé une étape décisive de cette formation au milieu du XVIIe siècle, avec l'avènement du calcul infinitésimal. Dans un paradigme newtonien, le carré d'une fréquence (ω^2) se déduit de grandeurs physiques, typiquement par des lois de proportion : pour un ressort, du rapport entre la raideur et la masse ; pour un pendule simple, du rapport entre la pesanteur et la longueur de la tige, sous une approximation aux petits angles. Point crucial, de telles grandeurs sont mesurables *au repos*, en l'absence même d'oscillation périodique. De plus, la formule est empiriquement réversible : avec la loi de Mersenne, par exemple, on peut inférer la tension d'une corde « à l'oreille » à partir de la fréquence fondamentale, pourvu que la masse linéique et la longueur soient connues par ailleurs.

En termes bachelardiens, la mécanique rationnelle de Newton instaure une *solidarité notionnelle*. « La notion de masse », écrit Bachelard, « se définit alors dans un *corps de notions* et non plus seulement comme un élément primitif d'une expérience immédiate et directe » (7, p. 27). Nous sommes tentés d'en dire autant de la notion de fréquence, dont la mathématisation par calcul infinitésimal peut être interprétée comme une relativisation ; ou mieux, comme une *relativation*, suivant la traduction que Solovine donne du terme allemand *Relativierung* chez Einstein (21). Bachelard définit les relativations comme des « extensions d'une notion qui était jadis entièrement solidaire de son caractère sensible » (6, p. 100). Ainsi, pour les Babyloniens, la fréquence demeure solidaire du va-et-vient ; pour Euler,

et *a fortiori* pour Fourier, elle devient le symbole ω dans une expression du type $e^{\mathrm{i}(\omega t+\varphi_0)}$, où la dérivation temporelle équivaut à une multiplication par $\mathrm{i}\omega$. Après relativation, cette expression est découplée de la phase initiale φ_0 : son existence n'est plus interprétable dans un « réalisme des choses », mais relève d'un « réalisme des lois » (7, p. 29).

3.2. La résonance, un obstacle épistémologique

Pourtant, l'application des séries de Fourier à la psychoacoustique, célébrée par Ohm, a rencontré un obstacle, interrompant pendant près d'un siècle le processus de relativation de la fréquence. Cet obstacle est la résonance, observable expérimentalement lorsqu'une corde S de fréquence fondamentale ξ excite une corde S' de fréquence fondamentale $\xi' = p\xi$ avec $p \geq 2$ entier. Le phénomène est connu depuis Mersenne, et joue un rôle décisif dans la théorie tonale au XVIIIe siècle (17). Pour expliquer la résonance, il est tentant de postuler l'existence de cordes de longueur p fois plus petite que S, vibrant à l'unisson avec S'. Chez Rameau, par exemple, « toute note est contenue dans son octave » ($p = 2$), et *idem* pour $p = 3$ et $p = 5$ (40). Cette hypothèse sera réfutée, bien plus tard, par les expériences de Seebeck sur les sirènes. Néanmoins, elle est si séduisante qu'elle devient presque irrésistible, spécialement pour un compositeur de la tradition tonale européenne comme Rameau.

La théorie ramiste des renversements d'accord marque alors un recul par rapport au rationalisme classique d'un Newton, soit la restauration d'un « réalisme des choses » contre le « réalisme des lois ». Cent ans plus tard, le mémoire de Fourier semble accréditer la croyance en l'existence objective des harmoniques, supposées « contenues » dans le son complexe ; puis à la modélisation de l'oreille corporelle comme banc de cordes vibrantes, erreur de modélisation rectifiée par en 1960 par von Békésy (57). Malgré les nombreuses précautions de Helmholtz concernant l'importance de l'oreille mentale, l'historiographie de la psychoacoustique a retenu, pour mieux la dénoncer, la croyance aux sons purs comme quasi-atomes de la résonance. C'est contre cette croyance que s'érige le psychoacousticien W. Dixon Ward dans un texte célèbre de 1970 :

> *For years musicians have been told that the ear is able to separate any complex signal into a series of sinusoidal signals—that it acts as a Fourier analyzer. This quarter-truth, known as Ohm's Other Law, has served to increase the distrust with which perceptive musicians regard scientists, since it is readily apparent to them that the ear acts in this way only under very restricted conditions.* (58)

Rétrospectivement, on peut relire les débuts de l'informatique musicale comme une tentative de dépasser ce discrédit. Il s'agissait, nous semble-t-il, de remettre en cause la distinction classique entre sons « simples » et « complexes » tout en conservant la puissance de la dualité temps–fréquence (20). La synthèse FM, que nous avons évoquée en introduction, s'inscrit dans cette trajectoire.

3.3. Néo-empirisme après Fourier

Que faire ? Une issue consisterait à maintenir le principe de superposition tout en relâchant certaines des contraintes définissant les modes de Fourier : $a_p \cos(2\pi p\xi t + \varphi_p)$. En cette matière, la décomposition modale empirique (EMD pour *empirical mode decomposition*) offre une solution intéressante (28). L'EMD est un algorithme récursif dont les itérations produisent des composantes à bande étroite, pourvues d'amplitudes lentement variables $a_p(t)$ et de phases $\varphi_p(t)$ dont les dérivées $\dot{\varphi}_p(t)$ jouent le rôle de fréquences instantanées ; ces composantes sont appelées IMF pour *intrinsic mode functions*. Comme avec une série de Fourier, la somme des IMF donne une formule de synthèse :

$$x(t) = \sum_{p=1}^{+\infty} a_p(t) \cos\big(2\pi\varphi_p(t)\big). \tag{5}$$

Mais contrairement à l'équation (3), les amplitudes ne sont plus nécessairement constantes et les phases ne sont plus des fonctions affines du temps dont la pente vaut nécessairement $p\xi$ pour une fondamentale ξ. Il nous paraît légitime de décrire l'EMD comme une *relativation* de la série de Fourier.

Fait remarquable, chaque itération de l'EMD débute par l'estimation des maxima et minima locaux du résidu. On peut y voir une réactivation de l'image ancienne, chère à Mersenne, du pendule simple allant "deçà-delà". Mais, pour empirique qu'elle soit, l'EMD se prête à une analyse mathématique : elle met effectivement en œuvre une décomposition hiérarchique, adaptative, et multiéchelle (24). En ce sens, l'EMD n'est pas un retour naïf au « réalisme des choses » d'avant Newton ; c'est plutôt une forme de néo-empirisme, instruite par le calcul infinitésimal et par la théorie de Fourier. On peut dire de même de l'interpolation d'enveloppes en cascade (CEI pour *cascaded envelope interpolation*), une variation récente de l'EMD proposée par Thoret (55) pour les sciences de l'audition.

3.4. Surrationalisme : Fourier transformé par Fourier

Après avoir reconnu, dans la sous-section précédente, la valeur de l'empirisme face à l'obstacle de la résonance, nous voulons à présent le prendre à contrepied. Si l'EMD conserve le principe de superposition mais fait de la fréquence un résultat de la décomposition, nous proposons au contraire de suspendre le principe de superposition tout en maintenant la primauté du concept de fréquence sur celui de signal. Pour ce faire, nous traitons la fréquence comme ce que Bachelard nomme un *surobjet* : « résultat d'une objectivation critique, d'une objectivité qui ne retient de l'objet que ce qu'elle a critiqué » (7, p. 139).

Concrètement, nous énonçons quelques croyances normatives souvent associées à la fréquence, notée ici ω, dans

le sillage historique du différend Ohm–Seebeck :

1. « ω doit être réel pour que $e^{\mathrm{i}\omega t}$ ait un sens ; en particulier, $|e^{\mathrm{i}\omega t}| = 1$. »
2. « ω doit être muni d'une structure d'ordre total, conformément à l'ordonnancement des hauteurs musicales du grave à l'aigu. »
3. « ω doit être non signé : remplacer ω par $-\omega$ ne fait que conjuguer $e^{\mathrm{i}\omega t}$, sans conséquence audible. »
4. « ω doit varier en raison inverse de la durée : étirer le temps d'un facteur $k > 0$ revient à multiplier les fréquences des coefficients de Fourier par $1/k$. »

Nous cherchons alors une critique pour chacune de ces croyances. Ce faisant, nous les interrogeons dans leur prétention à l'universalité ; nous les rapportons à un biais tonal occidental ; et nous posons, très simplement, la question bachelardienne : *pourquoi pas ?* (5) Pourquoi ne pas admettre des fréquences négatives ($\omega < 0$), des fréquences complexes ($\omega \notin \mathbb{R}$), voire des fréquences à valeur dans des espaces de dimension infinie ? Ce geste polémique, par lequel l'enquête scientifique s'efforce de miner ses propres jugements de valeur, Bachelard l'a théorisé sous le nom de *surrationalisme*, un terme calqué sur le surréalisme artistique et littéraire (7; 8; 14).

La liste ci-dessus nous servira donc d'anti-guide de lecture. À chaque item, nous demandons « pourquoi pas ? » et relisons l'histoire de l'informatique musicale en quête de ses possibles relativations.

3.5. Sons de Shepard

Shepard a montré que les jugements de hauteur relative possèdent un caractère circulaire, en construisant une famille de sons dont la *chroma* est nette (au sens de l'échelle chromatique), mais dont la hauteur absolue (grave/aigu) est ambiguë (54). On peut réécrire son modèle sous la forme :

$$x(t) = \sum_{j=1}^{+\infty} 2^{-j} \cos\Big(2\pi\, 2^{-j+\theta/(2\pi)}\, t\Big), \tag{6}$$

où $\theta \in [0, 2\pi)$ code la chroma. La fréquence du son de Shepard, comprise classiquement comme inverse d'une période, est mal définie, car x_θ n'est pas périodique. Il est plus fécond, et plus conforme à la perception, de redéfinir sa fréquence comme l'ensemble

$$\omega = \{2^{-j+\theta/(2\pi)} \mid j \in \mathbb{Z}\},$$

c'est-à-dire la classe d'équivalence de $2^{\theta/(2\pi)}$ sous l'action du groupe des transpositions à l'octave, isomorphe à $\mathbb{Z}$. La fréquence n'est donc plus un scalaire : elle est définie à un facteur puissance de deux près, comme les angles le sont modulo 2π. On relativise ainsi, d'un seul geste, l'item (1) (fréquence comme variable réelle unique) et l'item (2) (ordre total).

3.6. Synthèse modale

Un modèle classique de cordes pincées consiste à sommer P sinusoïdes exponentiellement amorties :

$$x(t) = \sum_{p=1}^{P} a_p \exp(-\beta_p t) \cos(\alpha_p t), \tag{7}$$

où les β_p sont positifs, de sorte que l'enveloppe décroît pour $t \geq 0$. Par la formule d'Euler, on peut réécrire :

$$x(t) = \mathrm{Re}\left(\sum_{p=1}^{P} a_p\, e^{\mathrm{i}(\alpha_p + \mathrm{i}\beta_p)t}\right). \tag{8}$$

Il est alors naturel de poser $\omega_p = \alpha_p + \mathrm{i}\beta_p \in \mathbb{C}$. Ceci est plus qu'un artifice de notation : autoriser $|e^{\mathrm{i}\omega_p t}| \neq 1$ est au cœur des transformées de Laplace et de Z (50). Les fréquences complexes relativisent les items (1), (2) et (3).

3.7. Cepstres

Le spectre de puissance d'un signal $\boldsymbol{x}$ est le module au carré de sa transformée de Fourier :

$$S_{\boldsymbol{x}}(\xi) = \left| \int_{-\infty}^{+\infty} \boldsymbol{x}(t) e^{-2\pi \mathrm{i}\xi t}\, \mathrm{d}t \right|^2, \tag{9}$$

où $\xi \in \mathbb{R}$ est une fréquence en Hertz. Bogert, Healy et Tukey ont proposé de lui appliquer au spectre de puissance un logarithme, puis une transformée de Fourier inverse, avant de prendre à nouveau le module au carré (13). D'où :

$$C_{\boldsymbol{x}}(\omega) = \left| \int_{-\infty}^{+\infty} \big(\log S_{\boldsymbol{x}}(\xi)\big) e^{2\pi \mathrm{i}\omega\xi}\, \mathrm{d}\xi \right|^2, \tag{10}$$

une fonction non négative de la variable réelle ω. Par goût du calembour, Tukey propose le mot *cepstrum* comme anagramme de *spectrum* ; et, de même, la variable fréquentielle ξ devient une *quefrency* ω : inverse d'une fréquence, donc homogène à une durée. De même *alanysis*, *liftering*, *rahmonics*, etc. Le cepstre est un descripteur efficace du modèle linéaire source–filtre de la parole et se montre robuste au phénomène psychoacoustique de la fondamentale absente (51). Les quéfrences relativisent l'item (4) et plus radicalement le principe de superposition, puisque $C_{x_1+x_2}$ n'est pas, en général, égal à $C_{x_1} + C_{x_2}$.

4. ASPECTS CRÉATIFS

4.1. Diffusion jointe temps–fréquence

Dans les deux sections précédentes, nous avons étudié la formation historique du concept de fréquence, depuis l'astronomie antique jusqu'à l'invention du calcul infinitésimal ; sa crise au XIXe siècle (Fourier, Ohm, Seebeck, Helmholtz) ; et son objectivation critique au XXe siècle avec l'invention des sons de Shepard, des transformées en z et des cepstres. Nous y avons trouvé des *relativations* de la fréquence, c'est-à-dire des extensions conceptuelles au-delà d'une série d'évènements identiquement espacés,

comme une contemplation naïve du flux et reflux l'aurait suggéré. Nous nous tournons maintenant vers un opérateur mathématique plus récent, dit *diffusion jointe temps–fréquence* (JTFS pour *joint time–frequency scattering*), qui nous a permis de réaliser plusieurs œuvres en informatique musicale depuis 2017 (36). On verra que la JTFS constitue une nouvelle relativation de la fréquence : son *corps de notions* s'élargit pour accueillir des entités mathématiques inédites, telles que le chemin *(path)* et le spin.

Un test d'écoute, réalisé par des étudiant-e-s en classe de composition du CNSMD de Paris, suggère que la distance induite par la JTFS est un bon prédicteur des jugements de dissimilarité entre sons isolés (37). Complémentairement, des indices neurophysiologiques suggèrent que la JTFS fournit un modèle idéalisé pertinent des champs récepteurs spectrotemporels dans le cortex auditif des humains (48) et de mammifères non humains (30). Ainsi, bien que la JTFS ait été initialement inventée pour des applications de classification, elle peut aussi servir en synthèse (2; 34; 37).

4.2. Ondelettes

Nous définissons les ondelettes de Morlet, de fréquence centrale $\lambda > 0$ et de facteur de qualité Q, par

$$\boldsymbol{\psi}_\lambda(t) = \lambda \exp\left(-\frac{\lambda^2 t^2}{2Q^2}\right) \times \big(\exp(2\pi \mathrm{i}\lambda t) - \kappa\big), \quad (11)$$

où le terme correctif κ assure que chaque $\boldsymbol{\psi}_\lambda(t)$ possède une valeur moyenne nulle. Dans un cadre discret, les fréquences acoustiques λ sont typiquement de la forme $2^{n/Q}$, où n est un entier, couvrant ainsi le champ auditif humain. Pour un signal $\boldsymbol{x}(t)$ d'énergie finie, nous définissons le scalogramme de $\boldsymbol{x}$ comme la matrice $\mathbf{U}_1\boldsymbol{x}(t,\lambda) = \big|\boldsymbol{x} * \boldsymbol{\psi}_\lambda\big|(t)$. Nous renvoyons à (43) pour une introduction à la théorie des ondelettes, et à (52) pour ses applications à l'informatique musicale.

4.3. Chemins et spins

Deuxièmement, nous définissons des ondelettes de Morlet de fréquences centrales respectives $\alpha > 0$ et $\beta \in \mathbb{R}$, avec facteur de qualité $Q = 1$. Avec un léger abus de notation, nous notons ces ondelettes $\boldsymbol{\psi}_\alpha(t)$ et $\boldsymbol{\psi}_\beta(\log\lambda)$, bien qu'elles n'aient pas nécessairement la même forme que les ondelettes $\boldsymbol{\psi}_\lambda(t)$ de l'Équation (11). Nous définissons le tenseur d'ordre quatre

$$\begin{aligned}\mathbf{U}_2\boldsymbol{x}(t,\lambda,\alpha,\beta) &= \big|\mathbf{U}_1\boldsymbol{x} \overset{t}{*} \boldsymbol{\psi}_\alpha \overset{\log_2\lambda}{*} \boldsymbol{\psi}_\beta\big|(t,\lambda) \\ &= \left|\iint \mathbf{U}_1\boldsymbol{x}(\tau,s)\,\boldsymbol{\psi}_\alpha(t-\tau)\,\boldsymbol{\psi}_\beta(\log_2\lambda - s)\,\mathrm{d}\tau\,\mathrm{d}s\right|.\end{aligned} \quad (12)$$

La Figure 2 montre la figure d'interférences du produit $\boldsymbol{\psi}_\alpha(t-\tau)\boldsymbol{\psi}_\beta(\log_2\lambda - s)$ pour différentes combinaisons de temps t, de fréquence λ, de taux α et d'échelle β. Nous appelons *chemins* de diffusion (44) les multi-indices (λ,α,β) issus de ces combinaisons.

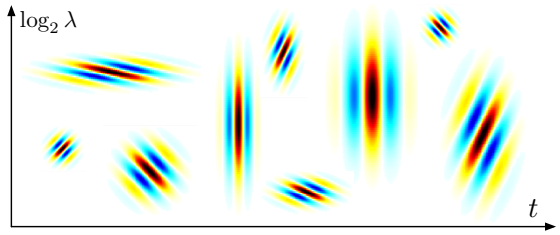

Figure 2. Figure d'interférence entre les ondelettes $\boldsymbol{\psi}_\alpha(t)$ et $\boldsymbol{\psi}_\beta(\log_2\lambda)$ dans le domaine temps–fréquence $(t, \log_2\lambda)$, pour différentes combinaisons de taux de modulation d'amplitude α et d'échelle de modulation de fréquence β. Les teintes rouges (resp. bleues) indiquent des valeurs positives (resp. négatives) de la partie réelle.

Les fréquences α, que nous nommons *taux* de modulation d'amplitude, sont mesurées en Hertz (Hz) et discrétisées comme 2^n pour n entier. Les fréquences β, que nous nommons *échelles* de modulation de fréquence, sont mesurées en cycles par octave (c/o) et discrétisées comme $\pm 2^n$ pour n entier. En ce sens, β ressemble à une *quéfrence* : voir la sous-section 3.7. De plus, puisque $\boldsymbol{\psi}_\alpha$ est à valeurs complexes, il en va de même du résultat de la convolution temporelle $(\mathbf{U}_1\boldsymbol{x} \overset{t}{*} \boldsymbol{\psi}_\alpha)$. Pour garantir la conservation de l'énergie, il est donc nécessaire d'inclure des fréquences négatives aussi bien que positives pour β. Le signe de β, noté s, est typiquement appelé le *spin* de l'ondelette $\boldsymbol{\psi}_\beta$ (3).

Le cas limite $\alpha = 0$ correspond à une ondelette $\boldsymbol{\psi}_\alpha(t)$ réduite à un filtre passe-bas gaussien $\boldsymbol{\phi}_T(t)$ de largeur temporelle T^{-1}, mesurée en Hertz. De même, le cas limite $\beta = 0$ correspond à une ondelette $\boldsymbol{\psi}_\beta(\log\lambda)$ réduite à un filtre passe-bas gaussien $\boldsymbol{\phi}_F(\log\lambda)$ de bande passante F^{-1}, mesurée en cycles par octave.

Enfin, nous effectuons une moyenne locale de $\mathbf{U}_1\boldsymbol{x}(t,\lambda)$ et de $\mathbf{U}_2\boldsymbol{x}(t,\lambda,\alpha,\beta)$ par des filtres passe-bas gaussiens $\boldsymbol{\phi}_T$ et $\boldsymbol{\phi}_F$, de fréquences de coupure respectives T^{-1} (en Hertz) et F^{-1} (en cycles par octave) :

$$\mathbf{S}_1\boldsymbol{x}(t,\lambda) = \big(\mathbf{U}_1\boldsymbol{x} \overset{t}{*} \boldsymbol{\phi}_T \overset{\log\lambda}{*} \boldsymbol{\phi}_F\big)(t,\lambda),$$

$$\mathbf{S}_2\boldsymbol{x}(t,\lambda,\alpha,\beta) = \big(\mathbf{U}_2\boldsymbol{x} \overset{t}{*} \boldsymbol{\phi}_T \overset{\log_2\lambda}{*} \boldsymbol{\phi}_F\big)(t,\lambda,\alpha,\beta).$$

4.4. Resynthèse

En raison des non-linéarités dues à l'application du module complexe, la JTFS ne satisfait pas le principe de superposition et ne possède pas d'inverse en forme fermée. Cependant, comme nous allons le voir, cette limitation peut être convertie en opportunité de création sonore.

Étant donnée la transformée de diffusion $\mathbf{S}\boldsymbol{x}$ d'un signal $\boldsymbol{x}$, nous définissons l'erreur quadratique moyenne

$$E_{\mathbf{S}\boldsymbol{x}}(\boldsymbol{y}) = \|\mathbf{S}\boldsymbol{x} - \mathbf{S}\boldsymbol{y}\|_2^2.$$

À partir d'un bruit gaussien $\boldsymbol{y}_0(t)$ dont la densité spectrale de puissance correspond à $\mathbf{S}_1\boldsymbol{x}(t,\lambda)$, nous le raffinons par

mises à jour additives de la forme $\boldsymbol{y}_{n+1}(t) = \boldsymbol{y}_n(t) + \boldsymbol{u}_n(t)$, où $\boldsymbol{u}_n(t)$ est défini récursivement par

$$\boldsymbol{u}_n(t) = m \times \boldsymbol{u}_n(t) - \mu_n \, \boldsymbol{\nabla} E_{\mathbf{S}\boldsymbol{x}}(\boldsymbol{y}_n)(t),$$

avec m est un terme de moment et μ_n une suite de pas d'apprentissage, elle aussi mise à jour itérativement (3).

Comme dans un réseau de neurones profond, le gradient $\boldsymbol{\nabla} E_{\mathbf{S}\boldsymbol{x}}(\boldsymbol{y}_n)$ s'obtient par rétropropoagation, c'est-à-dire en composant les adjoints hermitiens des opérateurs convolutifs et des non-linéarités dans l'ordre inverse de la diffusion directe. Nous nous appuyons sur la bibliothèque logicielle Kymatio pour effectuer cette rétropropagation par différentiation automatique en mode inverse (4).

Pour les besoins de la recherche en cognition musicale, la procédure ci-dessus permet de générer des stimuli contrôlés pour des tests d'écoute : des signaux $\boldsymbol{y}$ dont les coefficients de diffusion sont approximativement égaux à ceux de $\boldsymbol{x}$, après moyennage spectrotemporel par $\boldsymbol{\phi}_T$ en temps et $\boldsymbol{\phi}_F$ en log-fréquence, mais qui peuvent néanmoins sonner différemment à l'oreille humaine (32). Autrement dit, $(\boldsymbol{x}, \boldsymbol{y})$ forme une sorte de « paire métamérique » pour l'opérateur JTFS moyenné $\mathbf{S}$ (23). Dans ce cadre, on prend $\boldsymbol{y} = \boldsymbol{y}_N$, c'est-à-dire le N^{e} itéré d'une descente de gradient, pour une valeur N suffisamment grande.

4.5. *Scattering Impression Ear Brain Dance Mix* (2024)

La situation est différente dans le contexte de la création en informatique musicale. Ici, nous ne nous intéressons pas seulement à la comparaison binaire entre $\boldsymbol{x}$ et $\boldsymbol{y}_N$, comme dans une paire métamérique ; mais à la suite entière des itérés de reconstruction. Du point de vue créatif, les sons $\boldsymbol{y}_n$ pour $n < N$ sont des débris négligés d'un processus de calcul—à savoir, la descente de gradient. Pour le ou la scientifique, chaque $\boldsymbol{y}_n$ n'est qu'un état intermédiaire en vue de $\boldsymbol{y}_N$ et peut être jeté dès l'itération $(n+1)$. Pour le compositeur, au contraire, $\boldsymbol{y}_n$ peut contenir des propriétés sonores qui ne sont ni dans $\boldsymbol{x}$, ni dans $\boldsymbol{y}_0$, ni dans $\boldsymbol{y}_N$. Ces aspects inouïs des sous-produits computationnels sont au cœur de *Scattering Impression Ear Brain Dance Mix* (Hyperdub / Flatlines, 2024), œuvre dans laquelle les itérés $\boldsymbol{y}_{N_{\min}}$ à $\boldsymbol{y}_{N_{\max}}$ sont concaténés dans le domaine temporel. En prenant $N_{\min} = 50$ et $N_{\max} = 149$, le résultat est exactement cent fois plus long que $\boldsymbol{x}$.

À la lumière de ce qui précède, nous pouvons revenir à notre discussion méthodologique sur le relativisme spectrotemporel critique. Dans *Scattering Impression Ear Brain Dance Mix*, bien que certains motifs semblent revenir périodiquement, aucun fragment sonore $\boldsymbol{y}_n$ n'est présenté à l'oreille plus d'une fois en tant que tel. Les points de rupture structurels et les motifs sont plutôt la marque de la convergence de l'algorithme de descente de gradient. En termes mathématiques, $\boldsymbol{\nabla} E_{\mathbf{S}\boldsymbol{x}}(\boldsymbol{y}_n)$ perd en intensité et en saillance à mesure que n croît. Pourtant, nos oreilles s'accom modent de plus en plus finement à de faibles différences de contraste spectrotemporel, à mesure que l'œuvre se déploie et que le nombre d'itérations présentées augmente. En adjoignant le temps compositionnel (les fragments $\boldsymbol{y}_n$) et le temps computationnel (le pas d'apprentissage μ_n), *Scattering Impression Ear Brain Dance Mix* donne à entendre des différences qui sont typiquement effacées dans la paire métamérique $(\boldsymbol{x}, \boldsymbol{y}_N)$ lorsque N est grand. « Le relativiste », écrit Bachelard, « nous contraint à incorporer notre expérience dans notre conceptualisation. Il nous rappelle que notre conceptualisation est une expérience. » (5, p. 44)

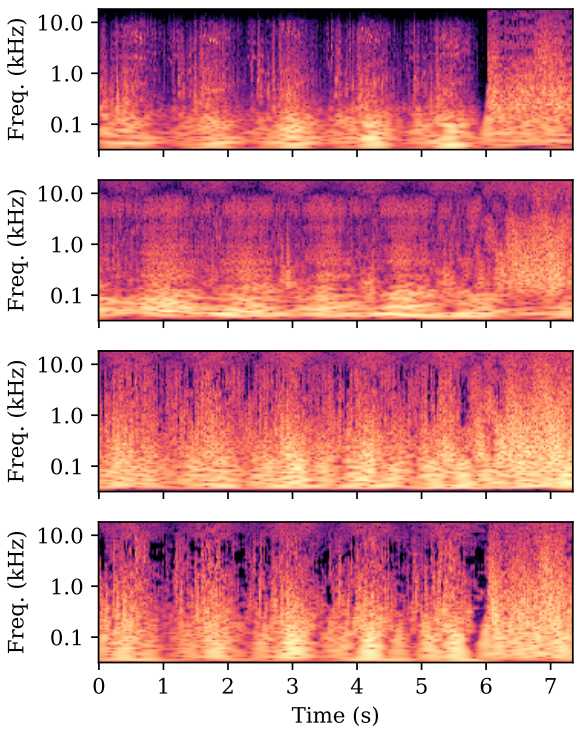

Figure 3. De haut en bas : signal d'origine $\boldsymbol{x}$ et resynthèses itératives $\boldsymbol{y_n}$ pour $n = 2, 10, 1000$. Représentation temps–fréquence à facteur de qualité constant (CQT).

L'œuvre a été créée par Florian Hecker, avec Cyrus Asfa comme réalisateur en informatique musicale (RIM). Cette œuvre appartient à un CD de compilation, "Dunwich Recursed", sorti en septembre 2024 sur Flatlines, une succursale du label Hyperdub consacrée aux essais audio et la fiction sonore. Le signal initial $\boldsymbol{x}$ est l'essai audio *By The North Sea* (42), lui aussi paru chez Hyperdub / Flatlines. Pour plus de détails, lire : (10).

4.6. *Syn 21845 (HEFT Sequenz)* (2022)

La création en informatique musicale permet de repenser la procédure de résynthèse, mentionnée plus haut, au-delà de la simple reconstruction d'un signal $\boldsymbol{x}$ à partir de ses coefficients JTFS moyennés $\mathbf{S}\boldsymbol{x}$. Pour cause, dans la définition de l'erreur quadratique moyenne $E_{\mathbf{S}\boldsymbol{x}}$, le terme $\mathbf{S}\boldsymbol{x}$ peut être remplacé par tout vecteur non négatif $\boldsymbol{c}$ de même dimension. Autrement dit, il n'est pas nécessaire de se référer à un signal préalablement observé $\boldsymbol{x}$ pour calculer l'erreur $E_{\boldsymbol{c}}(\boldsymbol{y})$ en toute généralité. De fait, un signal $\boldsymbol{x}$ satisfaisant $\mathbf{S}\boldsymbol{x} = \boldsymbol{c}$ peut ne pas exister ; ce qui signifie

Figure 4. Florian Hecker, *Syn 21845 (HEFT Sequenz)*, 2022. Son généré par ordinateur, dispositif de haut-parleurs. Photographie de l'installation © Markus Krottendorfer.

que le minimum global de $E_{\boldsymbol{c}}$ peut demeurer strictement positif. Cela ne pose aucun problème à la descente de gradient : le gradient $\boldsymbol{\nabla} E_{\boldsymbol{c}}(\boldsymbol{y})$ est calculé par rapport à $\boldsymbol{y}$, et $\boldsymbol{c}$ n'intervient que comme constante dans l'espace JTFS.

On peut ainsi inventer de nouveaux effets audionumériques en définissant $\boldsymbol{c}$ comme le résultat de transformations algébriques ou analytiques simples appliquées à $\mathbf{S}\boldsymbol{x}$ .Plus précisément, nous proposons d'échanger, dans le tenseur $\mathbf{S}\boldsymbol{x}$, les coefficients correspondant aux échelles de modulation β et $-\beta$, et ce pour tout quadruplet de temps t, de fréquence λ, d'échelle temporelle α et d'échelle fréquentielle β. Formellement :

$$\boldsymbol{c}(t, \lambda, \alpha, \beta) = \mathbf{S}\boldsymbol{x}(t, \lambda, \alpha, -\beta). \tag{13}$$

Nous appelons cette opération *retournement de spin* (*spin reversal*), en référence à la notion de spin $s \in \{-1, 0, 1\}$, qui n'est autre que le signe de β. Le retournement de spin ne se réduit ni à une inversion temporelle ni à une inversion fréquentielle : nous y voyons plutôt une nouvelle relativation de la notion de fréquence, au même titre que l'équation de Dirac constitue une relativation de la notion de masse en mécanique quantique relativiste (7, p. 33).

Le retournement de spin est analogue à la conjugaison de charge des bosons W et Z en physique des particules (1). Ces particules correspondent respectivement à des filtres $\beta \neq 0$ ($s = \pm 1$, passe-bande) et $\beta = 0$ ($s = 0$, passe-bas). Bien que la structure algébrique sous-jacente diffère, les deux opérations sont involutives et définissent des actions isomorphes de $\mathbb{Z}_2$, comportant un point fixe et une orbite libre. Il convient toutefois de rester prudent dans l'analogie : le spin physique, entendu comme moment cinétique quantifié, est catégoriquement distinct du spin de la JTFS. Répétons-le : c'est la charge, et non le spin, qui fournit l'équivalent le plus pertinent du spin JTFS dans le modèle standard.

Syn 21845 (HEFT Sequenz) est une installation sonore pour trois haut-parleurs, réalisée à l'occasion de l'exposition *Günther Domenig : Dimensional*, présentée du 12 juin au 16 octobre 2022 à HEFT, Hüttenberg (Autriche). L'Scattering Impression Ear Brain Dance Mix été créée par Florian Hecker, avec Vincent Lostanlen comme concepteur en informatique musicale. Sa durée est de 25 minutes et 29 secondes. Au cours de chaque journée d'exposition, quinze versions distinctes furent diffusées successivement. Ces versions diffèrent par les valeurs des paramètres J et Q, ainsi que par l'itération retenue pour chaque réalisation.

Une version stéréo de cette œuvre est parue en janvier 2026 sous le titre *Syn 21845 8 J15 Q12*, au sein du CD *Natural Selection*, publié par le label PAN. Voici un extrait du communiqué de presse :

> *The track takes microscopic glissandi fragments from a database of source material for Lostanlen's time–frequency scattering process and spatiotemporally* (sic) *reverses them, creating an unsettling, utterly beguiling audio illusion. What sounds like a barrage of whooshing, perpetually reversing acoustic impressions is actually a progressive metamorphosis of both forward and backward time—no audio is actually reversed.*

5. CONCLUSION

Nous devons veiller à ne pas réduire le concept de fréquence à celui de nombre d'onde, c'est-à-dire à un rapport de la forme N/T, où N désignerait un nombre d'événements et T une durée (ou toute autre dimension spatiale). Une telle image est non seulement restrictive, elle est aussi datée et, dans le cas de l'informatique musicale, contraire à la pratique. Nous avons vu que la fréquence peut être irrationnelle (synthèse FM), variable dans le temps (EMD et interpolation d'enveloppes en cascade), multivaluée (sons de Shepard), à valeurs complexes (synthèse modale et transformées en z), homogène au temps (cepstres), multi-indexée (diffusion en ondelettes), et négative (retournement de spin).

Nous avons présenté deux œuvres, fondées sur un algorithme de descente de gradient dans l'espace des coefficients de diffusion conjointe temps–fréquence (JTFS). Chaque itération intermédiaire de cet algorithme produit un son. Bien que ce son soit, en contexte scientifique, généralement écarté, nous l'intégrons à nos œuvres comme le débris négligé d'un processus computationnel. Ce geste ouvre la possibilité rare d'*écouter* ce que l'algorithme synthétise à chaque étape, et non pas seulement à l'issue du calcul : voir Figure 3. C'est un geste réflexif en ceci qu'il rappelle que le calcul pour la synthèse requiert du temps et des ressources.Ce temps écoulé est plus qu'un simple inverse de fréquence : il marque la reconnaissance que la musicalité machinique, à l'instar des modèles biologiques et à la différence des modèles purement mathématiques, doit affronter le problème du devenir (22).

Pour qu'un fil de synthèse JTFS devienne une œuvre fixée sur support, le flux et reflux de la rétropropagation du gradient doit, tôt ou tard, être interrompu. Sur ce point, nous rejoignons la compositrice Laurie Spiegel :

> *By analogy to [Turing's halting problem], the test for organic (human, versus machine)*

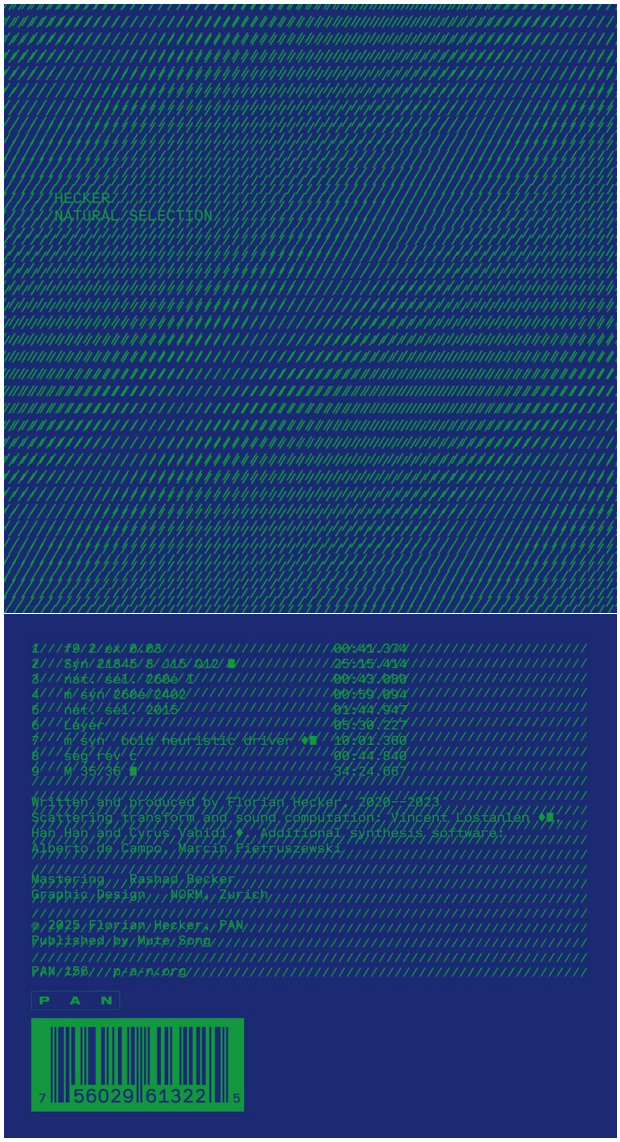

Figure 5. Couverture (haut) et quatrième de couverture (bas) de l'album *Natural Selection* (PAN, 2026), sur lequel figure *Syn 21845 8 J15 Q12*. Infographie © NORM.

> *music would be whether and when it stops. Musicians could be considered inhuman if they fail to stop playing after some finite period or number of creative steps. A computer can go on making music forever, whereas, at some point a human musician is bound to get tired and stops. If the music never resolves or halts, probably a machine is playing* (11).

Scattering Impression Ear Brain Dance Mix satisfait-elle au test de Spiegel ? Pour une auditrice humaine, cela est évident. En revanche, la question demeure insoluble—nous n'osons écrire *indécidable*—du point de vue d'une modélisation symbolique *a priori*. Ayant éprouvé les limites du relativisme spectrotemporel critique, nous nous efforçons d'intégrer les conditions de son exercice.

Références

[1] Aitchison, I. J. R. and Hey, A. J. G. *Gauge Theories in Particle Physics : Volume I : From Relativistic Quantum Mechanics to QED*. CRC Press, Boca Raton, 2013.

[2] Andén, J., Lostanlen, V. and Mallat, S. "Joint time-frequency scattering for audio classification", *Proceedings of the International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, 2015.

[3] Andén, J., Lostanlen, V. and Mallat, S. "Joint time–frequency scattering", *IEEE Transactions on Signal Processing*, 67(14), pp. 3704–3718, 2019.

[4] Andreux, M., Angles, T., Exarchakis, G., Leonarduzzi, R., Rochette, G., Thiry, L., Zarka, J., Mallat, S., Andén, J., Belilovsky, E. et al. "Kymatio : Scattering transforms in Python", *Journal of Machine Learning Research*, 21(60), pp. 1–6, 2020.

[5] Bachelard, G. *Le Nouvel Esprit scientifique*. PUF, 1934.

[6] Bachelard, G. *La Valeur inductive de la relativité*. Vrin, Paris, 1929.

[7] Bachelard, G. *La Philosophie du non*. Presses Universitaires de France, Paris, 1940.

[8] Bachelard, G. *L'engagement rationaliste*. Presses Universitaires de France, Paris, 1972.

[9] Bailache, P., Soulez, A. and Vautrin, C. *Helmholtz : Du son à la musique*. Vrin, Paris, 2011.

[10] Barton, J., Goodman, S. and Kronic, M. B. (eds.). *Sonic Faction : Audio Essay as Medium and Method*. Urbanomic, Falmouth, 2024.

[11] Belgum, E., Roads, C., Chadabe, J., Tobenfeld, T. E. and Spiegel, L. "A Turing test for "musical intelligence" ?", *Computer Music Journal*, 12(4), pp. 7–9, 1988.

[12] Blay, M. *La naissance de la mécanique analytique : la science du mouvement au tournant des XVIIe et XVIIIe siècles*. Presses Universitaires de France, Paris, 1994.

[13] Bogert, B. P., Healy, M. J. R. and Tukey, J. W. "The quefrency analysis of time series for echoes : Cepstrum, pseudoautocovariance, cross-cepstrum and saphe cracking", *Proceedings of the Symposium on Time Series Analysis*, pp. 209–243, 1963.

[14] Bontems, V. *Bachelard*. Les Belles Lettres, Paris, 2010.

[15] Bouveresse, J. "Helmholtz et la théorie physiologique de la musique", in *Parole et musique : Aux origines du dialogue humain*, S. Dehaene and C. Petit (eds.), Odile Jacob, Paris, 2009.

[16] Brack-Bernsen, L. "Babylonian astronomy : a new understanding of column Φ : Schematic astronomy, old prediction rules, riddles, loose ends, and new ideas", *Archive for History of Exact Sciences*, 74(6), pp. 605–640, 2020.

[17] Charrak, A. *Raison et perception : fonder l'harmonie au XVIIIe siècle*. Vrin, Paris, 2001.

[18] Chowning, J. M. "The synthesis of complex audio spectra by means of frequency modulation", *Journal of the Audio Engineering Society*, 21(7), pp. 526–534, 1973.

[19] Chowning, J. "Mathews' Diagram and Euclid's Line : Fifty Years Ago", *ICMC/SMC Keynote Address*, 2014.

[20] Cooley, J. W. and Tukey, J. W. "An algorithm for the machine calculation of complex Fourier series", *Mathematics of Computation*, 19(90), pp. 297–301, 1965.

[21] Einstein, A. *L'éther et la théorie de la relativité*. Trad. M. Solovine, Gauthier-Villars, Paris, 1921.

[22] Fagot-Largeault, A. *Ontologie du devenir : L'évolution, l'univers et le temps*. Odile Jacob, 2021.

[23] Feather, J., Leclerc, G., Mądry, A. and McDermott, J. H. "Model metamers reveal divergent invariances between biological and artificial neural networks", *Nature Neuroscience*, 26(11), pp. 2017–2034, 2023.

[24] Flandrin, P., Gonçalves, P. and Rilling, G. "EMD equivalent filter banks, from interpretation to applications", in *Hilbert–Huang Transform and Its Applications*, pp. 99–116, World Scientific, 2014.

[25] Gingras, Y. *Histoire des sciences*. Que sais-je ?, Paris, 2024.

[26] Haack, S. "Reflections on relativism : From momentous tautology to seductive contradiction", *Philosophical Perspectives*, 10, pp. 297–315, 1996.

[27] Herreman, A. "L'inauguration des séries trigonométriques dans la Théorie analytique de la chaleur de Fourier et dans la controverse des cordes vibrantes", *Revue d'histoire des mathématiques*, 19(2), pp. 151–243, 2013.

[28] Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C. and Liu, H. H. "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis", *Proceedings of the Royal Society of London. Series A : Mathematical, Physical and Engineering Sciences*, 454(1971), pp. 903–995, 1998.

[29] Kahane, J.-P. and Lemarié, P.-G. *Fourier Series and Wavelets*. Gordon and Breach, 1995.

[30] Kowalski, N., Depireux, D. A. and Shamma, S. A. "Analysis of dynamic spectra in ferret primary auditory cortex. I. Characteristics of single-unit responses to moving ripple spectra", *Journal of Neurophysiology*, 76(5), pp. 3503–3523, 1996.

[31] Koyré, A. *From the Closed World to the Infinite Universe*. Johns Hopkins University Press, Baltimore, 1957.

[32] Li, J., Baker, D. J., Burgoyne, J. A., Han, H., Henry, N., Lostanlen, V., Sadakata, M., Van der Vlist, M., van Schaik, T. M., Janmaat, K. R. L. and Honing, H. "Is Pitch Indispensable for Music Recognition ? : Robust Music Recognition Under Spectral and Temporal Degradations in BaYaka Hunter-Gatherers", *Proceedings of the International Conference on Music Perception and Cognition (ICMPC)*, 2025.

[33] Lostanlen, V. *Convolutional operators in the time-frequency domain*. PhD thesis, École normale supérieure, 2017.

[34] Lostanlen, V. "On time–frequency scattering and computer music", in *Florian Hecker : Halluzination, Perspektive, Synthese*, V. J. Müller (ed.), Sternberg Press, 2018.

[35] Lostanlen, V., Andén, J. and Lagrange, M. "Fourier at the heart of computer music : From harmonic sounds to texture", *Comptes Rendus. Physique*, 20(5), pp. 461–473, 2019.

[36] Lostanlen, V. and Hecker, F. "The shape of RemiXXXes to come : Audio texture synthesis with time-frequency scattering", *Proceedings of the International Conference on Digital Audio Effects (DAFX)*, 2019.

[37] Lostanlen, V., El-Hajj, C., Rossignol, M., Lafay, G., Andén, J. and Lagrange, M. "Time–frequency scattering accurately models auditory similarities between instrumental playing techniques", *EURASIP*

Journal on Audio, Speech, and Music Processing, 2021(1), p. 3, 2021.

[38] Lostanlen, V. and Hecker, F. "Synopsis Seriation : A computer music piece made with time-frequency scattering and information geometry", *Actes des Journées d'Informatique Musicale*, 2021.

[39] Lostanlen, V. "Phantasmagoria : Sound synthesis after the Turing test", *Speculative Sound Synthesis Symposium*, 2024.

[40] Lostanlen, V. "Rameau transformé par Fourier", in *La musique–science devant la question de l'harmonie*, dir. Gaulhiac, M. et Chouvel, J. M., Delatour, 2026.

[41] Mackay, R. "The Brain–Ear", article en ligne, February 26, 2015. https ://readthis.wtf/writing/the-brain-ear/

[42] Mackay, R. *By the North Sea*. Flatlines, FLAT003D (compact disc), London, United Kingdom, 2024.

[43] Mallat, S. *A Wavelet Tour of Signal Processing*. 2nd ed., Academic Press, San Diego, 1999.

[44] Mallat, S. "Group invariant scattering", *Communications on Pure and Applied Mathematics*, 65(10), pp. 1331–1398, 2012.

[45] Mersenne, M. *Harmonie universelle, contenant la théorie et la pratique de la musique*. Paris, 1636–1637. Available : https ://gallica.bnf.fr/ark :/12148/bpt6k5471093v

[46] Moore, B. C. J. *An Introduction to the Psychology of Hearing*. Brill, 2012.

[47] Neugebauer, O. *The Exact Sciences in Antiquity*. Courier Corporation, 1969.

[48] Norman-Haignere, S. V. and McDermott, J. H. "Neural responses to natural and model-matched stimuli reveal distinct computations in primary and nonprimary auditory cortex", *PLoS Biology*, 16(12), e2005127, 2018.

[49] Ohm, G. S. "Über die Definition des Tones, nebst daran geknüpfter Theorie der Sirene und ähnlicher tonbildender Vorrichtungen", *Annalen der Physik*, 135(8), pp. 513–565, 1843.

[50] Oppenheim, A. V., Willsky, A. S. and Nawab, S. H. *Signals and Systems*. 2nd ed., Prentice Hall, Upper Saddle River, NJ, 1997.

[51] Oppenheim, A. V. and Schafer, R. W. "From frequency to quefrency : A history of the cepstrum", *IEEE Signal Processing Magazine*, 21(5), pp. 95–106, 2004.

[52] Roads, C. *Microsound*. The MIT Press, 2002.

[53] Schwartz, L. *Théorie des distributions*. Hermann, 1951.

[54] Shepard, R. N. "Circularity in judgments of relative pitch", *The Journal of the Acoustical Society of America*, 36(12), pp. 2346–2353, 1964.

[55] Thoret, E., Ystad, S. and Kronland-Martinet, R. "Hearing as adaptive cascaded envelope interpolation", *Communications Biology*, 6(1), 671, 2023.

[56] Turner, R. S. "The Ohm-Seebeck dispute, Hermann von Helmholtz, and the origins of physiological acoustics", *The British Journal for the History of Science*, 10(1), pp. 1–24, 1977.

[57] von Békésy, G. *Experiments in Hearing*. McGraw–Hill Book Company, 1960.

[58] Ward, W. D. "Musical Perception", in *Foundations of Modern Auditory Theory*, J. V. Tobias (ed.), vol. 1, Academic Press, 1970, p. 438.