

DM de Traitement du signal

Utilisation de d'antennes sismique pour l'identification de phases sismiques profondes

Attendu : Un rapport au format pdf (10 pages max) et le code python (ou un jupyter notebook incluant des zones de textes avec vos commentaires). L'ensemble est à déposer sur la plateforme moodle dans la section dédiée au DM. Préciser le(s) nom(s) dans le nom des fichiers rendus.

Date de rendu : **21 Janvier 2022 à 23h59**

Notation : elle se fera sur la justesse des résultats et figures obtenus, la qualité des commentaires et discussions, la qualité du rapport, des figures, et du code python. Un bonus sera pris en compte en fonction de ce qui a été rendu lors des évaluations formatives.

Binômes : Groupe de 2 personnes autorisé. Bonus de 1 point pour les personnes seules.

1. Présentation du sujet

L'utilisation d'antennes sismiques permet de réduire le niveau de bruit en utilisant des méthodes de sommation (et donc d'augmenter la capacité de détection de signaux de faible amplitude), et d'avoir également des informations sur la directivité du signal. En sismologie, les premiers déploiements d'antennes avaient pour objectif de détecter des explosions nucléaires dans le cadre des traités de non-prolifération, mais par la suite ils se sont multipliés pour la recherche fondamentale : étude de la source des séismes ou imagerie haute résolution de la structure de la Terre. Les antennes présentent notamment des intérêts dans les régions hyper-urbanisées dans lesquelles le niveau de bruit est élevé, ainsi que pour détecter des signaux de faible amplitude, difficilement exploitable sur une seule station.

Dans le cadre de cette étude, on utilisera pour se faire des données de 13 stations constituant l'antenne de Gräfenberg (GRF) installée depuis les années 80 dans le sud de l'Allemagne. L'antenne a une ouverture de 50 km dans la direction EW et 100 km dans la direction NS. Le but du travail est de s'intéresser à des phases sismiques de très faible amplitude, apportant des informations importantes sur la couche D'' (couche à la base du manteau) comme montrer dans l'étude de Cobden et Thomas, 2013 (mis à disposition sur moodle). L'objectif est également de se familiariser avec le principe de traitements d'antennes, ce sujet est très bien présenté dans l'article Rost et al. 2002 (mis également à disposition sur moodle).

Sur moodle le fichier "*stations_GR.dat*" contient les coordonnées des stations de cette antenne sismique. Le fichier "*stations_GE.dat*" contient les coordonnées des stations d'une autre antenne qui ne sera utilisée que dans la partie 4.

2. Visualisation des données et pré-traitements

Sur moodle, vous disposez de plusieurs jeux de données `events_*.txt`. Chacun de ces jeux comportent les enregistrements (composante verticale) pour un séisme donné à chacune des stations. Chaque fichier contient 13 colonnes, chaque colonne contient l'enregistrement à une des stations. Si une station n'a pas d'enregistrement, la colonne ne contient que des 0.

Choisissez un de ces jeux de données sur lesquels vous ferez les analyses. Le premier échantillon des séries temporelles correspond au temps origine du séisme considéré + 600s. Les informations sur les différents séismes (localisation, magnitude, temps origine, distance épacentrale (δ), back-azimuth (θ)) sont dans le fichier "`events_info.dat`". Chaque séisme est identifié par un numéro correspondant à celui des fichiers `*.txt`. Les fichiers `*_phase.txt` contiennent le temps de trajet des ondes sismiques arrivant dans la fenêtre d'étude.

2.1 La fréquence d'échantillonnage de ces données est de 100 Hz. Quelle est la fréquence de Nyquist correspondante ? Quelles est la fréquence maximale que l'on pourra analyser dans ces données ?

2.2 Représenter (sur une même figure) les séries temporelles des enregistrements à l'ensemble des stations. On se focalisera sur les premières ondes sismiques visibles. Commenter.

2.3 Appliquer un filtre passe-bande de type Butterworth entre 1 et 10 secondes. Faire une extraction des données sur une fenêtre temporelle de 30s autour des premières arrivées. La suite des analyses s'effectuera sur ces fenêtres de signaux filtrés.

3. Traitement d'antenne

3.1 Introduction

Le traitement d'antenne se base généralement sur l'hypothèse que l'énergie arrivant à l'antenne peut être approximée par un front d'onde plan (figure 1), c'est-à-dire que tous les rais arrivant à l'antenne sont parallèles. Cette approximation revient à supposer que l'ouverture de l'antenne est petite devant la distance de propagation. Ce qui peut raisonnablement être considéré comme valide dans le cas de télé-séismes.

La figure 1a et 1b représente ce front d'onde dans les plans vertical et horizontal, le front d'onde sera caractérisé par 2 paramètres : son angle d'incidence sous l'antenne i , et le back azimuth θ qui représente l'angle entre le nord et la direction du grand cercle reliant la source et le centre de l'antenne.

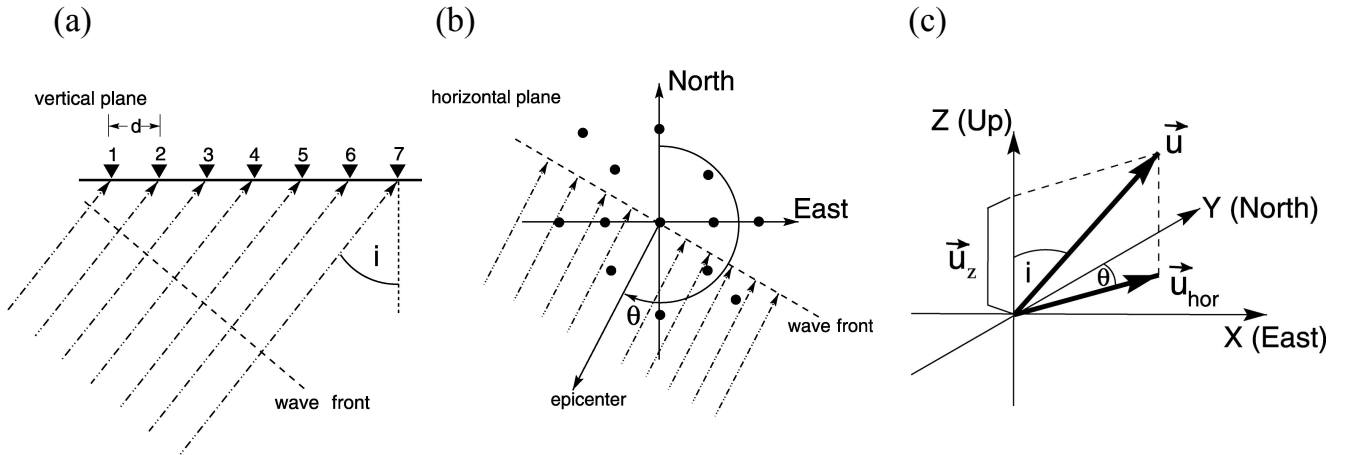


Figure 1 : (a) et (b) représentation verticale et horizontale d'un front d'onde plan incident sur l'antenne sismique avec un angle i et arrivant avec un back-azimut de θ , (b) représentation du vecteur lenteur \mathbf{u} . Le vecteur lenteur est normal au front d'onde incident.

Une onde sismique plane peut être représenté mathématiquement par un vecteur qui décrit à la fois sa direction de propagation et sa vitesse. Pour plus de simplicité, on travaillera avec la lenteur \mathbf{u} qui est l'inverse de la vitesse. La figure 1c montre la représentation de ce vecteur dans l'espace.

$$\mathbf{u} = \begin{pmatrix} u_x \\ u_y \\ u_z \end{pmatrix} = u_h \begin{pmatrix} \sin \theta \\ \cos \theta \\ 1/\tan i \end{pmatrix}$$

où u_h est la lenteur horizontale, l'angle i d'incidence sous l'antenne et θ le back-azimut.

Par la suite, on considérera une antenne purement horizontale, et on oubliera la composante verticale du vecteur \mathbf{u} :

$$\mathbf{u} = u_h \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix} = \mathbf{u}_h$$

Connaissant la lenteur, on peut définir le délai τ_i entre le temps d'arrivée de la phase sismique à la station i et celui à la station de référence (station au centre de l'antenne) :

$$\tau_i = \mathbf{r}_i \cdot \mathbf{u}_h$$

Chaque station est repérée par un vecteur \mathbf{r}_i représentant la distance entre la station i et celle de référence.

Si la structure sous l'antenne est homogène, les formes d'ondes enregistrées aux différentes stations de l'antenne seront quasiment identiques, à l'exception d'un délai en temps et le bruit local. Soit x_0 la trace sismique de référence (c'est-à-dire enregistrée au centre de l'antenne), la trace sismique enregistrée à la station i s'écrit :

$$x_i(t) = x_0(t - \tau_i) + n_i(t)$$

où $n_i(t)$ représente le bruit à la station i .

Si on corrige la trace par ce délai τ_i , on retrouve alors la trace de référence (plus du bruit également décalé temporellement) :

$$x_i(t + \tau_i) = x_0(t) + n_i(t + \tau_i)$$

Après avoir définie la position de référence (en général le centre de l'antenne, ou la position de la station la plus proche du centre), il est possible de corriger tous les enregistrements des stations de l'antenne avec des délais qui dépendent de la lenteur horizontale u_h et du back-azimut θ . Toutes les traces se ressembleront au bruit résiduel près.

3.2 Beamforming

Comme vu précédemment après correction des décalages temporels, tous les enregistrements (aux différentes stations) vont ressembler à la trace de référence, et donc ils vont pouvoir se sommer constructivement, à l'inverse du bruit aléatoire qui lui se sommera destructivement.

On appelle cette sommation, après correction des délais, "un beam". Il est défini de la façon suivante :

$$\begin{aligned} b(t) &= \frac{1}{N} \sum_{i=1}^N x_i(t + \tau_i) \\ &= x_0(t) + \frac{1}{N} \sum_{i=1}^N n_i(t + \tau_i) \end{aligned}$$

Cette sommation permet d'amplifier le rapport entre le signal cohérent et le bruit incohérent, et donc d'obtenir un meilleur rapport signal sur bruit. Ceci est un des nombreux avantages des antennes par rapport à une station seule.

3.2.1 Représenter, sur une même figure, l'ensemble des traces, ainsi que la trace sommée (sans correction de délais).

3.2.2 Choisir la station la plus proche du centre de l'antenne comme trace de référence, et calculer les délais en utilisant la corrélation.

3.2.3 Utiliser ces délais pour construire le meilleur "beam" (sommation corrigée des délais). On pourra utiliser la convolution pour construire la trace corrigée du délai. Comparer avec la figure précédente.

3.2.4 En déduire les paramètres des ondes (u_h et θ) correspondant à ces délais. Comparer avec la valeur de θ donnée pour l'évènement que vous aurez choisi (noté baz dans le fichier *events_info.txt*).

3.2.5 Quels peuvent-être à votre avis les avantages et les limitations d'une telle technique ?

3.3 Vespagramme (velocity spectral analysis)

D'autres méthodes ont été développées pour mesurer le vecteur lenteur du champs d'onde avec comme objectifs de localiser des séismes et d'identifier des phases sismiques. La technique VESPA calcule des beams pour différentes valeurs de lenteur ($u_{\min} \leq u_h \leq u_{\max}$) et une valeur de back-azimut fixée :

$$v_u(t) = \frac{1}{N} \sum_{i=1}^N x_i(t + \tau_i^u)$$

où τ_i correspond au délai calculé pour la station i à partir des valeurs de u_h et θ . On construit ainsi un "vespagramme".

3.3.1 Construire le vespagramme pour l'évènement considéré. Considérer des valeurs de lenteur comprise entre 2 et 8 s/degré. Faire le calcul pour la valeur de θ estimée précédemment, et une 2e fois pour une valeur un peu différente (celle donnée pour le séisme si elle diffère de celle calculée ou une valeur perturbée).

3.3.2 Représenter les 2 vespagrammes ainsi calculés en utilisant la fonction *pcolormesh* de python. Commenter. Vous pouvez comparer les figures obtenues avec les exemples des 2 articles mis à disposition sur le moodle.

Les enregistrements, que vous avez, correspondent à une configuration comparable à l'article de Cobden et Thomas 2013, c'est-à-dire dans laquelle les phases P, PcP et éventuellement PdP ont des temps d'arrivée proches. Pour rappel, la phase PcP correspond à une onde P se réfléchissant sur l'interface manteau-noyau, et la phase PdP est une phase se réfléchissant sur la couche D". La couche D" est une couche se situant à la base du manteau.

3.3.3 Combien de phases sismiques arrivez vous à distinguer ? La phase PdP devraient (si on l'a voit) arriver entre les phases P et PcP. Commenter.

3.3.4. Quels peuvent-être à votre avis les avantages et les limitations d'une telle technique ?

3.4 Sommation en phase (PWS : Phase weighted stack)

C'est une autre méthode pour mesurer le vecteur lenteur du champs d'onde, elle se base sur le signal analytique $S(t)$ défini par :

$$S(t) = x(t) + i \mathcal{H}[x(t)] = A(t) \exp(i\phi(t))$$

où $x(t)$ est la trace sismique considéré, \mathcal{H} correspond à la transformée de Hilbert. Le signal analytique peut s'écrire également en terme d'amplitude $A(t)$ et de phase instantanée $\phi(t)$.

La somme de 2 signaux analytiques à un temps t fixé ($t = \tau$) sera maximum si leur phases instantanées sont égales, c'est-à-dire si les 2 signaux sont cohérents. Ceci est d'autant

plus vrai si les amplitudes des 2 signaux analytiques sont comparables, ce qui peut être pris en compte en réalisant une normalisation à chaque pas de temps. Ceci permet d'amplifier les signaux cohérents de faibles amplitudes (comparés au bruit). On peut ainsi définir une sommation en phase dans le cas de N signaux :

$$c(t) = \frac{1}{N} \left| \sum_{i=1}^N \exp(i\phi_i(t)) \right|$$

où $c(t)$ représente la cohérence des signaux en fonction du temps indépendamment de leur amplitude. Cette fonction est comprise entre 0 et 1, respectivement pour du bruit incohérent et des signaux parfaitement cohérents.

En se basant sur cette propriété, on peut définir une sommation pondérée en phase de la façon suivante :

$$v_{PWS}(t) = \frac{1}{N} \sum_{i=1}^N s_i(t) \left| \frac{1}{N} \sum_{j=1}^N \exp(i\phi_j(t)) \right|^n$$

Cette sommation agit comme un filtre en cohérence (similarité des signaux). $s_i(t)$ est la trace à la station i corrigée du délai correspondant aux valeurs de u_h et θ , et n représente le caractère plus ou moins abrupte du filtre un peu comparable à ce qu'est l'ordre pour les filtres en fréquence. La phase instantanée $\phi_j(t)$ est obtenue en calculant le signal analytique de la trace $s_j(t)$.

3.4.1 Calculer cette sommation pondérée en phase pour différentes valeurs de lenteur horizontal u_h comme dans la section précédente (prendre le même intervalle de valeur) et la même valeur fixée de θ .

3.4.2 Représenter cette sommation en utilisant la fonction *pcolormesh*, et comparer avec la figure précédemment obtenue. Commenter.

3.4.3. Quels peuvent-être à votre avis les avantages et les limitations d'une telle technique ?

3.5 Analyse en lenteur et en azimuth

Dans les méthodes présentées précédemment, on fixait la valeur du back-azimut. Dans cette section, on fera une recherche systématique en lenteur et en back-azimut, afin de trouver les valeurs optimales de u_h et θ .

3.5.1 Définir une grille en lenteur et back-azimut.

3.5.2 Pour chacun des couples (u_h et θ), calculer le beam associé (cf section 3.2), et son énergie définie par :

$$E = \int_{-\infty}^{+\infty} b^2(t) dt$$

Pour cela, on considérera, une fenêtre de 3 secondes autour de la première phase (ou d'une des phases suivantes), pour réduire les calculs et pour éviter d'inclure plusieurs phases ayant potentiellement différentes valeurs de lenteur.

3.5.3 Représenter cette Energie en fonction de la lenteur et du back-azimut à l'aide de la fonction *pcolormesh*. Au besoin, utiliser une échelle de couleur logarithmique.

3.5.4 En déduire les valeurs optimales de (u_h et θ). Commenter.

4. Fonction de transfert de l'antenne

Il existe différentes configurations spatiales pour les antennes, elles sont généralement optimisées pour certains types d'ondes, certaines gammes de fréquences, certaines conditions de bruit et la région d'étude.

La fonction de transfert de l'antenne décrit sa sensibilité et sa résolution pour des signaux ayant des fréquences et des lenteurs différentes. Elle peut être définie par :

$$|A(\mathbf{k} - \mathbf{k}_0)|^2 = \left| \frac{1}{N} \sum_{i=1}^N \exp(i(\mathbf{k} - \mathbf{k}_0) \cdot \mathbf{r}_i) \right|^2$$

avec \mathbf{k} le vecteur nombre d'onde de l'énergie arrivant sur l'antenne ($\mathbf{k} = \omega \mathbf{u}$), \mathbf{k}_0 le vecteur de référence, \mathbf{r}_i le vecteur de localisation des stations, et N le nombre de stations de l'antenne considérée.

4.1 Définir une grille en lenteur et back-azimut. On pourra considérer le même intervalle de lenteur que précédemment, et un back azimut entre 0 et 360°.

4.2 Pour chacun des couples (u_h et θ), calculer la fonction ci-dessous. On considérera une fréquence de 1 Hz, et $\mathbf{u}_0 = 0$.

4.3 Représenter cette fonction de transfert de l'antenne en fonction de la lenteur et du back-azimut à l'aide de la fonction *pcolormesh*. Au besoin, utiliser une échelle de couleur logarithmique.

4.4 Refaire la même opération pour une autre antenne également localisée en Allemagne. Les coordonnées des stations sont dans le fichier stations_GE.dat. Faire une figure (à la même échelle) représentant la configuration des 2 antennes. Commenter.