# NOTE SUR LA DÉMARCHE INVERSE

# POUR L'ESTIMATION DU LIEU ET DE L'INSTANT

# D'UNE EXPLOSION À PARTIR DE L'ENREGISTREMENT

# D'UN UNIQUE SISMOMÈTRE.

Présentée à M. Bernard Beauzamy, P.D.G.
Société de Calcul Mathématique S.A., Paris

par

M. Riadh Zorgati
Docteur en Sciences Physiques
Ingénieur-Chercheur
Directeur de Recherches Associé au CNRS

05 Décembre 2008

**Résumé**. Cette note constitue un éclairage sur les possibilités et les limitations de la résolution du problème dit inverse de la détermination du lieu et de l'instant d'une explosion à partir de l'enregistrement d'un unique sismomètre distant d'environ 4 km du lieu de l'explosion. Nous rappelons dans un premier temps les pré-requis nécessaires à la résolution d'un tel problème : élaboration d'un modèle de propagation des ondes sismiques générées par l'explosion dans un milieu géologique supposé parfaitement connu, enregistrement des signaux en qualité et en quantité suffisante, élaboration d'un modèle de l'ensemble des perturbations affectant le processus d'estimation des inconnues. Nous présentons ensuite la démarche de résolution d'un problème inverse. Nous montrons les limitations intrinsèques de résolution du problème posé dues d'une part à sa nature mathématique et d'autre part à la dépendance des résultats aux hypothèses de travail, au réglage de paramètres, aux erreurs liées aux techniques de calcul numérique de la solution, aux incertitudes et aux erreurs de mesure. En conclusion, la résolution du problème posé n'est pas possible en toute généralité. En disposant de plus de mesures, il serait possible d'estimer les inconnues du problème sans que l'on puisse raisonnablement évaluer l'incertitude sur le résultat.

**Mots-clés**. Problème inverse, modèle de propagation, ondes sismiques, incertitudes, erreurs, information a priori.

## 1. POSITION DU PROBLÈME

Une explosion a eu lieu. Un sismomètre distant d'environ 4 km a enregistré les signaux issus de cette explosion. On demande s'il est possible de reconstituer le lieu et l'instant de l'explosion à partir de cet unique enregistrement.

## 2. RÉSOLUTION DU PROBLÈME

### 2.1 Principe de résolution

Le sol, soumis aux forces engendrées par l'explosion, va subir d'une part des compressions et dilatations successives et d'autre part des cisaillements. Il va se déplacer et se déformer de façon transitoire et locale. Ce processus de déplacement local et de déformation élastique va se propager de proche en proche au milieu avoisinant par l'intermédiaire d'ondes sismiques, volumiques. La vitesse de propagation et l'amplitude de ces ondes dépendent de la nature et de la structure géologique du milieu dans lequel elles se propagent et de la nature de la déformation subie par le sol. Ainsi, les compressions et dilatations successives engendrent des ondes de compression qui se déplacent dans la même direction que le déplacement du sol tandis que les cisaillements subis par le sol donnent lieu à des ondes dites de cisaillement dont le sens de propagation est perpendiculaire au sens du mouvement du sol. A noter que les ondes de compression sont plus rapides que les ondes de cisaillement. L'explosion va aussi engendrer des ondes surfaciques, car se propageant en surface, plus lentes que les ondes volumiques, susceptibles de provoquer des dégâts aux édifices, selon la force de l'explosion et sa distance à la surface, en raison en particulier des mouvements horizontaux générés par ces ondes.

Ces deux types d'ondes vont être enregistrées par un sismomètre sous la forme d'un signal temporel qui reflétera les caractéristiques liées à l'explosion, en particulier son début et sa localisation spatiale, celles liées aux milieux traversés et celles spécifiques au sismomètre. Il est donc clair que pour pouvoir estimer, à partir de l'enregistrement du sismomètre, l'instant de début de l'explosion et le lieu de cette explosion, il faudra pouvoir évaluer, dans ce signal, l'influence du milieu entre la zone de l'explosion et le sismomètre ainsi que l'erreur liée à l'instrument de mesure.

Le principe de résolution du problème posé consiste à estimer la vitesse de propagation des différentes ondes ayant donné lieu au signal enregistré par le sismomètre, en particulier la différence des temps d'arrivée des ondes de compression et de cisaillement, puis, à partir de ce paramètre, déterminer la distance entre la source de l'explosion et le lieu d'enregistrement. En pratique, les ondes de volume ont un comportement analogue aux ondes lumineuses de sorte qu'elles suivent les lois de l'optique : au passage d'un milieu à un autre, elles peuvent être réfléchies ou réfractées, selon les différences de caractéristiques entre les milieux et suivre ainsi des trajets d'autant plus complexes que la structure géologique traversée entre le lieu de l'explosion et le lieu d'enregistrement est complexe. Comme le temps de parcours des ondes dépend de ce trajet, donc de la structure du sol, l'estimation de la distance entre l'explosion et le sismomètre, sauf configuration géologique particulièrement favorable, est d'une grande difficulté.

## 2.2 Pré-requis

Pour être en mesure de résoudre le problème posé, il faut disposer d'un certain nombre d'éléments, les pré-requis du problème, qui peuvent chacun constituer un problème d'une grande difficulté.

Le premier pré-requis est de disposer d'un modèle mathématique de propagation des ondes générées par l'explosion dans un milieu supposé parfaitement connu. Bien sûr, lorsque le milieu est homogène, le modèle est plus simple mais, en pratique, la structure géologique d'un sol, même sur une distance de 4 km, peut être d'une grande complexité. On peut ainsi avoir affaire à un milieu stratifié à la manière d'un millefeuille ou d'une ardoise, ou encore à un milieu non seulement stratifié mais présentant des anticlinaux et des synclinaux, enfin à un milieu complètement hétérogène en structure et en nature (plusieurs types de roches). Le modèle de propagation des ondes sismiques de volume à l'intérieur de telles structures doit supposer une certaine homogénéité dans chaque couche et tenir compte, à la manière des rayons lumineux du soleil entrant dans un plan d'eau calme, de la réflexion subie par les ondes à chaque interface, de leur réfraction, compte tenu de la différence de caractéristiques physiques des couches, de l'atténuation de ces ondes, etc. Les physiciens ont développé de tels modèles dont les solutions ne sont qu'approchées en raison des imprécisions liées à la connaissance des phénomènes physiques, des techniques de résolution aussi bien mathématique que numériques (celles liées aux erreurs numériques introduites lors du calcul sur ordinateur de telles solutions). Un problème est dit direct si l'on cherche à évaluer par le calcul la réponse qui sera enregistrée par un sismomètre, en supposant parfaitement connu le milieu et ses paramètres.

Le second pré-requis est de disposer de la valeur des paramètres du modèle, comme par exemple la vitesse de propagation des différentes ondes dans les différents types de matériaux rencontrés dans le sous-sol. Or la détermination de la vitesse de propagation d'une onde de compression dans une roche de nature donnée (un granite, un schiste, …) nécessite l'estimation de paramètres mécaniques par le biais d'expérimentations nécessairement entachées d'erreurs de mesure et d'incertitudes et qui elle-même peut être pour partie fondée sur des modèles physico-mathématiques complexes.

Le troisième pré-requis est la nécessité d'enrichir le modèle initial par une modélisation du bruit affectant les mesures, ce bruit étant la résultante de nos connaissances incertaines des phénomènes physiques mis en jeu, des incertitudes sur les paramètres du problème, du comportement du système de mesure, de diverses perturbations comme la discrétisation numérique du problème initial en vue de son traitement sur ordinateur. Le bruit est considéré comme un processus aléatoire et sa modélisation est généralement effectuée dans un cadre statistique, à l'aide d'une loi de probabilité, dépendant de certains paramètres qui peuvent être déterminés empiriquement ou être estimés à partir d'observations. La loi du bruit permet d'exprimer la vraisemblance des mesures conditionnellement à une structure du sous-sol. Dans la pratique, la modélisation du bruit est d'une extrême complexité.

Le quatrième pré-requis, synthèse des trois précédents, est de disposer d'un modèle mathématique fiable et validé qui, à une structure de sol connue, fait correspondre des données qui sont suffisamment proches des mesures qui seraient générées par le sismomètre pour la structure de sol considérée.

Le cinquième pré-requis est de disposer de mesures fiables et en quantité suffisante par rapport au nombre d'inconnues du problème à résoudre.

### 2.3 Nature du problème, spécificités, difficultés de résolution

Par opposition au problème direct, estimer le lieu et l'instant de l'explosion à partir d'enregistrements d'un ou plusieurs sismomètres est un problème dit inverse. Or, si déjà la résolution d'un problème direct peut s'avérer d'une grande complexité mathématique en raison des difficultés propres aux phénomènes physiques décrits et aux procédés de calcul numérique, la résolution du problème inverse associé est encore plus difficile.

Notons qu'en raison des très nombreuses applications industrielles et des forts enjeux économiques et sociétaux associés (sûreté d'installations sensibles, maintenance, imagerie médicale, armes cachées dans des bagages, recherche pétrolière, …), la résolution des problèmes inverses constituent un axe de recherche très actif des mathématiques appliquées depuis les premiers travaux fondateurs dans les années 1970. Dans l'industrie, cette importance s'explique pour partie par le nombre de situations où il s'agit d'estimer les caractéristiques physiques d'une région inaccessible à l'observation directe (à la mesure), situation typique d'une démarche inverse, ce qui est le cas ici.

Une caractéristique fondamentale d'un problème inverse est son caractère dit « mal-posé » en ce sens qu'il n'est pas sûr qu'il existe une solution mathématique au problème considéré, que si par chance une telle solution existe, il peut en exister un grand nombre, et que si par grande chance une solution unique existe, il n'est pas garanti que cette solution soit stable lorsque les données du problème sont entachées d'incertitudes ou d'erreurs de mesure, même minimes. La solution reconstruite via un problème inverse mal-posé va alors varier de manière dramatique (solution instable) et de façon erratique par rapport aux petites erreurs sur les données du problème ou même par rapport au processus de discrétisation des équations du modèle direct. Les problèmes inverses sont donc en général très sensibles aux incertitudes, qui sont nombreuses et variées dans la physique expérimentale : la nature expérimentale des données peut ainsi conduire à une solution du problème inverse très différente de la solution exacte, voire dans certains cas à la non existence d'une solution lorsque les données disponibles sont incompatibles avec les paramètres du problème.

En résumé, à cause des imperfections du modèle, du caractère fini des mesures et des erreurs qui les affectent, estimer des paramètres physiques à partir de ces mesures est délicat. En effet, un modèle imparfait est susceptible de générer des paramètres impossibles tandis que certains résultats peuvent ne correspondre à aucun paramètre car entachés d'erreurs de mesure. La résolution du problème inverse est enfin rendue plus compliquée puisqu'à une faible variation des résultats peut correspondre une évolution dramatique des paramètres. Par opposition au problème bien posé -possible, déterminé et stable-, un problème physique, et particulièrement un problème inverse est généralement un problème mal posé, c'est-à-dire ne disposant pas pour toute mesure d'une solution unique dépendant continûment des données.

### 2.4 Méthodologie de résolution

La résolution d'un problème inverse mal posé consiste alors à "construire" un procédé permettant d'associer à toute mesure une et une seule solution de sorte qu'il n'y ait aucune

équivoque entre mesure et solution associée à cette mesure. Deux approches sont possibles pour aborder les problèmes de cette nature :

- Une approche fonctionnelle, d'ordre théorique. Elle est de nature mathématique et consiste à transformer le problème mal posé initial en un problème bien posé en modifiant les espaces décrivant les variables du problème et leur topologie ;
- Une approche par régulation, plus pratique, qui consiste à rechercher la solution d'un problème bien posé qui devra être « proche » de celle du problème initial, dépendre continûment des donnée et être suffisamment proche de la solution exacte (les données calculées doivent être proches des données observées) dans un contexte soit déterministe, soit stochastique.

La résolution d'un problème inverse, avec la seconde approche, est guidée par les deux règles suivantes :

- seul est retenu un résultat de mesure suffisamment proche d'un résultat calculé par le modèle de sorte que l'écart entre mesure et calcul puisse être interprété comme une erreur de mesure ;
- il existe une certaine quantité d'information a priori non formulée dans la définition de l'ensemble des paramètres qu'il est indispensable d'utiliser afin de définir et choisir une solution acceptable du problème, éventuellement une classe de solutions.

À ces deux règles, qui régissent les conditions d'existence et d'unicité d'une solution au problème inverse, s'ajoute l'étude du problème des erreurs. Assurer la condition de stabilité constitue le point le plus délicat de la résolution d'un problème mal posé. Un problème connexe à celui de la stabilité concerne la robustesse de la méthode d'inversion, c'est-à-dire la propriété de la solution d'être bien définie pour les résultats exacts, stable pour les résultats imprécis ou redondants et insensible aux résultats aberrants ou impossibles.

Les méthodes inverses doivent donc permettre d'extraire le maximum d'informations fournies par les données tout en contrôlant parfaitement le rôle de l'information a priori introduite au cours du processus d'inversion. Ce contrôle est assuré par la régularisation d'un problème mal posé.

La propagation des ondes sismiques à travers le milieu est accompagnée de pertes, les enregistrements sont limités et des perturbations de nature différente entachent les mesures. L'inversion à partir de ces mesures s'appuie, implicitement ou explicitement, sur des connaissances a priori venant de la physique du problème ou des expériences antérieures.

En pratique, la prise en compte formelle des erreurs et des incertitudes conduit à considérer toutes les variables du problème comme aléatoires. L'estimation bayésienne est une approche très utilisée qui permet de compléter de façon cohérente l'insuffisance des mesures par des connaissances a priori et des caractéristiques attendues de la solution, qui sont de nature probabiliste. Les connaissances a priori se traduisent par une loi probabiliste a priori sur la solution recherchée. Cette modélisation probabiliste reflète un état de connaissance et ne confère pas un caractère aléatoire à l'inconnue recherchée. Les informations a priori probabilistes que l'on voudra incorporer dans l'estimateur peuvent être de natures différentes. La solution est alors la fonction densité de probabilité associée aux inconnues recherchées et à la mesure.

Parler dans ces conditions de précisions au mètre ou au centième de seconde n'a pas grand sens en ce qui concerne des données expérimentales.

### 2.5 Mise en œuvre

En complément des difficultés générales de calcul sur ordinateur d'une solution (stabilité des algorithmes face aux problèmes d'approximations, d'erreurs d'arrondis, …), le calcul numérique de la solution du problème inverse comporte lui aussi son lot de difficultés et d'erreurs d'estimation qu'il faut avoir en tête lors de l'interprétation des résultats :

- Optimisation. La résolution numérique d'un problème inverse avec données réelles conduit à des problèmes d'optimisation mathématique assez complexes en raison de la taille des problèmes à traiter et de la nature mathématique de ce type de problèmes. La solution du problème variera en fonction du critère choisi pour résoudre le problème, en fonction de l'information a priori qui sera incorporée et de sa transcription mathématique, en fonction du choix des nombreux paramètres des méthodes, enfin en fonction de la pertinence des méthodes de calcul. Toute la pertinence de l'inversion résidera dans le choix d'une mesure d'optimalité fournissant des solutions qui satisferont les exigences de l'utilisateur et s'obtiendront avec des algorithmes réalistes et robustes. Or, il est notoire que bien des problèmes d'optimisation relèvent encore aujourd'hui de la recherche en mathématiques appliquées, même s'il existe des éléments de solutions. C'est le cas des fonctions non différentiables, non convexes. La grande taille des problèmes traités accroît encore les difficultés ;
- réglage du paramètre de régularisation du problème (appelé hyperparamètre de régularisation). Sauf situation simple, et très rare, où il existe une théorie efficace permettant de régler de façon adéquate ce paramètre, il s'agit en pratique d'une décision de l'utilisateur qui tente de faire un (bon) compromis entre la qualité et la pertinence des données expérimentales et l'apport de l'information a priori qu'il a lui-même incorporé dans la procédure de calcul.

### 2.6 Evaluation des performances

Il est enfin pertinent de se poser la question des performances du processus de calcul de la solution du problème inverse, c'est-à-dire quelles sont les qualités de l'estimateur de la solution en termes :

- d'optimalité, c'est-à-dire sa capacité à estimer de manière raisonnable la solution ;
- de stabilité, qui mesure sa capacité à résister au bruit ;
- de robustesse, qui reflète sa capacité à fournir une solution fiable ou du moins acceptable lorsque les hypothèses de départ ne sont pas entièrement satisfaites (le bruit n'est pas gaussien, le modèle d'observation n'est pas linéaire, etc.);
- de sensibilité, qui mesure les variations dans la solution entraînées par des changements dans les données et les paramètres. La facilité de réglage en est une conséquence. Ainsi, l'estimation assez précise d'une solution, à partir de données incomplètes, nécessite l'incorporation des informations *a priori* « fortes ». Cela rend l'estimation sensible mais en même temps diminue sa robustesse ;
- d'efficacité de l'algorithme qui reflète la complexité numérique nécessaire au calcul de la solution (un estimateur qui serait très performant mais qui nécessiterait des années de calcul en raison de sa complexité numérique n'aurait pas d'intérêt en pratique !).

## 3. CONCLUSIONS

Il est impossible de reconstruire la solution exacte du problème posé à partir des données réelles disponibles en raison :

- du bruit qui affecte inévitablement les mesures. Ce bruit étant par nature aléatoire, on peut espérer au mieux connaître ses statistiques ;
- du caractère nécessairement incomplet des mesures disponibles pour résoudre le problème (il faudrait au moins trois enregistrements pour espérer avoir une estimation du lieu de l'explosion). On dispose donc d'un nombre de données inférieur au nombre d'inconnues ;
- des erreurs du modèle de propagation des ondes dans un milieu qui peut être géologiquement complexe ;
- des erreurs liées au calcul sur ordinateur de la solution ;
- la résolution d'un problème inverse mal posé s'appuie nécessairement sur des hypothèses sur la solution, afin de compenser l'insuffisance des mesures.

A ces raisons factuelles se rajoute le comportement humain :

- souvent, des connaissances a priori sur l'inconnue peuvent être tirées de la nature physique du problème, du contexte actuel et des expériences précédentes. Lorsque de telles informations sont disponibles, leur intégration dans la procédure d'estimation permet de pallier les perturbations, l'insuffisance des mesures et d'« orienter » l'estimation de la solution du problème afin qu'elle soit plausible à l'utilisateur. Si bien des caractéristiques des structures sont factuelles et connues a priori (une vitesse de propagation est positive ou nulle par exemple), il se pose la question de la formalisation de ces connaissances et de la construction d'un estimateur qui en fait un usage adéquat dans le processus de calcul. En effet, ces informations a priori peuvent être formalisées dans des cadres analytique, probabiliste, logique, etc. et incorporées explicitement dans des estimateurs issus des concepts très différents et … laissées à l'appréciation humaine ;
- le réglage de certains paramètres peut avoir une influence notable sur les résultats obtenus. La détermination adéquate de la valeur de ces paramètres peut se révéler d'une complexité aussi grande que celle du problème inverse à résoudre. Aussi, dans la pratique, l'utilisateur procède à divers essais, nécessairement limités, ce qui introduit un caractère un peu arbitraire au réglage des paramètres en question, et *in fine*, aux résultats obtenus.