Le soustracteur de bruit : sens physique

Henri MERMOZ

Lorsqu'un bruit parasite A ressemble à (est corrélé avec) un autre bruit parasite B, on peut détruire partiellement A en lui opposant B, en proportion de cette ressemblance. La mesure de la corrélation permet de conditionner B en un nouveau bruit I, qui est, en quelque sorte, la meilleure image de A (à une translation près dans le temps) qu'on puisse extraire de B. Il en résulte que, retrancher de A cette image I, donne un bruit résiduel R = A – I, dont la puissance est minimale par rapport à celle de A. En même temps, le procédé « épuise » la ressemblance avec B, et le bruit résiduel R n'est plus corrélé du tout avec B ; il est donc inutile de songer à renouveler l'opération entre R et B.

Le « conditionnement » de B en I est, en fait, le filtrage linéaire de B par un filtre déduit de la corrélation (AB) et de la densité spectrale de B, l'une et l'autre mesurées au préalable. La génération du filtre et l'opération de filtrage peuvent se faire soit dans le domaine « temps » (réponse impulsionnelle et convolution) soit dans le domaine « fréquences » (gain complexe et multiplication) ; c'est une affaire de choix d'algorithme. Dans le domaine « fréquences » le gain complexe G du filtre est la fonction de la fréquence :

G = spectre croisé de A et B/densité spectrale de B et ce filtre est appliqué à B pour fournir I. On remarque que I ne dépend pas d'un facteur numérique arbitraire appliqué à B ; car G est *divisé* par ce facteur, ce qui ne change pas le résultat de « B filtré par G ». On peut donc normer en puissance le bruit B, au préalable, ce qui économise de la dynamique donc des bits de quantification.

Mais, de plus, on peut vérifier que I reste *invariant* par rapport à un *filtrage linéaire arbitraire* préalablement appliqué à B (pourvu qu'aucune composante spectrale porteuse d'énergie ne soit éliminée). Cette propriété remarquable permet, entre autres commodités, de « blanchir » préalablement le bruit B, à condition que l'opérateur utilisé soit un filtrage linéaire. Si le bruit B est ainsi blanchi et normé, le gain G se simplifie en :

G = spectre croisé de A et de B

De toutes façons, la stabilité de I en présence de tout filtrage linéaire de B, montre bien qu'il représente intrinsèquement le « degré de ressemblance » de A et B, un paramètre qui reste inchangé quand on remplace B par un quelconque de ses filtrages linéaires. On vérifie que la densité spectrale de I est nécessairement inférieure à toute fréquence, à celle de A, et que la soustraction (A – I) s'opère à la fois sur les valeurs instantanées et sur les densités spectrales, donc sur les puissances. Dans ce sens, c'est une soustraction un peu particulière. On remarque également que les niveaux de I et de R sont proportionnels à celui de A, toutes choses égales d'ailleurs.

L'opérateur global décrit ci-avant — avec ou sans filtrage préalable de B — est nommé « soustracteur de bruit » dans la littérature technique des années 70 et 80. Ce terme, un peu publicitaire, pourrait laisser croire à une « suppression » totale du bruit A dans tous les cas. Ce ne peut être là qu'un cas limite, improbable en pratique, où B se trouverait être précisément un filtrage linéaire de A (ou bien l'un et l'autre étant deux filtrages différents d'un même *bruit blanc*) et par conséquent en ressemblance *totale* avec A. On aurait alors une annulation « parfaite » ($R = 0$), en écartant lestement les erreurs d'estimation sur G.

Cet opérateur est intéressant lorsqu'on peut s'attendre à trouver, additionné au parasite A (mais non corrélé avec A et B) un *signal utile* faible. Le soustracteur permet de le faire apparaître inchangé mais additionné à R, c'est-à-dire dans un bruit minimisé, avec plus de chances de détection. Mais il est impératif que rien qui ressemble à ce signal ne vienne se mélanger à B. Sinon la mesure de la corrélation (AB) serait faussée, ainsi, par conséquent, que la génération de G et de I. Par essence ou par construction, le parasite B doit être une « référence bruit seul ». S'il en est bien ainsi, la génération de G est *indépendante de la présence ou de l'absence du signal* (non corrélé avec B) quelle que soit son amplitude. Ce signal est le même sur A et sur R ; le processus est donc bien linéaire par rapport au signal. En pratique le soustracteur n'est payant que s'il existe un degré significatif de corrélation entre A et B. Sinon le gain supputable peut être absorbé dans les erreurs d'estimation et les fluctuations résiduelles.

Lorsqu'on y regarde de près, le soustracteur n'est autre que le « traitement optimal d'antenne » dans le cas particulier de deux capteurs avec une distribution du signal : 1 sur le premier et zéro sur le second.

Cette distribution se trouve réalisée par essence dans certains cas. Par exemple, en acoustique, une sortie d'antenne A peut être polluée par le bruit d'une machine accessible (toutes précautions banales d'équilibrage et d'amortissement mécanique ayant été prises). Un capteur placé sur cette machine délivre un parasite B qu'on peut espérer corrélé avec A, tout en restant insensible aux signaux utiles reçus éventuellement par l'antenne. On pourra, alors, détruire partiellement A avec un soustracteur, et ceci, d'une façon indépendante de la présence ou de l'absence du signal.

Mais on peut aussi « construire » des « références bruit seul » à partir d'une antenne à N capteurs, si on connait parfaitement la *distribution spatiale* du signal sur les capteurs, par exemple, dans le cas le plus simple, la direction de l'onde plane idéale supposée transporter le signal (et, naturellement, la géométrie des capteurs eux-mêmes). On peut alors définir $(N - 1)$ combinaisons de capteurs, « spatialement orthogonales » au signal (cas simple : combinaisons présentant un zéro de directivité dans la direction du signal). Avec ces « références bruit seul », en principe indépendantes, on peut combattre le parasite sur la sortie d'une N^{me} combinaison qui, elle, est « spatialement adaptée » au signal (cas simple : combinaison présentant un maximum de directivité etc...). Le remplacement des N sorties de capteurs, par N combinaisons linéaires indépendantes (ici spécifiques, par le fait qu'une seule peut délivrer du signal) ne change rien à la performance ultime d'un traitement optimal d'antenne (c'est une propriété générale). C'est dire que, sans perte de généralité, on peut concevoir un tel traitement comme un « mécano » de soustracteurs emboités convenablement. C'est là surtout une satisfaction intellectuelle, mais elle fait apparaître le soustracteur comme une « brique » standard, un module universel de traitement d'antenne. De plus, dans un cas simple cité plus loin, l'application n'est pas hors de portée.

Deux sortes d'algorithmes peuvent réaliser un soustracteur. Dans le premier type qu'on appelera « direct », les opérateurs partiels décrits ci-avant sont réalisés « en ligne » dans leur ordre naturel. Dans le second type, dit « indirect » ou « bouclé », on utilise un algorithme de gradient qui corrige en permanence le filtre G et le filtrage de B, de façon à maintenir sur la sortie, l'une des deux propriétés caractéristiques et associées de R,

— puissance minimale par rapport à A
— non corrélation avec B

l'une et l'autre indépendantes de la présence ou de l'absence de signal.

L'article suivant « soustraction de bruit » passe en revue les algorithmes proposés, avec leurs avantages et leurs inconvénients. Lorsqu'on aborde les algorithmes bouclés, le langage spécifique de la « régulation de processus » tend à masquer parfois les propriétés physiques fondamentales du soustracteur. Il fallait les rappeler et tout algorithme se doit de les préserver.

On peut aussi rappeler que le soustracteur a eu des précurseurs, sous des noms beaucoup moins commerciaux : compensateur d'intercorrélation, doublet dissymétrique d'antenne etc. Certains, brevetés en leur temps, ont eu des applications opérationnelles. Mais leur structure était limitée à des bandes relatives de fréquence très étroites (moins de 10 %), telles que la bande opérationnelle d'un radar ou d'un « sonar actif » (émettant des signaux et recevant des échos).

Cette situation, en effet, simplifie d'une façon considérable l'architecture des opérateurs et on peut alors se passer des algorithmes de gradient. Si cette bande très étroite est soigneusement délimitée aux entrées du soustracteur, on peut considérer le bruit B comme suffisamment « blanc ». On le norme facilement en puissance, par un opérateur du type contrôle de gain. Mais, surtout, le gain du filtre G s'identifie au spectre croisé, et se résume, dans la bande, à un seul nombre complexe (amplitude et phase). Les parties réelles et imaginaires de ce nombre sont respectivement :

— la moyenne du produit $A \times B$
— la moyenne du produit $A \times$ (B déphasé de 90°),
l'une et l'autre moyenne indépendante de la présence ou de l'absence de signal.

L'image I se forme en pondérant B et (B déphasé de 90°), respectivement par les 2 moyennes précédentes, et en ajoutant.

On peut, de façon équivalente, faire intervenir les « composantes basses fréquences » de A et de B qu'on obtient en les démodulant

— par la fréquence centrale
— par la même fréquence déphasée de 90°.

On a un peu plus d'opérations mais on réduit la fréquence d'échantillonnage.

Cette simplicité fondamentale du soustracteur « à bande relative très étroite » permet de réaliser des « mécanos » du genre envisagé plus haut ; le traitement d'antenne, dans ces conditions particulières, est à portée de main avec d'honorables ancêtres.

Le soustracteur à bande relative large (supérieure à 10 %), même réduite à 1/2 ou 1/3 octave, est d'une toute autre difficulté, avec la dimension fréquence en plus. Des réalisations du type « direct » existent depuis longtemps (corrélofiltre) mais c'est aussi le domaine privilégié des algorithmes de gradient. Les applications potentielles sont nombreuses et variées : sonar passif (veille acoustique), veille électromagnétique, suppression d'échos sur les circuits de télécommunications, auscultation médicale du cœur du fœtus en « éliminant » le cœur de la mère etc. etc.

Faisons une dernière remarque sur le problème de la constante de temps dans les procédés examinés. Tous les systèmes adaptatifs reposent sur un compromis entre le temps de réaction à des variations externes (qu'on souhaite bref) et le temps de bonne appréciation des paramètres estimés (qu'on souhaite long, si les conditions le permettent). Il existe toujours, plus ou moins sous-jacente, une notion de « durée de pseudo-stationnarité » ; elle est bien nécessaire pour parler de puissance, de densité spectrale ou de spectre croisé, toutes quantités qui n'ont un sens absolu qu'au cas idéologique de stationnarité parfaite ; ce qui d'ailleurs rendrait inutile tout traitement adaptatif ! Il faut donc faire avec la physique. Par exemple dans le cas d'un procédé « direct » (corrélofiltre ou bande relative très étroite) la constante de temps de compromis est celle qui est affichée au programme pour la durée d'intégration, dans le calcul des valeurs successives de la fonction d'intercorrélation ou du spectre croisé, ou dans le calcul des produits moyens. Cette constante de temps est prise assez longue pour raffiner les estimations autant que le permet la pseudo-stationnarité des informations reçues (du moins l'idée qu'on s'en fait), et assez courte pour corriger aussi vite que possible les défauts de stationnarité. On voit bien par exemple, dans les systèmes à bande relative étroite, qu'on a un compromis honnête en intégrant sur 50 ou 100 fois l'inverse de la largeur de bande. C'est un ordre de grandeur pour lequel les théories de l'estimation sont indulgentes tant sont faibles les biais et fluctuations. Mais souhaitons que le temps de pseudo-stationnarité soit du même ordre ou plus long, sinon le raffinement de l'estimation n'aurait aucun sens. Quoi qu'il en soit, dans les procédés « directs » c'est la même constante de temps qui gouverne estimation et réaction.

C'est peut être un peu moins clair pour les systèmes bouclés. La puissance de sortie à minimiser est bien estimée avec une constante de temps affichée. Mais dans les systèmes régulés de l'aval vers l'amont la constante de temps de régulation est généralement d'autant plus courte que l'écart à réguler est plus grand ; ce qui peut être un avantage à l'égard des non-stationnarités intempestives. Mais que devient la notion de constante de temps globale ?

Peut être le choix d'une constante de temps n'est-il pas tout à fait de même nature en « direct » et en « bouclé », bien que le résultat de tous les systèmes soit le même en régime juste assez lentement variable pour justifier la procédure adaptative.

En conclusion — très provisoire — le problème du soustracteur est une des clés, peut être la principale, dans le domaine de l'antenne adaptative. C'est un sujet qui tient l'affiche depuis vingt cinq ans, chaque année un peu plus foisonnant. La synthèse réalisée dans cette édition est donc des plus nécessaires. Il est souhaitable qu'on en vienne rapidement à des réalisations concrètes, voire commerciales, qui sont l'ultime justification des études théoriques.