

# La formulation classique

La tomographie des temps de première arrivée a pour objectif de retrouver un modèle de vitesse de propagation des ondes sismiques à partir de temps de première arrivée mesurés à la surface. Mathématiquement, cela consiste à résoudre un problème inverse, à savoir retrouver les paramètres d'un modèle physique qui expliquent au mieux les données observées. La résolution de ce problème nécessite une étape de modélisation, appelée problème direct, qui permet, à partir d'un modèle donné et en utilisant les lois de la physique, de simuler des données observées (Fig. 2.1).

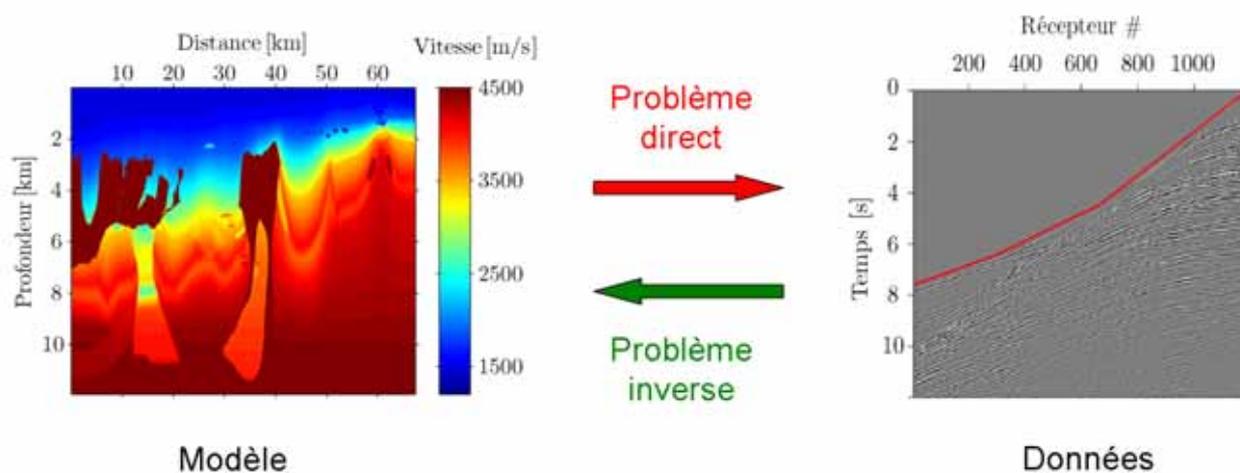


Fig. 2.1 – Les problèmes direct et inverse en tomographie des temps de première arrivée relient un modèle de vitesse aux temps de première arrivée pointés (en ligne rouge) sur un sismogramme.

Dans ce chapitre, nous développons tout d'abord la formulation mathématique classiquement utilisée pour la définition et la résolution des problèmes direct et inverse. Nous présentons ensuite la mise en œuvre pratique d'un algorithme de tomographie de temps de première arrivée défini à partir de cette formulation. Enfin, nous exposons les raisons qui nous ont conduits à considérer une nouvelle formulation pour la résolution du problème inverse.

## 2.1 Le problème direct

L'objectif de cette partie est de déterminer la relation physique  $\mathbf{g}$  qui existe entre les données  $\mathbf{d}$  et le modèle physique étudié  $\mathbf{m}$

$$\mathbf{d} = \mathbf{g}(\mathbf{m}). \quad (2.1)$$

Dans le cadre de ce travail, les données sont les temps de première arrivée et le modèle étudié est un modèle de vitesse de propagation des ondes sismiques. Nous exposons pour commencer, les hypothèses et les approximations faites pour permettre la définition de la loi de modélisation physique. Puis nous illustrons le développement de l'équation eikonale dont la résolution numérique permet le calcul des temps de première arrivée pour un modèle de vitesse donné.

### 2.1.1 Hypothèses et approximations

Nous nous intéressons dans ce travail à la propagation *asymptotique haute fréquence* d'ondes sismiques dans un milieu élastique *isotrope non homogène*. L'approximation asymptotique haute fréquence, aussi appelée approximation des rais, considère que la longueur d'onde des signaux est très faible devant les longueurs caractéristiques du milieu de propagation. Le champ d'onde peut alors être découpé entre un champ d'onde scalaire (ondes P) et un champ d'onde vectoriel (ondes S) qui se propagent indépendamment. Leur temps de trajet et leur amplitude sont alors définis à partir d'une solution particulière de l'équation de l'élastodynamique, à savoir respectivement l'équation eikonale et l'équation de transport [Červený, 2001], [Chapman, 2004].

### 2.1.2 L'équation eikonale

Nous choisissons d'illustrer le calcul de l'équation eikonale dans le cas acoustique [Baina, 1998]. La propagation d'une onde élastique P dans un milieu solide est alors assimilée à la propagation d'une onde de pression dans un fluide.

#### De l'équation des ondes à l'équation eikonale

Nous développons ici les calculs permettant d'obtenir l'équation eikonale à partir de l'équation d'onde scalaire en temps hors de la zone source [Červený, 2001]

$$K(\mathbf{x}) \nabla \cdot \left[ \frac{1}{\rho(\mathbf{x})} \nabla P(\mathbf{x}, \tau) \right] = \frac{\partial^2}{\partial \tau^2} P(\mathbf{x}, \tau), \quad (2.2)$$

où  $P(\mathbf{x}, \tau)$  est la pression acoustique subie par la particule au point courant  $\mathbf{x}$  à l'instant  $\tau$ ,  $K(\mathbf{x})$  est le module d'incompressibilité et  $\rho(\mathbf{x})$  est la masse volumique ou densité. Après transformation de Fourier sur la variable temporelle  $\tau$ , on obtient dans le domaine fréquentiel l'équation d'Helmholtz

$$K(\mathbf{x}) \nabla \cdot \left[ \frac{1}{\rho(\mathbf{x})} \nabla P(\mathbf{x}, \omega) \right] + \omega^2 P(\mathbf{x}, \omega) = 0. \quad (2.3)$$

Si on adopte pour solution générale une solution analogue à celle obtenue dans un milieu homogène [Baina, 1998], on a

$$P(\mathbf{x}, \omega) = A(\mathbf{x}, \omega) e^{i\omega t(\mathbf{x})}, \quad (2.4)$$

où  $A(\mathbf{x}, \omega)$  et  $t(\mathbf{x})$  sont respectivement l'amplitude et le temps de parcours de l'onde sismique. Cette solution une fois injectée dans l'équation (2.3), fournit après développement

$$|\nabla t(\mathbf{x})|^2 = \frac{1}{v^2(\mathbf{x})} + \omega^{-2} \frac{\nabla^2 A(\mathbf{x}, \omega)}{A(\mathbf{x}, \omega)} - \omega^{-2} \nabla(\ln \rho(\mathbf{x})) \cdot \frac{\nabla A(\mathbf{x}, \omega)}{A(\mathbf{x}, \omega)} + \frac{i\omega^{-1}}{\rho(\mathbf{x}, \omega) A^2(\mathbf{x}, \omega)} [\nabla(\rho(\mathbf{x}) A^2(\mathbf{x}, \omega) \nabla t(\mathbf{x}))], \quad (2.5)$$

où  $v(\mathbf{x}) = \sqrt{\frac{K}{\rho}(\mathbf{x})}$  est la vitesse de propagation des ondes de compression au point  $\mathbf{x}$ . En découplant la partie réelle de la partie imaginaire, on obtient l'équation de transport

$$\nabla \cdot (\rho(\mathbf{x}) A^2(\mathbf{x}, \omega) \nabla t(\mathbf{x})) = 0, \quad (2.6)$$

et l'équation dite hyper-eikonale [Menahem and Beydoun, 1985]

$$|\nabla t(\mathbf{x})|^2 = \frac{1}{v^2(\mathbf{x})} \left[ 1 + \left( \frac{\omega_c}{\omega} \right)^2 \right], \quad (2.7)$$

où l'on a utilisé la notation suivante :

$$\omega_c^2 = \frac{\nabla^2 A(\mathbf{x}, \omega)}{A(\mathbf{x}, \omega)} - \nabla(\ln \rho(\mathbf{x})) \cdot \frac{\nabla A(\mathbf{x}, \omega)}{A(\mathbf{x}, \omega)}. \quad (2.8)$$

Le système d'équations couplées formé par (2.6) et (2.7) est équivalent à l'équation d'Helmholtz puisqu'il a été obtenu sans aucune approximation.

### Une interprétation physique

L'équation de transport (2.6) exprime la conservation de l'énergie dans un tube de rais. L'équation hyper-eikonale (2.7) donne l'expression de la vitesse de phase en fonction de la fréquence de l'onde. Elle peut être vue comme une équation eikonale dans laquelle la lenteur effective du milieu serait corrigée par un terme d'autant plus important que la fréquence du signal est basse et que la variation locale de courbure des fronts d'onde est forte. Cette correction est une sorte de filtre lissant localement le modèle de lenteur et les fronts d'onde produits par la même occasion. Il en résulte que, dans un modèle donné, les singularités des fronts d'onde sont d'autant plus rares, et prennent leur origine d'autant plus loin de la position de la source que la fréquence du signal propagé est basse.

### Application de l'approximation haute fréquence

L'approximation asymptotique de la théorie des rais est obtenue lorsque  $\frac{\omega_c}{\omega} \ll 1$ , c'est-à-dire lorsqu'on se place dans un régime haute fréquence  $\omega \rightarrow \infty$ . Dans ce cadre asymptotique l'équation hyper-eikonale se réduit à l'équation eikonale :

$$|\nabla t(\mathbf{x})|^2 = \frac{1}{v^2(\mathbf{x})}. \quad (2.9)$$

Les temps de trajet  $t(\mathbf{x})$  deviennent alors indépendants de l'amplitude et le système d'équations découplées eikonale - transport peut alors être résolu d'une manière séquentielle. Notons que dans ce travail qui traite de la tomographie des temps de première arrivée, nous n'exploitons pas l'information contenue dans l'équation de transport.

### 2.1.3 Une relation non linéaire

Mises à part certaines lois explicites et particulières, comme celle à gradient constant de vitesse ou celle à gradient constant de lenteur au carré, il n'existe pas de solution analytique à l'équation eikonale [Červený, 2001]. Dans le cas général, il est indispensable de recourir à des méthodes de résolution numérique. Ces méthodes ont pour objectif de fournir une solution à l'équation eikonale (2.9) qui correspond aux temps de première arrivée. La relation  $\mathbf{g}$  qui lie les temps de première arrivée au modèle de vitesse est une relation non-linéaire.

## 2.2 Le problème inverse

La résolution du problème inverse se propose de retrouver les paramètres physiques permettant d'expliquer au mieux les données observées. Tarantola et Valette [1982] posent les bases de la théorie stochastique de l'inversion. Le principe est de considérer le problème inverse en terme d'états de connaissance *a priori* et *a posteriori* sur les données et les modèles. Ce formalisme permet la prise en compte d'information *a priori* et des statistiques des incertitudes, expérimentales et théoriques, sur les modèles et les données. A partir d'hypothèses faites sur la nature de ces incertitudes il est possible de définir une fonction coût, dite des moindres carrés, qui mesure l'écart entre les données observées et celles théoriques obtenues pour un modèle donné. La minimisation de cette fonction coût permet alors de déterminer le modèle qui explique au mieux les données observées au sens des moindres carrés [Tarantola and Valette, 1982].

### 2.2.1 La fonction coût des moindres carrés

Nous présentons ici les hypothèses permettant de définir la fonction coût classiquement utilisée en tomographie des temps de première arrivée [Tarantola, 1987a]. Nous considérons comme *gaussiennes* les statistiques des incertitudes sur les modèles et les données. Cette hypothèse gaussienne est justifiée par le théorème central limite, mais elle ne peut pas toujours être vérifiée pour une expérience réelle, i.e. avec un nombre fini de réalisations des variables mises en jeu. Le modèle de vitesse  $\mathbf{m}$  le plus vraisemblablement à l'origine des données observées  $\mathbf{d}_{obs}$  se situe alors au minimum de la fonction coût  $\mathcal{C}$ , aussi appelée fonction des moindres carrés,

$$\mathcal{C}(\mathbf{m}) = \frac{1}{2} \left[ (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs})^t \mathbf{C}_{\mathcal{D}}^{-1} (\mathbf{g}(\mathbf{m}) - \mathbf{d}_{obs}) + (\mathbf{m} - \mathbf{m}_{prior})^t \mathbf{C}_{\mathcal{M}}^{-1} (\mathbf{m} - \mathbf{m}_{prior}) \right], \quad (2.10)$$

où  $\mathbf{C}_{\mathcal{D}}$  et  $\mathbf{C}_{\mathcal{M}}$  sont respectivement les matrices de covariance *a priori* sur l'espace des données  $\mathcal{D}$  et l'espace des modèles  $\mathcal{M}$ ,  $\mathbf{m}_{prior}$  est un modèle connu *a priori* et  $^t$  désigne l'opérateur transposée. Les méthodes d'inversion stochastiques cherchent donc à minimiser la fonction coût  $\mathcal{C}$  par rapport au modèle  $\mathbf{m}$ .

### 2.2.2 La méthode de Gauss-Newton

L'approche choisie pour minimiser la fonction coût  $\mathcal{C}$  est l'utilisation de méthodes itératives locales. Ces approches locales convergent par un cheminement dans l'espace des modèles. Pour la méthode de Gauss-Newton, ce cheminement est basé sur l'utilisation du gradient de la fonction

coût, qui est lié à la direction de descente, et du Hessien de la fonction coût qui lui en donne la courbure. Le schéma itératif, à l'itération  $n + 1$ , de la méthode de Gauss-Newton s'écrit sous la forme [Tarantola, 1987a]

$$\mathbf{m}_{n+1} = \mathbf{m}_n - \mathbf{H}_n^{-1} \boldsymbol{\gamma}_n, \quad (2.11)$$

où  $\boldsymbol{\gamma}_n$  représente le gradient,

$$\boldsymbol{\gamma}_n = \left( \frac{\partial \mathcal{C}}{\partial \mathbf{m}} \right)_{\mathbf{m}_n}, \quad (2.12)$$

et  $\mathbf{H}_n$  est le Hessien de la fonction coût  $\mathcal{C}$  au point  $\mathbf{m}_n$

$$\mathbf{H}_n = \left( \frac{\partial^2 \mathcal{C}}{\partial \mathbf{m}^2} \right)_{\mathbf{m}_n}. \quad (2.13)$$

Schématiquement cela revient à faire un pas dans la direction donnée par le gradient à partir du modèle courant  $\mathbf{m}_n$  jusqu'au minimum du paraboloïde défini par le Hessien tangent à la fonction coût en ce point (Fig. 2.2). La méthode de Gauss-Newton a pour propriété de converger en une itération lorsque la fonction coût à minimiser est quadratique. On choisit alors classiquement de linéariser localement le problème direct (2.1) autour du modèle courant  $\mathbf{m}_n$ , ce qui permet de rendre la fonction coût  $\mathcal{C}$  à minimiser quadratique [Baina, 1998]. Ainsi à chaque itération, la méthode de Gauss-Newton fournit la solution exacte au problème localement linéarisé.

### 2.2.3 La solution des moindres carrés

#### La linéarisation du problème direct

La linéarisation du problème direct autour du modèle courant  $\mathbf{m}_n$  s'écrit

$$\mathbf{g}(\mathbf{m}) \simeq \mathbf{g}(\mathbf{m}_n) + \mathbf{G}_n (\mathbf{m} - \mathbf{m}_n) = \mathbf{g}(\mathbf{m}_n) + \mathbf{G}_n \delta \mathbf{m}_n, \quad (2.14)$$

où l'opérateur linéaire  $\mathbf{G}_n$ , aussi appelé matrice des dérivées de Fréchet, représente la dérivée de  $\mathbf{g}$  au point  $\mathbf{m} = \mathbf{m}_n$

$$\mathbf{G}_n = \left( \frac{\partial \mathbf{g}}{\partial \mathbf{m}} \right)_{\mathbf{m}_n}, \quad (2.15)$$

et

$$\delta \mathbf{m}_n = \mathbf{m} - \mathbf{m}_n. \quad (2.16)$$

Le symbole  $\simeq$  dans l'équation (2.14) signifie que les termes de second ordre peuvent être négligés par rapport aux erreurs d'observation et de modélisation, i.e. comparés aux valeurs d'écart-type et de corrélation présentes dans  $\mathbf{C}_D$ . La fonction coût  $\mathcal{C}$  donnée par (2.10) s'écrit alors

$$\begin{aligned} \mathcal{C}(\mathbf{m}) &= \mathcal{C}(\mathbf{m}_n + \delta \mathbf{m}_n) \\ &= \frac{1}{2} \left[ (\mathbf{G}_n \delta \mathbf{m}_n - \delta \mathbf{d}_n)^t \mathbf{C}_D^{-1} (\mathbf{G}_n \delta \mathbf{m}_n - \delta \mathbf{d}_n) \right. \\ &\quad \left. + (\delta \mathbf{m}_n - \Delta \mathbf{m}_n)^t \mathbf{C}_M^{-1} (\delta \mathbf{m}_n - \Delta \mathbf{m}_n) \right], \end{aligned} \quad (2.17)$$

où l'on a utilisé les notations suivantes :

$$\delta \mathbf{d}_n = \mathbf{d}_{obs} - \mathbf{g}(\mathbf{m}_n), \quad (2.18)$$

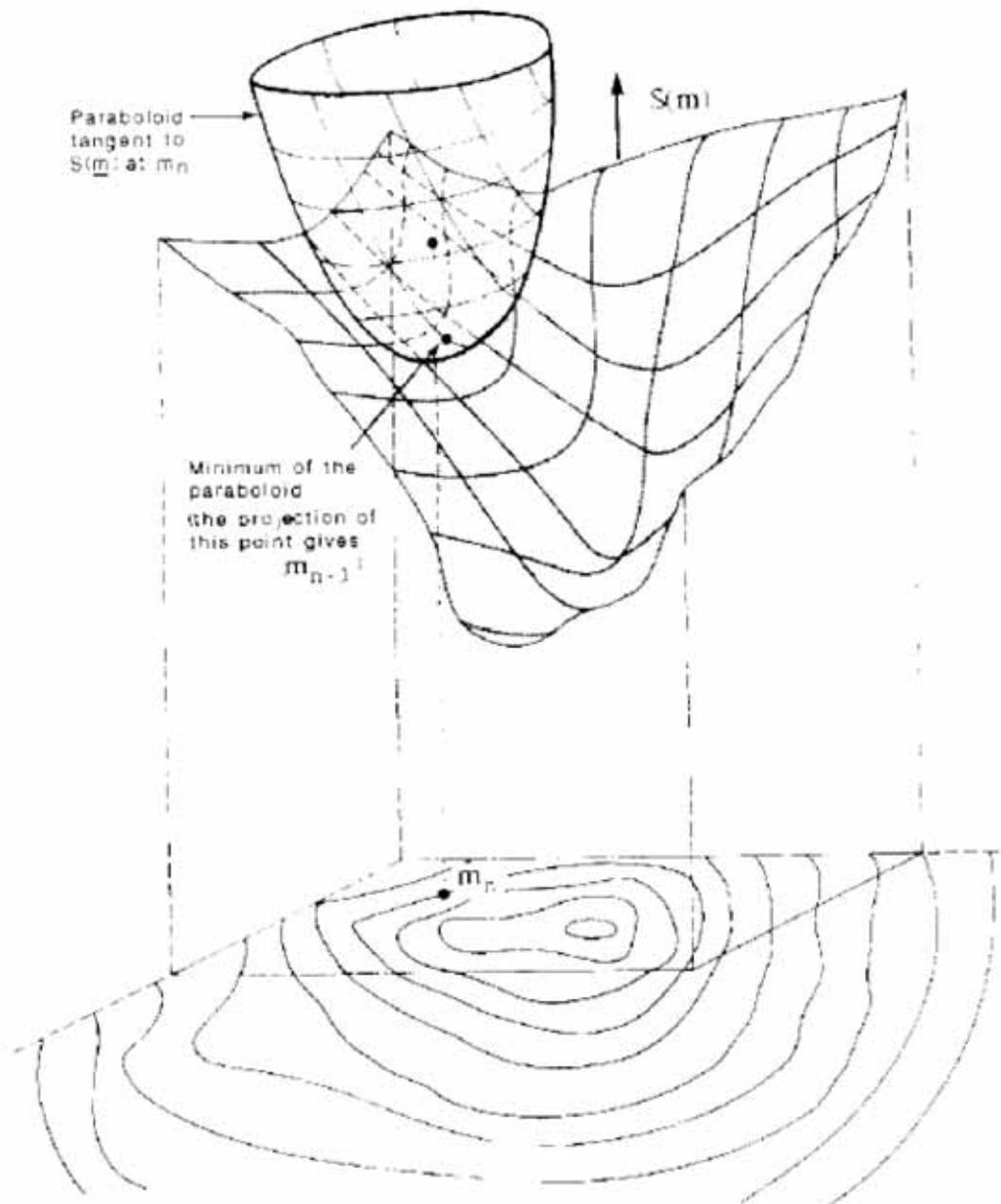


Fig. 2.2 – Minimisation d'une fonction coût par une méthode de Gauss-Newton. Cette figure illustre, pour un problème à deux dimensions, les courbes de niveau et la surface représentant la fonction coût. Le parabolôide tangent à la fonction coût au point courant est défini en utilisant l'information de courbure. Le point mis à jour correspond à la projection du minimum du parabolôide, extrait de [Tarantola, 1987a].

et

$$\Delta \mathbf{m}_n = \mathbf{m}_{prior} - \mathbf{m}_n, \quad (2.19)$$

La fonction coût  $\mathcal{C}$  est alors quadratique en  $\delta \mathbf{m}_n$ .

### La solution des moindres carrés

On peut alors exprimer directement la solution des moindres carrés  $\delta \mathbf{m}_n$  associée au problème linéarisé localement de minimisation de la fonction coût  $\mathcal{C}$  [Tarantola, 1987a]

$$\delta \mathbf{m}_n = (\mathbf{G}_n^t \mathbf{C}_D^{-1} \mathbf{G}_n + \mathbf{C}_M^{-1})^{-1} (\mathbf{G}_n^t \mathbf{C}_D^{-1} \delta \mathbf{d}_n + \mathbf{C}_M^{-1} \Delta \mathbf{m}_n). \quad (2.20)$$

En choisissant  $\mathbf{m} = \mathbf{m}_{n+1}$ , on obtient alors

$$\mathbf{m}_{n+1} = \mathbf{m}_n + \delta \mathbf{m}_n. \quad (2.21)$$

Par identification avec (2.11), on peut en déduire l'expression du gradient

$$\gamma_n = - (\mathbf{G}_n^t \mathbf{C}_D^{-1} \delta \mathbf{d}_n + \mathbf{C}_M^{-1} \Delta \mathbf{m}_n), \quad (2.22)$$

et du Hessien de la fonction coût  $\mathcal{C}$  au point  $\mathbf{m}_n$

$$\mathbf{H}_n = \mathbf{G}_n^t \mathbf{C}_D^{-1} \mathbf{G}_n + \mathbf{C}_M^{-1}. \quad (2.23)$$

### Un schéma récapitulatif

Le schéma (Fig. 2.3) reprend les principaux calculs mathématiques réalisés pour une itération de la méthode de Gauss-Newton associée au problème localement linéarisé (2.21).

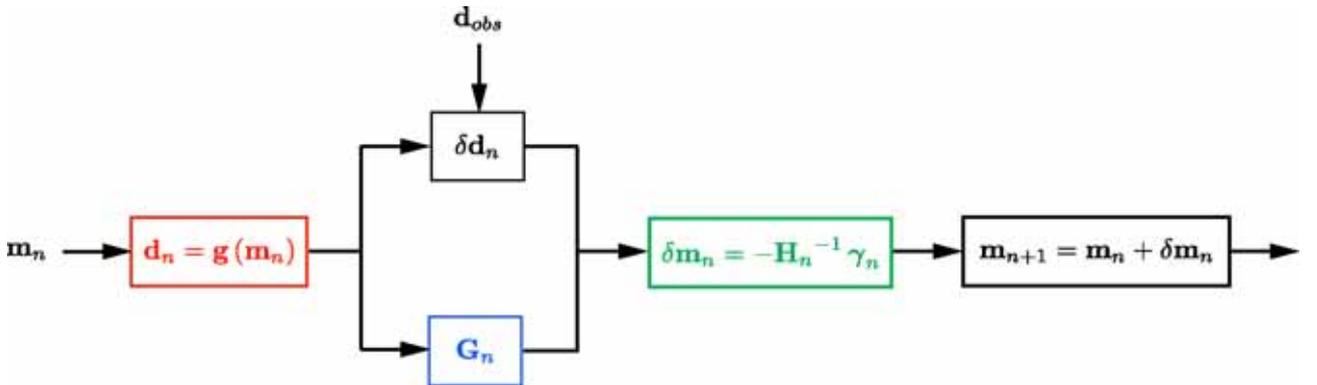


Fig. 2.3 – Diagramme représentant les principaux calculs mathématiques réalisés pour une itération de la méthode de Gauss-Newton associée au problème localement linéarisé. Les notations mathématiques sont définies à la partie 2.2.

Le processus itératif fournit alors un modèle  $\mathbf{m}$  le plus vraisemblablement à l'origine des données observées  $\mathbf{d}_{obs}$ .

### 2.2.4 Application à la tomographie des temps de première arrivée

On vient de voir les principaux développements mathématiques de la théorie de l'inversion et de la méthode de minimisation de Gauss-Newton. On va maintenant appliquer ces différentes formules au problème de la tomographie des temps de première arrivée. Le modèle recherché  $\mathbf{m}$  est donc un modèle de vitesse  $\mathbf{v}$ , ou indifféremment un modèle de lenteur  $\mathbf{s}$  liés par la relation

$$\mathbf{v} = \frac{1}{\mathbf{s}}. \quad (2.24)$$

### Le problème direct

Le schéma (2.4) est un exemple de discrétisation en grille cartésienne d'un modèle de lenteur  $\mathbf{s} = (s^j)_j$ . Nous avons aussi représenté sur ce schéma la trajectoire d'un rai entre une source et un récepteur. Un rai correspond, selon le principe de Fermat, à la trajectoire des temps extrema, ou d'un point de vue de la propagation des ondes, en milieu isotrope, à la trajectoire perpendiculaire aux fronts d'onde.

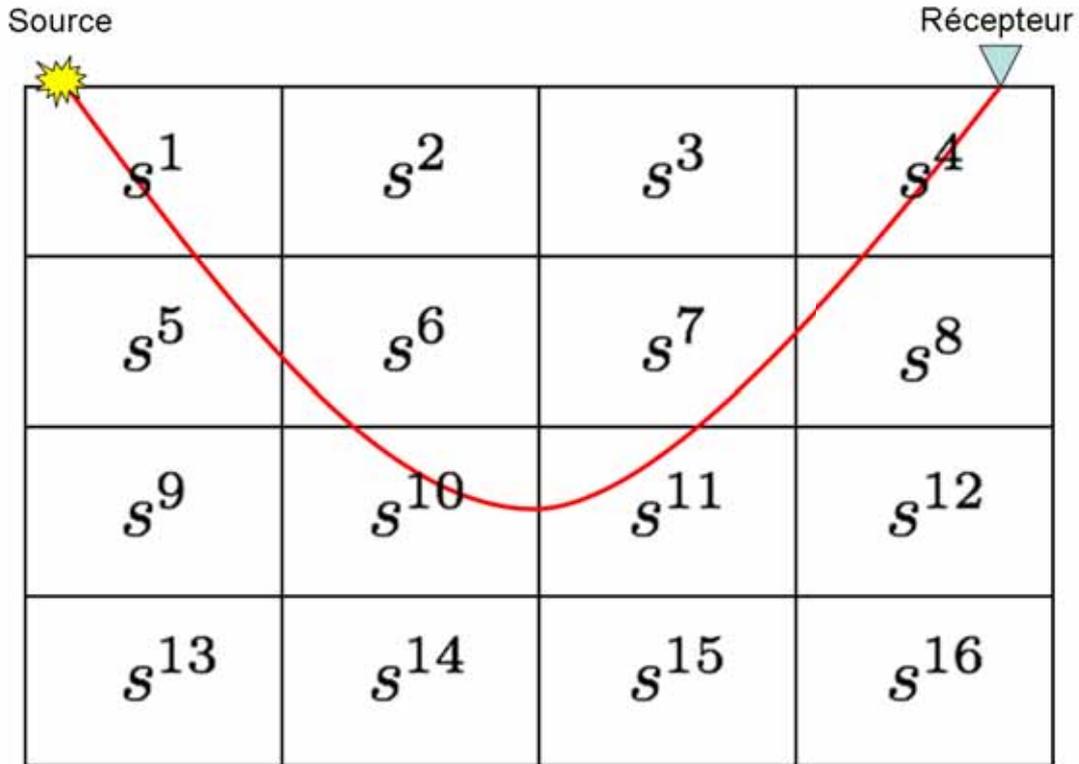


Fig. 2.4 – Schéma illustrant la discrétisation en grille cartésienne d'un modèle de lenteur  $\mathbf{s} = (s^j)_j$  et la trajectoire d'un rai entre une source et un récepteur.

Les données observées  $\mathbf{d}_{obs}$  correspondent aux temps de première arrivée pointés sur les sismogrammes  $\mathbf{t}_{obs}$ . Les données synthétiques  $\mathbf{d}$  sont les temps de première arrivée  $\mathbf{t} = (t^i)_i$  calculés pour un modèle  $\mathbf{s}$  donné par la résolution de l'équation eikonale (2.9). On peut alors établir la relation entre le temps de première arrivée  $t^i$ , le modèle de lenteur  $(s^j)_j$  et les longueurs des segments du rai  $l^{i,j}$  dans chaque maille du milieu discrétisé

$$t^i = \sum_j l^{i,j} s^j. \quad (2.25)$$

En utilisant une notation matricielle, on obtient la formule suivante :

$$\mathbf{t} = \mathbf{L} \mathbf{s}, \quad (2.26)$$

avec  $\mathbf{L} = (l^{i,j})_{i,j}$  la matrice des longueurs de rais. Cette équation correspond à la formulation du problème direct (2.1).

### Le problème inverse

Nous faisons tout d'abord quelques choix permettant de simplifier l'expression de la fonction coût à minimiser (2.10). Nous ne prenons en compte aucune information *a priori*, i.e.  $\mathbf{C}_{\mathcal{M}}^{-1} = \mathbf{0}$ . De plus nous supposons que les erreurs sur les données sont indépendantes et possèdent la même variance égale à 1, i.e.  $\mathbf{C}_{\mathcal{D}}^{-1} = \mathbf{I}$ . Sous ces hypothèses, la fonction coût à minimiser s'écrit

$$\mathcal{C}(\mathbf{s}) = \frac{1}{2} (\mathbf{L}\mathbf{s} - \mathbf{t}_{obs})^t (\mathbf{L}\mathbf{s} - \mathbf{t}_{obs}) . \quad (2.27)$$

Considérons maintenant un modèle de lenteur de référence  $\mathbf{s}_n$ , la matrice des longueurs de rais  $\mathbf{L}_n$  et les temps de première arrivée  $\mathbf{t}_n$  qui lui sont associés. Soient  $\mathbf{s} = \mathbf{s}_n + \delta\mathbf{s}_n$ , avec  $\delta\mathbf{s}_n$  une perturbation de lenteur, et  $\mathbf{t}$  le temps correspondant. Le principe de Fermat permet d'établir qu'une petite perturbation  $\delta\mathbf{s}_n$  de la loi de lenteur n'affecte les temps d'arrivée qu'au second ordre, c'est-à-dire que la trajectoire du rai n'est pas perturbée lorsque la variation de lenteur locale est faible, ce qui nous donne

$$\mathbf{t} \simeq \mathbf{L}_n \mathbf{s}_n + \mathbf{L}_n \delta\mathbf{s}_n = \mathbf{t}_n + \mathbf{L}_n \delta\mathbf{s}_n . \quad (2.28)$$

Cette étape correspond à la linéarisation du problème direct formulée à l'équation (2.14). On peut alors établir que  $\mathbf{L}_n = \mathbf{G}_n$ , c'est dire que la matrice des longueurs de rais correspond à la matrice des dérivées de Fréchet. En posant  $\delta\mathbf{t}_n = \mathbf{t}^{obs} - \mathbf{t}_n$ , on obtient

$$\delta\mathbf{t}_n = \mathbf{L}_n \delta\mathbf{s}_n . \quad (2.29)$$

La solution des moindres carrés associée au problème de minimisation de la fonction coût  $\mathcal{C}$  est donnée par (2.20)

$$\delta\mathbf{s}_n \simeq (\mathbf{L}_n^t \mathbf{L}_n)^{-1} \mathbf{L}_n^t \delta\mathbf{t}_n . \quad (2.30)$$

Nous utilisons ici le symbole  $\simeq$  car la matrice  $\mathbf{L}_n$  est en général mal conditionnée et donc l'inverse de  $\mathbf{L}_n^t \mathbf{L}_n$  est difficile à estimer. La solution donnée par (2.30) ne peut donc pas en général être calculée directement. On utilise alors des méthodes de résolution de système algébrique ou d'estimation de l'inverse généralisé correspondant associé au système linéaire (2.29) équivalent. Le gradient et le Hessien de la fonction coût  $\mathcal{C}$  sont alors

$$\boldsymbol{\gamma}_n = -\mathbf{L}_n^t \delta\mathbf{t}_n , \quad (2.31)$$

et

$$\mathbf{H}_n = \mathbf{L}_n^t \mathbf{L}_n . \quad (2.32)$$

Le schéma itératif donné par la méthode de Gauss-Newton (2.21) permet alors de déterminer un modèle de lenteur  $\mathbf{s}$  qui explique au mieux les temps de première arrivée observés à la surface  $\mathbf{t}_{obs}$ .

## 2.3 La mise en œuvre pratique

Nous venons de présenter la formulation mathématique conventionnelle des problèmes direct et inverse associés à la tomographie de temps de trajet de première arrivée. Nous nous intéressons maintenant à la mise en œuvre pratique d'un algorithme de tomographie des temps de première arrivée défini à partir de cette formulation. Nous présentons dans cette partie les méthodes classiquement utilisées pour le calcul des temps de première arrivée  $\mathbf{t}$ , pour la construction de la matrice des dérivées de Fréchet  $\mathbf{L}_n$  et pour la résolution du système linéaire qui permet la détermination de la perturbation de lenteur  $\delta\mathbf{s}_n$ .

### 2.3.1 La discrétisation du modèle de vitesse

Le choix de la discrétisation retenue pour décrire le modèle du vitesse dépend de l'application considérée et des algorithmes mis en œuvre. Dans le cadre de la tomographie des temps de première arrivée, le modèle de vitesse est classiquement discrétisé sur une grille cartésienne. Cette discrétisation permet la prise en compte de variations latérales de vitesse et elle ne nécessite aucune connaissance *a priori* du milieu, contrairement à la discrétisation par couches où un nombre donné de couches est considéré. De plus, elle est totalement adaptée aux algorithmes mis en œuvre pour le calcul des temps de première arrivée et pour la résolution du problème inverse.

### 2.3.2 Le calcul des temps de première arrivée

Le calcul des temps de première arrivée repose sur une résolution numérique de l'équation eikonale (2.9). L'approche la plus couramment utilisée est la résolution par différences finies introduite par Vidale [1988]. Les algorithmes ainsi définis calculent les temps de première arrivée sur des grilles cartésiennes, à partir d'un modèle de vitesse lui aussi défini sur une grille cartésienne avec le même espacement de grille.

L'implémentation d'un solveur eikonale peut être décomposée en deux problèmes :

- un problème local, qui a pour objectif de définir le schéma aux différences finies utilisé pour calculer le temps de première arrivée en un point donné de la grille en fonction des valeurs de temps des points voisins ;
- un problème global, qui permet de déterminer dans quel ordre les points de la grille doivent être considérés.

L'article de Vidale [1988] a rapidement été suivi par des dizaines d'articles introduisant des variantes de schémas aux différences finies développés pour améliorer la rapidité, la robustesse, la stabilité, ou la précision, par exemple [van Trier and Symes, 1991], [Podvin and Lecomte, 1991], [Qin et al., 1992], [Pica, 1997]. De nombreux articles continuent d'être publiés chaque année sur le sujet donnant ainsi naissance à des nouveaux noms de méthodes telles que la *Fast Marching Method* [Sethian, 1996], [Popovici and Sethian, 1998], ou la *Fast Sweeping Method* [Boué and Dupuis, 1999], [Zhao, 2005]. Des variantes de ces méthodes permettent aussi de prendre en compte par exemple l'anisotropie du milieu [Lecomte, 1993], [Qian et al., 2007] ou bien encore de réaliser une implémentation parallèle de l'algorithme [Zhao, 2006].

### 2.3.3 Le tracé de rais *a posteriori*

La résolution par différences finies de l'équation eikonale fournit une grille des temps de première arrivée en tout point du modèle. Il est alors ensuite possible de tracer les rais qui relient la source aux récepteurs de manière *a posteriori* [Podvin and Lecomte, 1991], [Baina, 1998]. Une méthode de calcul de la trajectoire des rais *a posteriori* repose sur le principe de stationnarité proposé par Vidale [1988]. On part du récepteur, pour une meilleure stabilité numérique, et on avance itérativement vers la source en suivant la direction opposée au gradient des temps. Ce gradient est estimé localement par un schéma aux différences finies centré sur un nœud du réseau ou bien sur le centre de la maille selon leur proximité de la position du point courant. Le tracé de rais *a posteriori* réalisé pour toutes les positions des couples {source - récepteur}

permet alors de calculer la matrice des dérivées de Fréchet  $\mathbf{L}_n$ . On calcule pour cela en chaque maille du modèle (Fig. 2.4) la longueur du segment du rai qui la traverse.

### 2.3.4 La résolution du système linéaire tomographique

La détermination de la perturbation de lenteur  $\delta \mathbf{s}_n$  passe par la résolution au sens des moindres carrés du système linéaire tomographique (2.29). Plusieurs méthodes de résolution de système algébrique ou d'estimation de l'inverse généralisé correspondant existent. On peut en distinguer deux classes : les méthodes de résolution directe comme la factorisation de Cholesky, la décomposition LU ou la décomposition par valeurs singulières et les méthodes de résolution itérative comme les méthodes de projection et de reconstruction. Les caractéristiques de la matrice des dérivées de Fréchet  $\mathbf{L}_n$  sont sa grande taille [*nombre de données*  $\times$  *nombre de mailles*], le fait qu'elle soit creuse, c'est-à-dire qu'elle contient essentiellement des éléments nuls, et son mauvais conditionnement, du à la couverture inhomogène des rais. Ces caractéristiques font que les méthodes itératives sont les mieux adaptées à la résolution du système tomographique [van der Sluis and van der Vorst, 1987], [Baina, 1998].

#### Les méthodes de reconstruction

Les méthodes de reconstruction ont pour principe général de rétropropager à chaque itération les résidus des temps sur les mailles traversées par les rais calculés pour un modèle donné. Cette rétropropagation est pondérée par des coefficients liés à la couverture des rais dans chaque maille. La méthode de reconstruction la plus répandue est la méthode *Simultaneous Iterative Reconstruction Technique* (SIRT). van der Sluis et van der Vorst [1987] ont prouvé théoriquement la convergence de cette méthode vers une solution au sens des moindres carrés pondérés. A partir de cette formulation commune, de nombreuses variantes ont été développées pour améliorer la rapidité de convergence [Zelt and Barton, 1998] ou bien prendre en compte les zones de Fresnel [Watanabe et al., 1999], [Grandjean and Sage, 2004]. L'avantage principal des méthodes de reconstruction est leur simplicité de mise en œuvre. Par contre, ces méthodes souffrent d'une renormalisation intrinsèque qui modifie le système linéaire à résoudre et donc la solution obtenue [van der Sluis and van der Vorst, 1987].

#### Les méthodes de projection

La méthode de projection la plus utilisée est la méthode *Least Squares QR* (LSQR) développée par Paige et Saunders [1982]. Dans son principe, cette méthode ressemble à la fois à la méthode itérative du gradient conjugué et à l'inversion par décomposition par valeurs singulières, de façon à tirer profit des avantages des deux méthodes simultanément [Baina, 1998]. La méthode LSQR, en calculant à chaque itération des estimateurs de la valeur du conditionnement, permet de prévenir les instabilités numériques. Nolet [1985] a montré, lors d'une étude comparative, que l'algorithme LSQR est supérieur aux algorithmes de reconstruction, tant du point de vue de la vitesse de convergence que de la stabilité numérique. De nombreux algorithmes de tomographie des temps de première arrivée utilisent l'algorithme LSQR pour déterminer la perturbation de lenteur  $\delta \mathbf{s}_n$ , par exemple [Spakman and Nolet, 1988], [Le Meur, 1994], [Baina, 1998], [Zelt and Barton, 1998].

### 2.3.5 Schéma de l'algorithme de tomographie

Le diagramme (Fig. 2.5) reprend les principales étapes algorithmiques aboutissant à la détermination de la perturbation de lenteur  $\delta s_n$  pour une itération de la méthode de Gauss-Newton. A partir d'un modèle de lenteur courant  $s_n$ , la résolution du problème direct pour  $N$  points de tirs fournit les cartes des temps de première arrivée en tous points de la grille. Pour toutes les positions {source - récepteur}, un tracé de rais *a posteriori* permet la construction de la matrice des dérivées de Fréchet. Cette matrice et les résidus calculés à partir des temps observés à la surface sont ensuite utilisés pour la résolution itérative du système linéaire tomographique. Cette résolution permet de déterminer la perturbation de lenteur  $\delta s_n$  à appliquer au modèle courant.

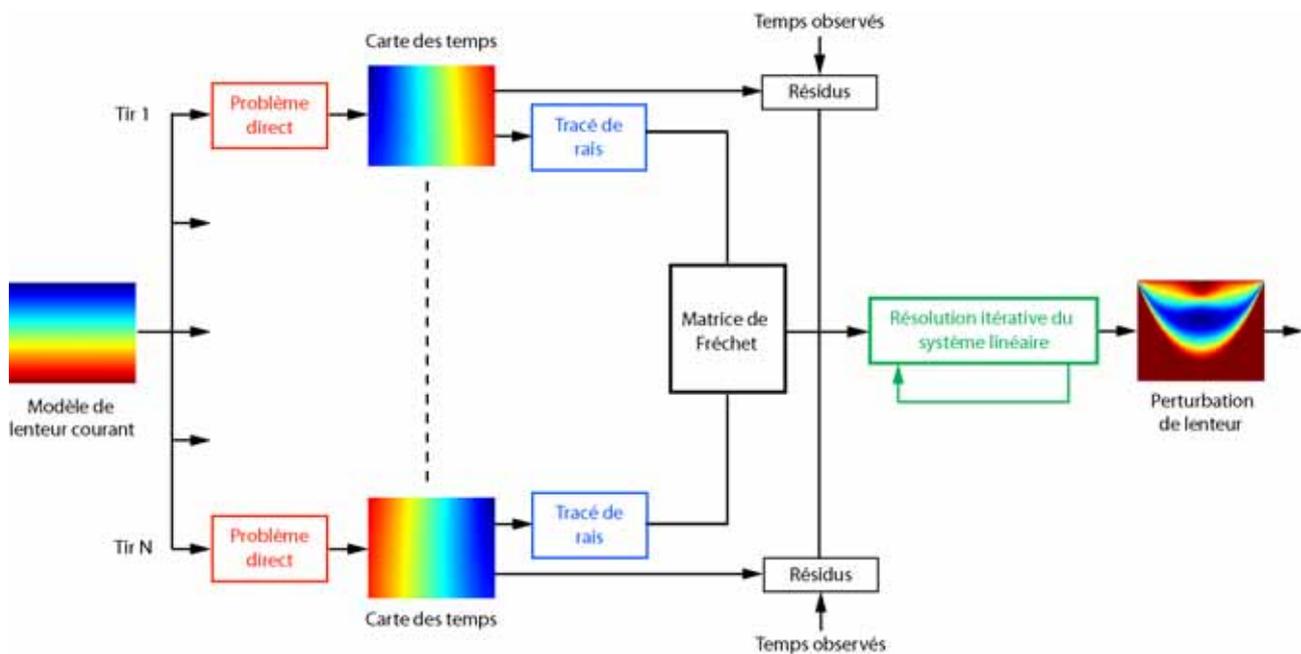


Fig. 2.5 – Diagramme illustrant les principales étapes de l'algorithme de tomographie de temps de trajet de première arrivée défini à partir de la méthode de Gauss-Newton.

Le schéma itératif donné par la méthode de Gauss-Newton (2.21) permet alors de déterminer un modèle de lenteur  $s$  qui explique au mieux les temps de première arrivée observés à la surface  $t_{obs}$ .

## 2.4 Vers un changement de méthode mathématique

Nous venons de présenter la formulation et les principaux éléments d'un algorithme classiquement défini pour résoudre le problème de tomographie des temps de première arrivée. La formulation repose sur le choix de la méthode de Gauss-Newton pour la minimisation de la fonction coût des moindres carrés. Nous présentons dans cette partie le contexte, géophysique et informatique, de la tomographie des temps de première arrivée qui est en pleine évolution. Ce changement de contexte permet de révéler les limitations pratiques rencontrées par de nom-

breux algorithmes de tomographie actuels. Ces limitations pratiques sont essentiellement dues à l'estimation de l'inverse du Hessien nécessaire au schéma itératif de Gauss-Newton.

### 2.4.1 Un contexte qui évolue

#### Le contexte géophysique

La taille des dispositifs d'acquisition sismique et la dimension des volumes étudiés ne cessent de croître. Vesnaver [2008] donne quelques chiffres caractéristiques des acquisitions sismiques 3-D actuelles et futures. Les systèmes d'enregistrement avec plusieurs milliers de récepteurs sont devenus des standards, et des acquisitions géophysiques pétrolières mettant en œuvre jusqu'à 36 000 canaux sont maintenant réalisées aussi bien sur mer que sur terre. Pour une acquisition sismique 3-D de taille raisonnable, par exemple sur une surface de 200 km<sup>2</sup> avec une densité de 250 points de tir par km<sup>2</sup>, on obtient au total 50 000 points de tir. Si chaque point de tir est composé de 36 000 traces sismiques, 1.8 milliards de sismogrammes sont enregistrés. Si on considère uniquement les temps de première arrivée, l'algorithme de tomographie doit alors être capable de gérer 1.8 milliards de temps observés. De plus si on considère une profondeur étudiée de l'ordre de 5 km et une discrétisation du modèle de vitesse avec une grille cubique de 12.5 m, le nombre de mailles est d'environ 600 millions. Un algorithme de tomographie des temps de première arrivée doit donc être capable d'assimiler un nombre très important de données observées et de gérer des modèles de vitesse de grande dimension.

#### Le contexte informatique

Parallèlement à l'évolution du contexte géophysique, le contexte informatique connaît lui aussi une évolution très rapide. Les moyens et la puissance de calcul disponibles ne cessent de croître. Cette augmentation de la puissance de calcul est essentiellement due à l'architecture parallèle des supercalculateurs actuels. Ils combinent en effet un nombre élevé de cœurs, ou unités de calcul, qui possèdent chacun une quantité de mémoire partagée ou distribuée et sont reliés entre eux par des connexions très rapides. La réalisation de calculs parallèles consistent en l'exécution d'un traitement partitionné en tâches élémentaires réparties entre plusieurs cœurs opérant simultanément. Il est ainsi possible de réduire significativement le temps de calcul nécessaire à l'exécution de ce traitement. Un algorithme de tomographie des temps de première arrivée, pour tirer pleinement profit de la puissance de calcul d'une architecture parallèle, doit donc pouvoir être facilement partitionné en tâches élémentaires.

### 2.4.2 Les limites pratiques de la méthode de Gauss-Newton

Le choix de la méthode de Gauss-Newton a des conséquences sur la mise en œuvre pratique d'un algorithme de tomographie des temps de première arrivée. Nous décrivons ici les limites pratiques dues au choix de cette méthode. Ces limitations sont essentiellement dues à la taille et au mauvais conditionnement de la matrice des dérivées de Fréchet  $\mathbf{L}_n$ . Cette matrice permet d'obtenir une estimation de l'inverse du Hessien  $\mathbf{L}_n^t \mathbf{L}_n$  utilisé par le schéma itératif de la méthode de Gauss-Newton.

## Des contraintes informatiques

Une caractéristique de la matrice des dérivées de Fréchet est sa grande taille [*nombre de données*  $\times$  *nombre de mailles*], même si de par sa construction cette matrice est le plus souvent creuse. Les valeurs caractéristiques des acquisitions sismiques 3-D actuelles données à la partie (2.4), 1.8 milliards de temps pointés et 600 millions de mailles, permettent d'établir un ordre de grandeur de l'espace mémoire occupé par cette matrice. En considérant un codage des réels sur 4 octets, l'occupation mémoire des temps pointés est d'environ 7.2 Go. En supposant, qu'en moyenne, un rai est décrit sur 100 mailles, l'occupation mémoire des éléments non nuls de la matrice des dérivées de Fréchet est de l'ordre de 720 Go. Une telle quantité de données ne peut être que très difficilement maintenue en mémoire vive et nécessite donc un stockage sur disque. On peut noter que l'occupation mémoire dépend essentiellement du nombre de données observées. Pour résoudre le système linéaire tomographique, l'algorithme de tomographie doit alors effectuer un grand nombre d'accès disque qui pénalisent l'algorithme en temps de calcul. Pour limiter l'occupation mémoire, une solution consiste à décimer le nombre de données observées utilisées et/ou à réduire la discrétisation du modèle étudié. Mais cette solution peut provoquer des pertes d'information et/ou de résolution.

Par ailleurs, les méthodes itératives de résolution du système linéaire tomographique ne se prêtent guère à une parallélisation des calculs à réaliser. Ainsi la méthode LSQR est très difficilement parallélisable, pour un gain en temps très faible. Les méthodes de reconstruction possèdent *a priori* une formulation adaptée au calcul parallèle. Cependant, cette parallélisation des calculs se fait au niveau des mailles du modèle, ce qui nécessite de fait un temps de calcul très important quand la taille du modèle considéré est grande.

## Une paramétrisation complexe

La résolution numérique du système linéaire tomographique est une tâche délicate à cause du mauvais conditionnement, synonyme de valeurs propres faibles, de la matrice des dérivées de Fréchet  $\mathbf{L}_n$  qui rend l'estimation de l'inverse du Hessien  $\mathbf{L}_n^t \mathbf{L}_n$  instable. La couverture inhomogène des rais est à l'origine du mauvais conditionnement de la matrice des dérivées de Fréchet. Pour prévenir les instabilités, il est nécessaire de procéder à des régularisations numériques. Ces régularisations peuvent prendre différentes formes [Tikhonov and Arsenin, 1977], [Yao and Roberts, 2002], l'approche la plus commune consiste à introduire un facteur d'amortissement, aussi appelé *damping*, dans la fonction coût à minimiser [Baina, 1998]. Par ailleurs, une paramétrisation adaptée des méthodes de résolution est indispensable pour permettre une estimation correcte de la perturbation de lenteur. Cette paramétrisation peut s'avérer complexe, par exemple les "boutons de réglage" de la fonction LSQR ou la définition des poids à appliquer au schéma itératif des méthodes de reconstruction.

### 2.4.3 Le choix d'une autre méthode

Les limitations pratiques rencontrées par les algorithmes de tomographie de première arrivée définis à partir de la méthode de Gauss-Newton sont essentiellement dues à l'estimation de l'inverse du Hessien de la fonction coût. Nous avons alors choisi d'utiliser une méthode de gradient pour la minimisation de la fonction coût des moindres carrés associée au problème de tomographie. L'utilisation d'une méthode de gradient permet généralement une mise en œuvre et une

résolution plus simples des problèmes inverses de grande taille<sup>1</sup>. Par exemple en géophysique, la résolution du problème d'inversion de forme d'onde [Lailly, 1984], [Tarantola, 1984], repose sur une méthode de gradient associée à la méthode de l'état adjoint [Lions, 1971] pour calculer le gradient de la fonction coût par rapport aux paramètres du modèle. Nous choisissons donc de formuler la résolution du problème de tomographie des temps de première arrivée à partir d'une méthode de gradient.

---

1. "The Newton method is not well adapted to large-sized inverse problems, which are more easily solved using gradient methods", d'après [Tarantola, 1987a].

