

Zéros du spectrogramme et filtrage temps-fréquence

Patrick FLANDRIN¹

¹Laboratoire de Physique, CNRS - École Normale Supérieure de Lyon

patrick.flandrin@ens-lyon.fr

Résumé – Pour un choix convenable de la fenêtre d’analyse, on sait qu’une transformée de Fourier à court-terme est complètement caractérisée par ses zéros, qui coïncident avec ceux du spectrogramme associé. Une représentation simplifiée de la structure temps-fréquence d’un signal peut alors être donnée par la triangulation de Delaunay définie sur les zéros du spectrogramme. Dans le cas de signaux AM-FM multicomposantes bruités, il s’avère que chaque région temps-fréquence attachée à une composante donnée peut se voir comme l’union de triangles de Delaunay adjacents dont au moins un côté est anormalement long en comparaison avec la distribution dans le cas bruit seul. Identifier de telles régions offre une nouvelle possibilité de séparer les différentes composantes dans le plan temps-fréquence et de reconstruire les signaux correspondants.

Abstract – For a proper choice of the analysis window, a short-time Fourier transform is known to be completely characterized by its zeros, which coincide with those of the associated spectrogram. A simplified representation of the time-frequency structure of a signal can therefore be given by the Delaunay triangulation attached to spectrogram zeros. In the case of multicomponent AM-FM signals embedded in white Gaussian noise, it turns out that each time-frequency domain attached to a given component can be viewed as the union of adjacent Delaunay triangles whose edge length is an outlier as compared to the distribution in noise-only regions. Identifying such domains offers a new way of disentangling the different components in the time-frequency plane, as well as of reconstructing the corresponding waveforms.

1 Introduction

Le spectrogramme est sans doute la méthode d’analyse temps-fréquence (TF) la plus simple et la plus naturelle [8]. Dans le cas de signaux AM-FM ayant un nombre limité de composantes, c’est une représentation relativement parcimonieuse, avec seulement quelques bandes d’énergie localisées le long des trajectoires TF des différentes composantes. Le filtrage de ces composantes est souvent effectué en identifiant les domaines d’intérêt et en inversant la transformée après masquage. Cette question a récemment connu un regain d’intérêt en lien avec le développement de techniques qui, comme la réallocation et le “synchrosqueezing” [1,6,21], ou encore les “contours” [18], reposent sur une idée de bassins d’attraction. Dans la plupart des cas, l’identification des “régions signal” se base sur des critères énergétiques locaux et/ou des arêtes [7] associées aux maxima locaux.

En contraste avec de telles approches reposant sur de *grandes valeurs* de la représentation, on propose ici de faire usage des *zéros* comme points caractéristiques.

2 STFT et spectrogramme

2.1 Définitions et rappels

Étant donné un signal $x(t)$, la Transformée de Fourier à Court-Terme (TFCT) $F_x^{(h)}(t, \omega)$ de fenêtre $h(t)$ est définie comme le

Ce travail a été réalisé avec le soutien de l’ANR au titre du projet ASTRES, ANR-13-BS03-0002-01.

produit scalaire entre $x(t)$ et les translatées (en temps et fréquence) de $h(t)$, i.e., comme $F_x^{(h)}(t, \omega) = \langle x, \mathbf{T}_{t\omega} h \rangle$, où $\mathbf{T}_{t\omega}$ est un opérateur de translation conjointe. Les translations individuelles ne commutent pas, cet opérateur n’est pas unique et la TFCT est définie à un terme de phase près que l’on fixera ici en adoptant la forme explicite [4, 5] :

$$F_x^{(h)}(t, \omega) = \int_{-\infty}^{+\infty} x(s) \overline{h(s-t)} \exp \left\{ -i\omega \left(s - \frac{t}{2} \right) \right\} ds. \quad (1)$$

Le *spectrogramme* correspondant s’en déduit selon :

$$S_x^{(h)}(t, \omega) = \left| F_x^{(h)}(t, \omega) \right|^2. \quad (2)$$

Il est bien connu que $F_x^{(h)}(t, \omega)$ n’est pas n’importe quelle fonction 2D dans la mesure où (1) satisfait l’identité :

$$F_x^{(h)}(t', \omega') = \int_{-\infty}^{+\infty} K(t', \omega'; t, \omega) F_x^{(h)}(t, \omega) dt \frac{d\omega}{2\pi}, \quad (3)$$

dans laquelle le noyau reproduisant $K(t', \omega'; t, \omega)$ n’est autre (à un terme multiplicatif près) que la TFCT de la fenêtre :

$$K(t', \omega'; t, \omega) = \frac{\exp \{ i(\omega t' - \omega' t) / 2 \}}{\|h\|_2^2} F_h^{(h)}(t' - t, \omega' - \omega). \quad (4)$$

Ce faisant, il est facile de justifier que toute TFCT (ou tout spectrogramme) possède naturellement une redondance interne car le noyau reproduisant (4) ne peut pas être arbitrairement concentré en temps et en fréquence (voir , e.g., [17] ou [13] pour des formulations précises). Dans le cas particulier de la fenêtre gaussienne (d’énergie unité) dite “circulaire” :

$$g(t) = \pi^{-1/4} \exp \{ -t^2 / 2 \}, \quad (5)$$

le noyau reproduisant est de concentration maximale et définit pour le spectrogramme (appelé “distribution de Husimi” [14, 15] en Physique) un “domaine d’influence” circulaire dont l’aire effective est celle de la gaussienne 2D de variance 2.

2.2 Spectrogramme et bruit blanc gaussien

Le spectrogramme est classiquement défini pour des signaux certains d’énergie finie, mais il peut aussi être utilisé pour l’analyse de processus harmonisables de puissance moyenne finie [8]. On s’intéressera ici au cas idéalisé du bruit blanc gaussien circulaire tel que :

$$\mathbb{E}\{n(t)\overline{n(t')}\} = \gamma_0 \delta(t - t') \quad ; \quad \mathbb{E}\{n(t)n(t')\} = 0. \quad (6)$$

Il suit de (2) et (6) que l’espérance du spectrogramme est dans ce cas constante :

$$\mathbb{E}\{S_n^{(h)}(t, \omega)\} = \gamma_0^2, \quad (7)$$

et que la covariance relative à deux positions dans le plan ne dépend que de leurs différences :

$$\text{cov}\{S_n^{(h)}(t, \omega), S_n^{(h)}(t', \omega')\} = \gamma_0^2 S_h^{(h)}(t' - t, \omega - \omega'). \quad (8)$$

Dans le cas particulier (5), ceci se réduit à

$$\text{cov}\{S_n^{(g)}(t, \omega), S_n^{(g)}(t', \omega')\} = \gamma_0^2 \exp\{-d^2((t, \omega), (t', \omega'))/2\} \quad (9)$$

où

$$d((t, \omega), (t', \omega')) = \sqrt{(t - t')^2 + (\omega - \omega')^2} \quad (10)$$

mesure la distance TF euclidienne entre les points. Comme fonction de cette seule distance, le spectrogramme du bruit blanc peut être vu comme un champ stationnaire. Il est attendu que cette homogénéité se transfère aux points caractéristiques de la surface (maxima locaux et zéros), en y incluant des contraintes dues au noyau reproduisant.

3 Zéros du spectrogramme

3.1 Bargmann

Il est intéressant de voir le temps et la fréquence comme des coordonnées du plan complexe. Introduisant ainsi $z = \omega + it$, un calcul direct montre que, dans le cas de la fenêtre (5), la TFCT (1) peut s’écrire :

$$F_x^{(g)}(t, \omega) = \exp\{-|z|^2/4\} \mathcal{F}_x(z), \quad (11)$$

où $\mathcal{F}_x(z)$ est la transformée de Bargmann de x [2, 4].

Combinant la propriété $|\mathcal{F}_x^{(g)}(t, \omega)| \leq \|x\|$ avec la factorisation (11), on obtient

$$|\mathcal{F}_x(z)| \leq \|x\| \exp\{|z|^2/4\}, \quad (12)$$

d’où l’on déduit que $\mathcal{F}_x(z)$ est une fonction entière d’ordre 2 [3]. En conséquence, elle admet une factorisation de Weierstrass-Hadamard de la forme [3, 12, 15, 22]

$$\mathcal{F}_x(z) \propto \prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right) \exp\left\{\frac{z}{z_n} + \frac{1}{2} \left(\frac{z}{z_n}\right)^2\right\}, \quad (13)$$

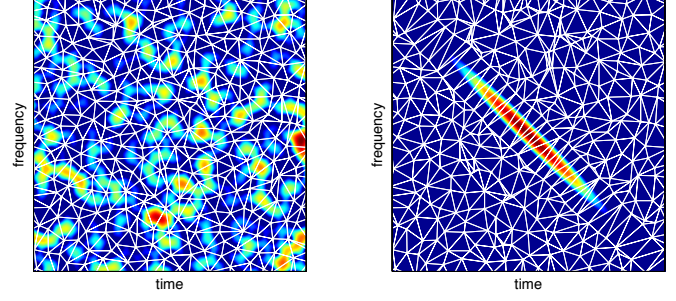


FIGURE 1 – **Triangulation de Delaunay**. Gauche : bruit seul. Droite : mélange signal + bruit.

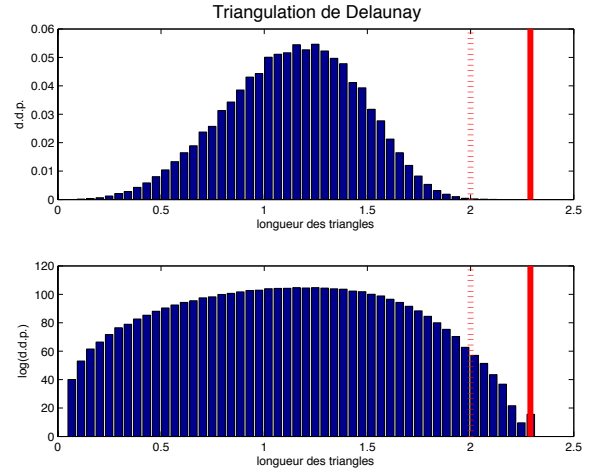


FIGURE 2 – **Triangulation de Delaunay — Distribution des longueurs de côté dans le cas du bruit blanc gaussien**. Haut : échelle linéaire ; Bas : échelle logarithmique (dB).

où les variables $z_n = \omega_n + it_n$ représentent les zéros de la transformée de Bargmann qui, par construction, sont les mêmes que ceux de la TFCT et du spectrogramme. Quoique la factorisation (13) ne puisse que difficilement être utilisée de façon directe, sa signification est que la transformée de Bargmann (et, par suite, la TFCT comme le spectrogramme) est entièrement caractérisée par la distribution de ses zéros.

3.2 Triangulation de Delaunay

Les zéros caractérisant complètement une TFCT, on peut s’attendre à ce que leur distribution dans le plan—appelée “stellar representation” en Mécanique Quantique [16] (voir aussi [12] pour un point de vue TF relié)—mette en évidence des propriétés distinctives attachées à la nature spécifique du signal analysé. Dans cet esprit, on peut obtenir une description géométrique de la structure TF d’un signal en utilisant des diagrammes reliant les zéros, le plus simple étant la *triangulation de Delaunay* [20]. Un exemple est donné en Figure 1. On y observe que l’homogénéité attendue dans le cas du bruit seul est brisée lorsqu’on ajoute un signal, la “région signal” étant ca-

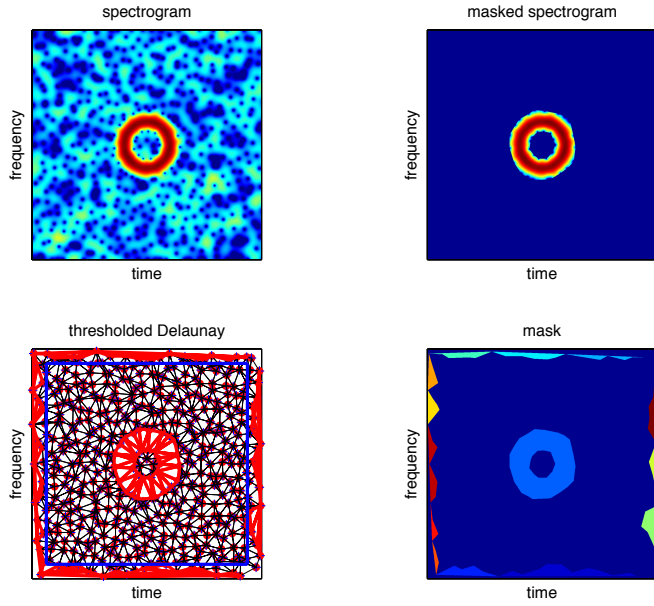


FIGURE 3 – Filtrage temps-fréquence — Signal synthétique (1/2). En haut à gauche : spectrogramme d’une fonction d’Hermite bruitée (RSB = 10 dB). En bas à gauche : triangulation de Delaunay construite sur les zéros du spectrogramme, avec les côtés anormaux (voir texte) marqués en rouge. En bas à droite : régions TF obtenues en concaténant les triangles de Delaunay adjacents à côtés anormaux, chaque région étant référencée par une couleur. En haut à droite : spectrogramme masqué en retenant comme “région signal” l’anneau bleu.

caractérisée non seulement par de grandes valeurs de l’amplitude mais aussi par des triangles de Delaunay dont au moins un des côtés est plus grand que dans les régions de bruit seul.

4 Filtrage temps-fréquence

4.1 Principe

Les considérations théoriques précédentes et le comportement observé en Figure 1 suggèrent que les “régions signal” soient construites en identifiant les triangles de Delaunay dont le comportement s’écarte de celui attendu pour du bruit, ce qui nécessite de caractériser cette situation de référence. C’est ce qui est fait en Figure 2 pour les longueurs de côté. Il en ressort que la distribution de celles-ci est essentiellement bornée par une longueur maximale $L_{max} \sim 2.2$ (trait plein), la probabilité qu’une longueur $d(z_m, z_n)$ dépasse 2 (trait pointillé) étant de l’ordre de 10^{-3} . Sélectionner les triangles de Delaunay sur la base d’un seuil relatif à leur plus grand côté est ainsi une façon simple d’identifier des domaines dont la concaténation définit les supports—bordés par des zéros—de masques TF binaires que l’on pourra appliquer à la TFCT avant reconstruction des composantes correspondantes.

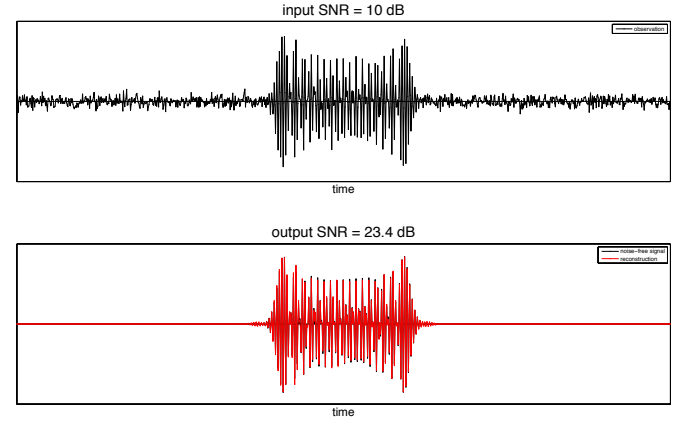


FIGURE 4 – Filtrage temps-fréquence — Signal synthétique (2/2). En haut : observation bruitée. En bas : reconstruction après masquage binaire de la TFCT de la Figure 3 (en haut à droite), à comparer à la fonction d’Hermite non bruitée.

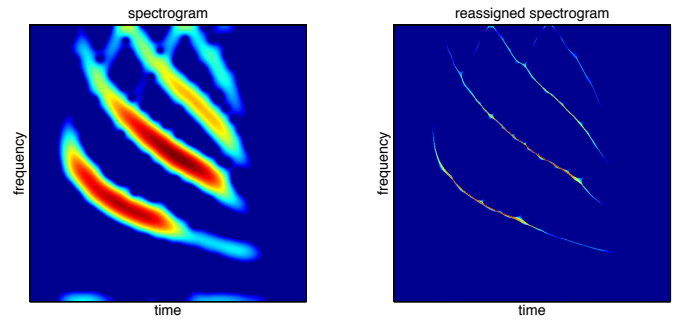


FIGURE 5 – Filtrage temps-fréquence — Signal réel (1/3). Gauche : spectrogramme du signal “benchmark” de chauve-souris. Droite : spectrogramme réalloué.

4.2 Algorithme

En se basant sur les éléments précédents, l’algorithme de filtrage TF est immédiat et peut s’exprimer comme suit :

1. Calculer la triangulation de Delaunay sur les zéros z_m de la TFCT ;
2. Identifier les côtés anormaux tels que $d(z_m, z_n) > 2$;
3. Garder les triangles ayant au moins un côté anormal ;
4. Grouper de tels triangles adjacents en régions connexes et disjointes \mathcal{D}_j ;
5. Multiplier la TFCT avec des masques $\mathbf{1}_{\mathcal{D}_j}(t, \omega)$;
6. Reconstruire les composantes région par région.

La dernière étape (reconstruction) peut être effectuée en utilisant, soit la formule 2D standard [5], soit la forme simplifiée

$$x_j(t) = \frac{1}{h(0)} \int_{(t, \omega) \in \mathcal{D}_j} F_x^{(h)}(t, \omega) \frac{d\omega}{2\pi} \quad (14)$$

qui est en usage par exemple en synchrosqueezing [21].

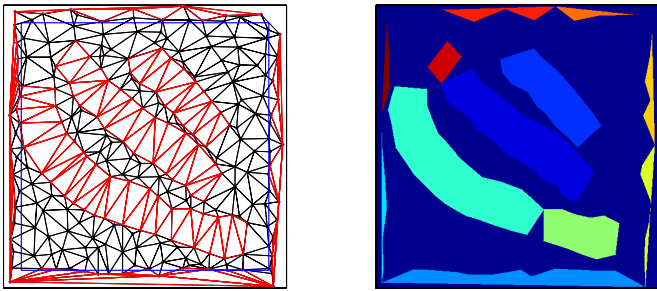


FIGURE 6 – Filtrage temps-fréquence — Signal réel (2/3). Gauche : triangulation de Delaunay construite sur les zéros du spectrogramme, avec les triangles anormaux (voir texte) marqués en rouge. Droite : Régions TF obtenues en concaténant les triangle adjacents anormaux, chaque région étant référencée par une couleur.

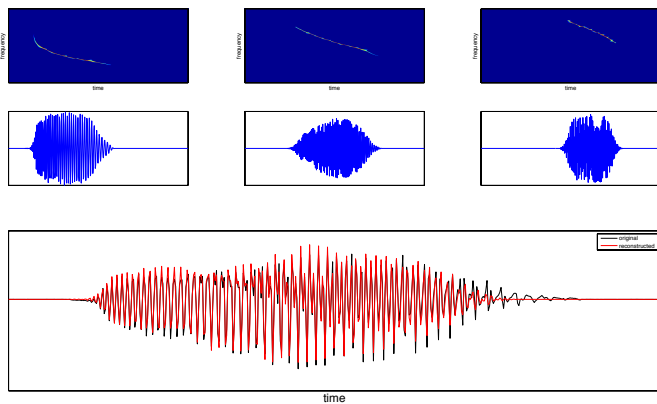


FIGURE 7 – Filtrage temps-fréquence — Signal réel (3/3). Haut : spectrogrammes réalloués masqués des 3 composantes principales. Milieu : les signaux correspondants obtenus par inversion des TFCT masquées. Bas : somme de ces 3 composantes, avec le signal original pour comparaison.

4.3 Exemples

Le premier exemple est celui d'une fonction d'Hermite dont la "trajectoire" est un cercle [9]. La procédure complète de filtrage (triangulation, sélection des triangles anormaux, regroupement et masquage) est résumée en Figure 3, avec le résultat correspondant de reconstruction en Figure 4.

Le deuxième exemple correspond au signal "benchmark" de chauve-souris*, dont le spectrogramme et sa version réallouée sont tracés en Figure 5. La sélection des régions correspondant aux différentes composantes est présentée en Figure 6, avec les reconstructions individuelles des 3 principales tracées en Figure 7, ainsi que leur recombinaison à comparer avec le signal original.

Références

- [1] F. Auger, P. Flandrin, Y.T. Lin, S. McLaughlin, S. Meignen, T. Oberlin and H.T. Wu, "Time-frequency reassignment and synchrosqueezing," *IEEE Sig. Proc. Mag.*, Vol. 30, No. 6, pp. 32-41, 2013.
- [2] V. Bargmann, "On a Hilbert space of analytic functions and an associated integral transform," *Commun. Pure Appl. Math.*, Vol. 14, pp. 187-214, 1961.
- [3] R.P. Boas, *Entire Functions*, Academic Press, New York, 1954.
- [4] E. Chassande-Mottin, I. Daubechies, F. Auger, and P. Flandrin, "Differential reassignment," *IEEE Signal Proc. Lett.*, Vol. 4, No. 10, pp. 293-294, 1997.
- [5] I. Daubechies, *Ten Lectures on Wavelets*, SIAM, Philadelphia (PA), 1992.
- [6] I. Daubechies, J. Lu, and H.-T. Wu, "Synchrosqueezed wavelet transforms : An empirical mode decomposition-like tool," *Appl. and Comp. Harm. Anal.*, Vol. 30, No. 1, pp. 243-261, 2011.
- [7] N. Delprat, B. Escudié, P. Guillemain, R. Kronland-Martinet, P. Tchamitchian, and B. Torrèsani, "Asymptotic wavelet and Gabor analysis : extraction of instantaneous frequencies," *IEEE Trans. on Info. Theory*, Vol. 38, No. 2, pp. 644-664, 1992.
- [8] P. Flandrin, *Time-Frequency/Time-Scale Analysis*, Academic Press, 1999.
- [9] P. Flandrin, "A note on reassigned Gabor spectrograms of Hermite functions," *J. Fourier Anal. Appl.*, Vol. 19, No. 2, pp. 285-295, 2012.
- [10] P. Flandrin, F. Auger, and E. Chassande-Mottin, "Time-frequency reassignment? From principles to algorithms," in *Applications in Time-Frequency Signal Processing*, A. Papandreou-Suppappola, Ed., chapter 5, pp. 179-203, CRC Press, Boca Raton (FL), 2003.
- [11] D. Gabor, "Theory of communication," *J. IEE*, Vol. 93, pp. 429-457, 1946.
- [12] T. J. Gardner and M. O. Magnasco, "Sparse time-frequency representations," *Proc. Nat. Acad. Sc.*, Vol. 103, No. 16, pp. 6094-6099, 2006.
- [13] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston (MA), 2011.
- [14] K. Husimi, "Some formal properties of the density matrix," *Proc. Phys. Math. Soc. Jpn.*, Vol. 22, pp 264-314, 1940.
- [15] H.J. Korsch, C. Muller, and H. Wiescher, "On the zeros of the Husimi distribution," *J. Phys. A : Math. Gen.*, Vol. 30, pp. L677-L684, 1997.
- [16] P. Lebeuf and A. Voros, "Chaos-revealing multiplicative representation of quantum eigenstates," *J. Phys. A Math. Gen.*, Vol. 23, pp. 1765-1774, 1990.
- [17] E.H. Lieb, "Integral bounds for radar ambiguity functions and Wigner distributions," *J. Math. Phys.*, Vol. 31, pp. 594-599, 1990.
- [18] Y. Lim, B.G. Shinn-Cunningham and T.J. Gardner, "Sparse contour representations of sound," *IEEE Signal Proc. Lett.*, Vol. 19, No. 10, pp. 684-687, 2012.
- [19] S. Meignen, T. Oberlin, and S. McLaughlin, "On the mode synthesis in the synchrosqueezing method," *Proc. European Signal Proc. Conf. EUSIPCO-12*, Bucharest (RO), 2012.
- [20] A. Okabe, B. Boots, K. Sugihara, and S.N. Chiu, *Spatial Tessellations — Concepts and Applications of Voronoi Diagrams* (2nd ed.), John Wiley, 2000.
- [21] G. Thakur, H.-T. Wu, "Synchrosqueezing-based recovery of instantaneous frequency from nonuniform samples," *SIAM J. Math. Anal.*, Vol. 43, pp. 2078-2095, 2011.
- [22] M. Toda, "Phase retrieval problem in quantum chaos and its relation to the origin of irreversibility I.," *Physica D*, Vol. 59, pp. 121-141, 1992.

*Merci à C. Condon, K. White et A. Feng du Beckman Institute of the University of Illinois.