

## Imagerie Sismique de la Terre Profonde

## 2- Tomographie des temps de parcours télésismiques ("Travel time tomography")

Barbara Romanowicz Chaire de Physique de l'Intérieur de la Terre Collège de France, Paris

5 Novembre 2019

## Problème direct / Problème inverse

- Données (observations) en temps et ou en espace: -> vecteur données (d)
- Paramètres physiques du modèle: modèle (m)





## Problème direct / Problème inverse

- Données (observations) en temps et ou en espace: -> vecteur données (d)
- Paramètres physiques du modèle: modèle (**m**)





## Problème direct / Problème inverse

- Cas linéaire (ou linéarisé):
- `Données (observations) en temps et ou en espace: -> vecteur données ( $\delta d$ )
- Paramètres physiques du modèle: -> vecteur modèle ( $\delta \mathbf{m}$ )

 $\delta d = G \delta m$ 

Solution par "moindres carrés":

 $G^{T}\delta d = G^{T}G \delta m$ mais G<sup>T</sup>G N'est pas inversible en général => Solution non unique  $\delta m_{est} = (G^T G)^{-1} G^T \delta d$ 

## Exemple du trésor enterré sur la plage





## Exemples de problèmes directs

#### Linéaire

## • $y(x) = m_1 x + m_2$

 e.g. prédire le champ de gravité étant donné la densité d'un objet de localisation et taille donnée



#### Non-linéaire

• 
$$y(x) =$$
  
 $sin(xm_1) + m_1m_2$ 

 e.g. prédire la forme d'onde étant donné un modèle 3D particulier de la structure de la Terre



## Problème inverse sous-determiné

 Avec une seule donnée (un point sur la droite, on ne peut pas déterminer de façon unique les 2 paramètres de la droite

• Infinité de solutions pour m<sub>1</sub> and m<sub>2</sub>!



## Problème inverse sur-déterminé

Trouver une droite passant par un ensemble de points у `(x<sub>j</sub>,y<sub>j</sub>) • Avec >2 points, la  $y(x) = m_1 x + m_2$ ligne droite ne passe pas forcément par tous les points. • Laquelle choisir? • Quelle est la meilleure solution...? Х

## La fonction coût

- On introduit une fonction coût  $\phi$  dont la valeur est minimale lorsque la solution (la droite choisie) explique les données de façon optimale
- points de mesure par rapport à la solution:  $\varepsilon_i$
- => par exemple on minimise la somme des distances au carré:

Trouver une droite passant par un ensemble de points У



## Choix de fonctions coût

- L1
  - Somme des valeurs absolues des résidus
  - Atténue l'effet des données erronnées (*outliers*)

$$\varphi(m) = \sum_{j=1}^{N_d} \left| d_j - g_j(m) \right|$$

- L<sup>2</sup>
  - Somme des carrés des résidus
  - Tend à favoriser les "outliers"
  - "Moindres-carrés"
- L∞
  - Minimiser le résidu le plus grand
  - Solution en général ni très mauvaise ni très bonne
  - "Minimax"

$$\varphi(m) = \sum_{j=1}^{N_d} (d_j - g_j(m))^2$$

$$\varphi(m) = max \big| d_j - g_j(m) \big|$$

Données: temps de parcours télésismiques des ondes P Théorie des rais

 $\delta d_{i} = \sum_{j=1}^{M} G_{ij} \delta m_{j} \quad i = 1, N \quad \stackrel{\delta \vec{d} = \text{Vecteur des données = perturbations } \delta \mathsf{T}_{i}}{\delta \vec{m} = \text{vecteur du modèle = perturbations en vitesse } \delta \mathsf{v}_{j}/\mathsf{v}_{j}}$   $G_{ij} = -\frac{l_{ij}}{v_{0}^{j}}$ 



"Solution par moindres carrés"

$$\delta \hat{m} = (G^T G)^{-1} G^T \delta d^{-1}$$

Equivalent à minimiser la fonction coût (L<sup>2</sup>):

$$\varphi(m) = \sum_{j=1}^{N_d} (d_j - g_j(m))^2$$

Résidus (au carré)

## Erreurs dans les données

- Si l'erreur de mesure est plus grande sur un des points que sur d'autres, la solution ne doit pas nécessairement passer près de ce point
- => la fonction coût doit en tenir compte





# Trouver une droite qui passe par un ensemble de points (x,y)

- Problème linéaire:
- $y_j = m_1 x_j + m_2$  $d_j = (x_j, y_j)$



Notation matricielle:

$$\varphi = (d - Gm)^T C_D^{-1} (d - Gm)$$

Minimiser  $\varphi$  par rapport à m = (m<sub>1</sub>, m<sub>2</sub>)

$$\frac{\partial \varphi}{\partial m} = 0$$

La solution est:

$$m_{est} = (G^T C_D^{-1} G)^{-1} G^T C_D^{-1} d$$

Solution par "moindres carrés" pondérée: mest

Données: temps de parcours télésismiques des ondes P Théorie des rais

$$\delta d = G \delta m$$

 $\delta \vec{d}$  = Vecteur des perturbations  $\delta T_i$  $\delta \vec{m}$  = vecteur du modèle = perturbations en vitesse  $\vec{\sigma}$  = vecteur des erreurs sur les données

$$C_D = \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$$

$$G_{ij} = -\frac{l_{ij}}{v_0^j} \qquad C_D^{-1}\delta d = C_D^{-1}G\delta m^*$$

$$G^{T}C_{D}^{-1}\delta d = G^{T}C_{D}^{-1}G\,\delta m_{est}$$

$$\delta m_{est} = (G^T C_D^{-1} G)^{-1} G^T C_D^{-1} \delta d$$

"Solution par moindres carrés pondérés"

-> Equivalent à minimiser la fonction coût (L<sup>2</sup>):

$$\varphi(m) = \sum_{j=1}^{N_d} \frac{\left(d_j - g_j(m)\right)^2}{\sigma_j^2}$$



"Solution par moindres carrés pondérés"

$$\delta m_{est} = (G^T C_D^{-1} G)^{-1} G^T C_D^{-1} \delta d$$



Certains éléments de G<sub>ij</sub> sont nuls ( l<sub>ij</sub>=0) -> La vitesse dans ces blocs ne peut pas être déterminée de manière unique

->  $G^{T}C_{D}^{-1}G$ , en général, est singulière:

- => Régularisation:

Par exemple, on ajoute un terme  $\epsilon^{2''}$ à la diagonale de G<sup>T</sup>G, suffisamment grant pour que la matrice ( $G^TG$ +  $\epsilon^2 I$ ), où I est la matrice identité MxM, soit inversible.

$$\delta m_{est} = (G^T C_D^{-1} G + \varepsilon^2 I)^{-1} G^T C_D^{-1} \delta d$$

solution "Par moindres carrés attenués et pondérés"

Solution par moindres carrés atténuée et pondérée

$$\delta m_{est} = (G^T C_D^{-1} G + \varepsilon^2 I)^{-1} G^T C_D^{-1} \delta d$$



Inverse "généralisée:

$$\begin{split} \delta m_{est} &= G^{-g} \ \delta d \quad \text{où:} \quad G^{-g} = (G^T C_D^{-1} G + \varepsilon^2 C_M^{-1} I)^{-1} G^T C_D^{-1} \\ m_{est} &= < m > + (G^T C_D^{-1} G + \varepsilon^2 C_M^{-1} I)^{-1} \ G^T C_D^{-1} (d - G < m >) \\  = \text{Modèle de référence} \end{split}$$

#### En résumé:

- Pour obtenir des résultats robustes, il faut pondérer les données de manière adéquate pour tenir compte des erreurs et des redondances (C<sub>D</sub><sup>-1</sup>).
- Il faut aussi spécifier les caractéristiques *a-priori* du modèle en précisant le type de régularisation (C<sub>M</sub><sup>-1</sup>)
- Ceux-ci ne sont pas en général déterminés de manière unique, et donc peuvent être "ajustés", souvent de manière subjective
- Une fois qu'on a obtenu une solution, on doit se poser les questions suivantes:
  - Le modèle obtenu est-il fiable?
  - Les données permettent-t-elles de déterminer les paramètres choisis?

#### "Trade offs": exemple de la solution par moindres carrés

- Lorsqu'on augmente le facteur de damping ε<sup>2</sup>, on augmente le poids de la norme (ou amplitude) du modèle que l'on souhaite contrôler
  - Cela diminue l'amplitude des variations du modèle, mais cela augmente les résidus (les données sont moins bien expliquées)
  - On obtiendra un modèle qui s'écarte peu du modèle de référence
- Pour choisir ε<sup>2</sup> de façon optimale, il faut explorer une suite de modèles



#### Implémentation pratique:

- On résoud le problème inverse pour une suite de valeurs des paramètres de régularisation
- On représente le résidu en fonction de la norme (amplitude) des perturbations du modèle pour ces différents cas
- On choisit un modèle près du coude de la courbe "L"



#### Matrice de résolution

- Comment l'image est elle dégradée?
- Quel est le plus petit objet qu'on peut résoudre?
- Matrice de résolution R:

$$\hat{m} = G^{-g} d^{obs} = G^{-g} G m_{true} = R m_{true}$$
$$R = G^{-g} G$$

- R est une sorte de filtre spatial que l'on applique au vrai modèle test pour evaluer la dégradation du modèle estimé
  - Si R=I la résolution est parfaite
  - Les termes en dehors de la diagonale nous renseignent sur la dégradation de l'image.

#### Matrice de résolution

- On considère un modèle particulier, souvent de type "damier"
- On le "passe" par la matrice de résolution pour voir ce que le processus d'inversion dégrade
- Cet exercice ne permet pas de tenir compte des erreurs dans la théorie ni dans l'estimation des incertitudes sur les données



🛚 🕬 🛛 • Si le résultat est mauvais, le modèle obtenu avec les vraies données n'est pas de bonne qualité

🦮 • Si le résultat est bon, cela ne prouve pas que ce modèle est robuste



#### Incertitude sur le modèle

- Si les données sont indépendantes, avec une variance commune  $\sigma_d^2$ :
  - Etant donné que:

 $\hat{m} = Md + v$ 

• On peut écrire

 $[\operatorname{cov} m] = M[\operatorname{cov} d]M^T$ 

• Pour la solution par moindres carrés:

 $[\operatorname{cov} m] = [[G^{T}G]^{-1}G^{T}]\sigma_{d}^{2}[[G^{T}G]^{-1}G^{T}]^{T} = \sigma_{d}^{2}[G^{T}G]^{-1}$ 

# Ingrédients essentiels pour la tomographie sismique globale ou régionale

- Quelles données ? -> illumination des structures
  - Ondes de volume/ondes de surface
  - Temps de parcours/formes d'onde



- Théorie de propagation des ondes /méthode d'inversion
  - Théorie des rais/effets de fréquence finie
  - Calcul plus ou moins précis du champ des ondes sismiques
  - Méthode d'optimisation (linéaire or quadratique) ou....Monte Carlo
- Paramétrisation du modèle
  - Physique: hypothèses (anisotropie?..., atténuation?...)
  - Géométrique: paramétrisation locale ou globale?..

Paramétrisation sur la sphère

$$f(\theta, \varphi) = \sum_{i=1}^{i=M} c_i \cdot g_i (\theta, \varphi)$$

 $\theta$  = colatitude = 90°-latitude  $\varphi$  = longitude

- Choix des fonctions de base  $g_i$ 
  - Blocs (fonctions "locales")

$$g_i(\theta, \varphi) = \begin{array}{cc} 1 & si(\theta, \varphi) \ \texttt{a} \ l'intérieur}{0} \ \text{du bloc i} \\ si(\theta, \varphi) \ \texttt{a} \ l'extérieur \end{array}$$

# Paramétrisation globale sur la sphère en harmoniques sphériques

$$f(\theta, \varphi) = \sum_{i=1}^{i=M} c_i \cdot g_i (\theta, \varphi) \qquad \begin{array}{l} \theta = \text{colatitude} = 90^\circ\text{-latitude} \\ \varphi = \text{longitude} \end{array}$$

• *Harmoniques sphériques*: base complète orthogonale de fonctions régulières sur la sphère

$$g_{i}(\theta,\varphi) = Y_{l}^{m}(\theta,\varphi) = P_{l}^{m}(\theta) \sum_{\sin m\varphi}^{\cos m\varphi}$$
  
Fonctions de Legendre  
Polynômes en cos( $\theta$ )  $f(\theta,\varphi) = \sum_{l} \sum_{m} (c_{l}^{m}\cos m\varphi + s_{l}^{m}\sin m\varphi) P_{l}^{m}(\theta)$ 

de rayon unité 
$$\begin{bmatrix} c_l^m \\ s_l^m \end{bmatrix} = \iint_{\Omega} f(\theta, \varphi) Y_l^{m*}(\theta, \varphi) d\Omega$$

 $\Omega$  = sphère de rayon unité

Harmoniques sphériques de degré 2





# Les premiers modèles tomographiques de l'intérieur de la terre

- Données: anomalies de temps de parcours des ondes P
- *Global*: Dziewonski, Hager and O'Connell, 1977
- *Régional*: Aki, Christofferson and Husebye, 1977

# Les premiers modèles tomographiques de l'intérieur de la terre

- Global: Dziewonski, Hager and O'Connell, 1977
- <u>Données</u>: anomalies de temps de parcours des ondes P
  - Principalement: Bulletins ISC\* 1967-1970
  - Mesurées par rapport aux tables de Jeffreys-Bullen

\*International Seismological Center http://www.isc.ac.uk/



-> Temps de parcours pour 50,000 séismes enregistrés sur bande bande magnétique (ISC)

-> 2,000,000 temps de parcours d'ondes P dans 1400 stations distribuées globalement

-> Hypothèse: erreurs aléatoires

-> 27° < $\Delta$ <105° écarts  $\Delta$ t < 5s par rapport aux tables de Jeffreys-Bullen

- ->II reste 728,072 données de  $\Delta t$ 
  - Correction d'ellipticité
  - Correction d'anomalies de station



Dziewonski, Hager and O'Connell (1977)



Dziewonski, Hager and O'Connell (1977)



Anomalies relatives dVp/Vp par rapport au modèle moyen à 2500 km de profondeur



Anti-corrélation des anomalies de gravité (géoide) et des anomalies de vitesses sismiques dans les derniers 1800 km du manteau inférieur



Coef.de correlation ~ - 0.66

Dziewonski, Hager and O'Connell (1977)

$$\Delta v_i(\theta,\,\phi) = \alpha \Delta \rho_i(\theta,\,\phi)$$

Alpha négatif:

- Du aux variations latérales de composition
- Présence d'éclogite dans les régions de vitesse plus rapide (provenant de la croûte basaltique)
  Subduction
- Ou bien interpretation dynamique: effet de la convection – dû à la déformation de la surface et de la CMB dans un milieu visqueux

#### Corrélations dans les structures de degré 2 & 3



Structure sismique à 2800 km de profondeur

Subduction 0 – 120 Ma Richards & Engerbretsen, 1992



Dziewonski, Lekic & Romanowicz, 2010

## Tomographie régionale par temps de parcours télesismiques:



Aki, Christofferson et Husebye (1977)

Modèles globaux par tomographie des temps de parcours d'ondes P et S

- Paramétrisation locale du modèle (blocs de vitesse constante ou splines/ondelettes)
  - Grand nombre d'inconnues (par exemple blocs de 2,5°x2,5°x200 km Grand et al., 1997; 3°x3°x150 km, Karason et van der Hilst, 2000)
    - Inversion par méthodes itératives (LSQR, SIRT)
    - Paramétrisation en blocs de taille variable

#### Densité de rais, données ondes P bulletins ISC, télésismiques



CMB= limite noyau-manteau

Courtesy of D. Vasco

Données ISC globales, étude régionale Pacifique Ouest



Données ISC globales, étude régionale Pacifique Ouest



Conclusion: Discontinuité de 660 km: barrière partielle à la pénétration dans le manteau inférieur

Fukao et al., 1992



Karason and van der Hilst, 2000

#### Anomalies relatives de vitesses des ondes P et S

Profondeur = 1300 km



#### Anomalies relatives de vitesses des ondes P (dVp/Vp)

Profondeur = 1300 km



Karason and van der Hilst, 2000





slow	fast



Analyse de résolution du modèle Vp de Karason et van der Hilst, 2000

Karason & van der Hilst, 2000



Karason and van der Hilst, 2000

Note - *Pour Pdiff: noyaux de sensibilité "de fréquence finie"* 



non seulement à 660 km, mais dans certaines régions (e.g. Java), vers

Fukao, 2001



- •11,3 millions de temps de parcours P (ISC 1964-2008)
- 60,000 temps de parcours mesurés manuellement ouest Pacifique
- 15,000 temps différentiels PP-P (stations large bande, cross-correlation
- 4300 temps differentiels entre stations fond de mer
- Noyaux de sensibilité de "fréquence finie"

Fukao and Obayashi, 2013









Tonga Kermadec



Le doublement du "slab" est du à la courbure du slab de Tonga (illustration origami)

★ séismes les plus profonds

Fukao and Obayashi, 2013





#### Amérique du Sud





Fukao and Obayashi, 2013



#### Relation entre les images tomographiques et l'histoire de la subduction



□volume de lithosphère engloutie (normalisé) 1-2 10<sup>8</sup> km<sup>3</sup>

+0.56% Wen and Anderson, 1995

## Evolution spatio-temporelle des plaques tectoniques

- On combine les images tomographiques des "slabs" identifiés comme tels avec les reconstructions géologiques pour prolonger dans le passé l'histoire de la subduction.
- Régions de collision continentale, telles la collision Inde/Asie (eg. Replumaz et al., 2005)
  - Déformation importante des 2 plaques, raccourcies de plusieurs centaines de kilomètres voir plus
- Déformation représentée par des modèles de blocs lithosphériques cohérents déterminés à partir des traces de failles obtenues par des études sur le terrain et les observations satellitaires (SPOT, LANDSAT)
- Reconstruction des mouvements de blocs par pas de quelques Millions d'années correspondants aux changements majeurs dans le régime de déformation

## Position de l'Inde et des blocs du sud-Est de l'Asie depuis 50 Ma



A chaque époque (50-5Ma):

-> Rotation des blocs sans déformation

-> Puis ajustements de la surface des blocs le long de certaines frontières tectoniques de caracteristiques connues

Replumaz et al., EPSL, 2005

Modèle de Vp , Karason et van der Hilst (2000)



Inversion globale de temps de parcours d'ondes P (8 Millions), pP (~1 Million) et PKP (~1Million) pour 300,000 séismes de 1964 à 2000 (ISC)

Blocs de taille variable: 0,5° x 0,5° dans les régions de couverture dense

Théorie: "fréquence finie" permettant de rajouter des données de plus longue période (temps différentiels PP-P)

Replumaz et al., EPSL, 2005





TRIASSIC 200 million years ago



Replumaz et al., EPSL, 2005