

# INTRODUCTION AUX FONDEMENTS, AUX TECHNOLOGIES ET AUX APPLICATIONS POTENTIELLES DE LA SYNTHÈSE DE CHAMPS ACOUSTIQUES POUR LA SPATIALISATION AUDIO SUR RÉSEAUX DE HAUT-PARLEURS

P.-A. Gauthier<sup>†</sup>, A. Berry<sup>†</sup>, W. Woszczyk<sup>\*</sup>  
philippe\_aubert\_gauthier@hotmail.com

<sup>†</sup> GAUS (Groupe d'Acoustique de l'Université de Sherbrooke), 2500 boul. de l'Université, Sherbrooke, Québec, Canada J1K 2R1

<sup>\*</sup> CIRMMT (Centre for Interdisciplinary Research in Music, Media, and Technology), McGill University, 555 Sherbrooke Street West, Montréal, Quebec, Canada, H3A 1E3

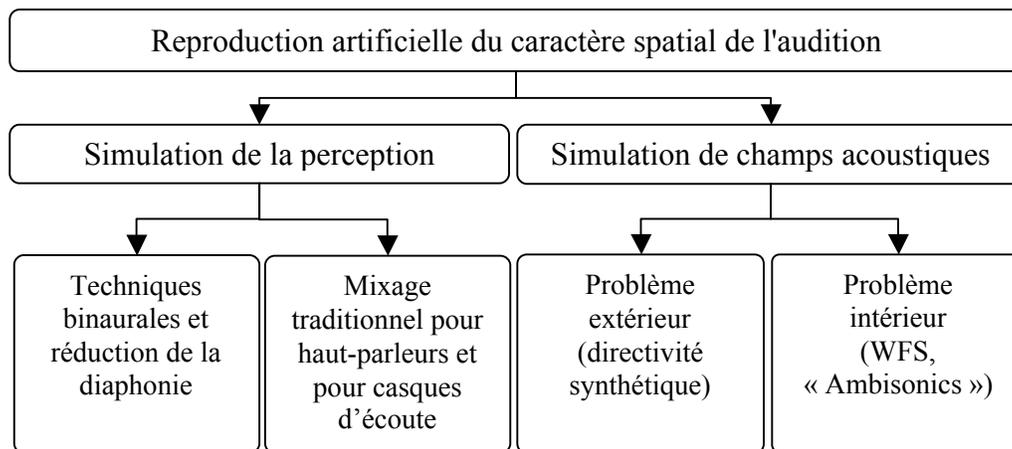
**Résumé.** Principalement dédié aux artistes ou aux compositeurs intéressés par la recherche et la création avec des réseaux de haut-parleurs qui peuvent inclure jusqu'à quelques centaines de sources, cet article introduit la technologie de reproduction du son qu'est la synthèse de champs acoustiques (WFS pour « Wave Field Synthesis »). Une classification des méthodes de reproduction spatiale du son est aussi présentée. Au sein des frontières de ce texte d'introduction, j'espère exposer une technique innovatrice de reproduction spatiale du son, bien connue en acoustique ou en traitement des signaux, qui souffre malheureusement d'un manque de visibilité chez les artistes ou les compositeurs. Telle que présentée dans cette revue, la WFS se pare de maintes possibilités et ouvertures pour des compositions spatiales. Les aspects les plus fascinants de cette technologie viennent du fait que le mixage spatial traditionnel (comme dans le cas du « surround 5.1 ») n'est plus réalisé par un ingénieur de son ou le compositeur. Parfois fastidieux et fragile, ce processus est maintenant automatisé – sur la base de principes physiques – et crée un lien entre la description virtuelle d'une scène acoustique, prescrite par le compositeur, et le réseau de haut-parleurs. Réseau qui constitue le système de reproduction. En bout de ligne, la technique en question procure la reconstruction du champ acoustique virtuel sur une zone étendue de l'espace de reproduction. Puisque la WFS est considérée comme une solution des plus prometteuses pour l'avenir, les intéressés par la diffusion et la composition spatiale, ou la reproduction sur réseau de haut-parleurs, pourront être informés par cette brève introduction au sujet.

## INTRODUCTION ET PROBLÉMATIQUE

Accaparé par des activités de recherches sur la présentation et la représentation multimédia, le protagoniste de telles recherches sera inévitablement confronté à la problématique, fort complexe, de la reproduction multisensorielle. Le problème est complexe en ce sens que les modalités sensorielles sont couplées : soumises à une certaine hiérarchie et toutes sujettes à une structure d'influence. Ce problème technologique de la reproduction multimédia peut, pour réduire le niveau de paramètres impliqués, être ramené à ses composantes sensorielles distinctes. Ainsi, selon cette relaxation du problème et en ce qui concerne l'objectif technologique de la reproduction spatiale du son, maints travaux ont été conduits par les chercheurs lors des dernières décennies.

À ce stade, la définition de la spatialisation audio à laquelle nous adhérons se doit d'être introduite. Cette définition ne se limite pas nécessairement à la reproduction parfaite de la sensation ressentie en une situation naturelle (reproduire la qualité sonore d'un lieu ou un événement) enregistrée par des moyens quelconques, elle prend plutôt sa base dans l'intérêt de reproduire la qualité de l'immersion naturelle (enveloppement, cohérence, réalisme). Bref, en spatialisation, on pourra dire que l'on tente de créer une sensation auditive qui tente d'atteindre la qualité de l'immersion naturelle.

Parmi les recherches effectuées sur le sujet, il est possible d'introduire une première dichotomie quant aux principales approches possibles. La première classe de méthodes est la *simulation de la perception* et la seconde, selon une école de pensée opposée qui répond à l'intérêt de reproduire un champ acoustique sur une zone plus ou moins étendue, est la *simulation de champs acoustiques*. Cette première distinction des deux approches générales est représentée à la figure 1 et peut aisément être illustrée par quelques exemples pratiques.



**Fig. 1 – Classification des approches fonctionnelles et des écoles de pensées en reproduction artificielle du caractère spatial de l'audition.**

Ainsi, l'ensemble de ce qui est couramment identifié comme des technologies binaurales appartient à la simulation de la perception. De telles technologies supposent, avec raison, que la seule reproduction de la pression acoustique aux tympanes de l'auditeur peut être efficace lors de la reproduction d'une impression spatiale auditive. La validité de cette hypothèse repose sur un fait maintenant bien connu et vérifié. On sait en effet qu'un filtrage fréquentiel opère sur les sons qui se rendent aux oreilles selon leur direction d'arrivée (par rapport à l'oreille externe qui inclut le buste, la tête et le pavillon de l'oreille). Ce filtrage, essentiellement attribuable à l'oreille externe, procure des indications binaurales qui permettent la localisation auditive des sources acoustiques puisque les signaux auditifs binauraux sont légèrement colorés selon la direction d'arrivée et selon la distance. Cette légère coloration permet de localiser inconsciemment les sources acoustiques. L'approche technologique binaurale (qui peut inclure : enregistrements binauraux avec mannequins, synthèse binaurale ou même filtrage binaural des prises de sons directes par les réponses impulsionnelles des oreilles externes)

de la reproduction spatiale du son appartient donc à la simulation de la perception puisque cette solution technologique inclut une partie de la chaîne de la perception auditive du spatial, qui est, dans ce cas, l'oreille externe. Les techniques de prises de sons et de stéréophonie conventionnelles se basent quant à elles sur l'observation, ou la connaissance plus ou moins empirique, de phénomènes de la perception auditive pour établir des règles de mixage manuel (images fantômes par le biais de la stéréophonie d'intensité ou de délai). Encore une fois, une part de la chaîne de la perception spatiale du son est prise en compte par le biais d'expériences et d'observations des perceptions résultant de certaines actions typiquement « audio ». Ce qui motive également l'appartenance des techniques de stéréophonies conventionnelles à la classe de la simulation de la perception.

Peu de systèmes se basent sur une approche dite de simulation de champs acoustiques puisque les technologiques et les concepts physiques impliqués sont habituellement plus complexes et requièrent un certain bagage en acoustique et en traitement de signal. En pratique, le nombre de sources de reproduction et le traitement de signal impliqués en simulation de champs sonores sont, tous deux, plus importants que ceux évoqués par une approche de simulation de la perception. Tel que mentionné rapidement plus tôt, l'hypothèse propre à la simulation de champs acoustiques est la suivante : en reproduisant un champ acoustique dans l'espace (ou dans une partie de l'espace plus ou moins grande) avec une fidèle distribution spatiale de la pression, le système auditif complet (ce qui inclut le filtrage naturelle par l'oreille externe) sera excité par un stimulus physique qui correspondra au stimulus virtuel, soit celui que l'on cherche à reproduire. Clairement, la tâche présente une nature plus physique et ne peut plus être atteinte par la seule compréhension théorique ou intuitive de la perception spatiale du sonore. Étant donné le nombre de sources et le traitement de signal nécessaires, on peut comprendre la popularité moindre de la simulation des champs acoustiques. La WFS, comme son appellation le suggère, est une méthode de simulation de champs acoustiques puisque le système tente de reproduire (objectivement ou, de façon équivoque, physiquement) un champ acoustique donné dans une zone entourée par des haut-parleurs. La création d'un motif de directivité synthétique autour d'une source, qui peut inclure plusieurs haut-parleurs, est aussi un exemple de technologie qui appartient à la simulation de champs acoustiques. Il importe de noter que ces deux problèmes présentent une complémentarité fondamentale. Le premier type (reproduction d'un champ acoustique à l'intérieur d'un réseau de sources, comme pour la WFS) est un problème intérieur et le deuxième type (reproduction d'une directivité synthétique autour d'une source complexe) est un problème que l'on pourrait qualifier d'extérieur. Cette nouvelle division, purement descriptive en apparence, a d'importantes répercussions sur les enjeux physiques et technologiques propres à ces deux objectifs complémentaires.

Dans cette hiérarchie, la place de la diffusion sur plusieurs haut-parleurs ne tient pas qu'à une seule case spécifique. L'appellation « multi haut-parleurs » est, quoique à propos et judicieusement rassembleuse en maintes situations, trop générique pour cela. Elle peut tout aussi bien désigner une diffusion sur « Surround 5.1 » par simulation de la perception que l'usage de sources à directivités synthétiques industriellement disposées sur une scène pour un concert de haut-parleurs. En forçant la note, la présentation

simultanée de matériel audionumérique, variant selon les canaux, sur une série de casques d'écoute, pourrait recevoir cette même appellation de diffusion sur système « multi haut-parleurs ». Ainsi, quoiqu'il soit difficile d'attribuer précisément des parties de l'organisation de la figure 1 à la présentation sur plusieurs haut-parleurs, cette expression fait écho à certaines pratiques musicales électroacoustiques et y trouve pleinement son utilité.

À l'exception de certaines combinaisons et hybridations possibles, la classification des approches telle que présentée à la figure 1 regroupe l'ensemble des technologies possibles, voire imaginables, en spatialisation audio. On note par contre qu'une telle taxinomie n'est pas nécessairement reconnue; ainsi, selon les points de vue, d'autres qualificatifs peuvent être utilisés pour édifier des classifications qui pourraient s'attarder aux résultats obtenus (système en deux dimensions, trois dimensions, distribution homogène ou hétérogène des directions reproductibles, etc.).

### **BUTS ET DÉFINITION PRATIQUE DE LA SYNTHÈSE DE CHAMPS ACOUSTIQUES, UN EXEMPLE DE SYSTÈME MULTI HAUT-PARLEURS**

La WFS est une application spécifique de la simulation de champs acoustiques. C'est d'ailleurs cette spécificité qui motive ici l'usage de son appellation anglophone (« Wave Field Synthesis ») plutôt que sa traduction française (synthèse de champs acoustiques) pour réduire la confusion entre WFS et simulation de champs acoustiques. À ce sujet, on pourrait simplement clarifier en soulignant le fait que la WFS est un hyponyme de simulation de champs acoustiques.

Avant même de dévoiler les détails de la WFS, supposons a priori qu'elle serait sollicitée pour reproduire une onde sphérique progressive qui émanerait d'une position prescrite. Cette simple définition, soit une source de type sphérique<sup>1</sup> en une telle position, serait donc la scène virtuelle. Cette source virtuelle devra être alimentée par un flux audionumérique. C'est la physique des ondes qui va permettre d'établir le noyau du système de WFS. Ce noyau est le traitement de signal numérique qui doit justement relier, automatiquement et en temps réel, un flux audionumérique monophonique à une importante série de haut-parleurs pour obtenir le champ acoustique attribuable à la source virtuelle désirée. Quoique ces remarques sont formulées pour une source virtuelle unique, une scène complexe est obtenue par la simple superposition de sources virtuelles.

C'est pour reproduire un champ acoustique sur une zone étendue, qui pourrait inclure une audience en mouvement, que la WFS fut initialement introduite. Dans sa formulation originale et selon son usage le plus simple, la WFS suppose que l'espace de reproduction (où se trouve le système de son et l'audience) est un espace anéchoïque, c'est-à-dire sans écho ni réflexion acoustique par les murs ou les objets sur place. Cette simplification, qui ne correspond évidemment pas à la réalité d'un studio, d'une salle de concert ou d'écoute, allège énormément, mais de façon nécessaire, le développement des bases théoriques et des technologies de WFS. Malgré l'importante approximation attribuable à

---

<sup>1</sup> On pourrait définir ce type d'onde par le biais d'une analogie plus ou moins exacte mais néanmoins illustrative : comme des ondes circulaires à la surface de l'eau quittant un point où un objet serait lancé.

la supposition que l'espace de reproduction est anéchoïque, cette hypothèse de champ libre (c'est-à-dire de lieu de reproduction sans obstacle ni réflexion) est justifiable de par une observation quant à la perception spatiale du son. Ce qui est baptisé l'effet de précédence en localisation auditive nous montre en fait que la construction spatiale d'une scène auditive<sup>2</sup> peut éradiquer l'influence de réflexions (la localisation opère selon la direction d'arrivée du premier front d'onde dans une série cohérente de fronts<sup>3</sup>), dans la mesure où ces dernières ne sont pas trop tardives [1]. Lors de la reproduction sur système WFS, c'est donc le système auditif qui compense et annule l'effet perceptif attribuable à des réflexions par le biais de cet effet de précédence. Mais en pratique, ces réflexions, qui demeurent bien réelles, peuvent nuire à la reproduction objective et subjective, notamment au sujet de la reproduction des réflexions acoustiques d'un lieu virtuel par WFS. Une part des recherches actuelles sur la WFS concerne cet aspect de la reproduction, c'est la compensation de salle.

En situation pratique, pour atteindre une telle reconstruction spatiale d'un champ virtuel à l'aide de la WFS, la zone d'écoute est entourée par un réseau linéaire de haut-parleurs (dans le plan horizontal pour la reproduction en deux dimensions) tous séparés par une petite distance. Pour une zone d'écoute couvrant 12 mètres carrés (un cercle d'environ 5.5 mètres de diamètre), on pourrait avoir près de 100 haut-parleurs qui formeraient une ceinture autour de l'audience. En pratique, cette distance de séparation est simplement limitée par la taille des transducteurs. Plus les sources seront rapprochées, plus la WFS sera spatialement efficace en hautes fréquences. Il doit y avoir en effet au moins deux sources par longueur d'onde d'intérêt. Dans l'exemple précédent des 100 sources, la WFS ne serait efficace qu'en dessous d'environ 1kHz. Cela, se démontrant avec quelques arguments de la perception du son qui vont au-delà des besoins de la présente introduction, est tout de même suffisant – du point de vue subjectif – en terme de qualité de reproduction.

Cette première section a passé en revue la partie tangible et visible de la WFS. La prochaine section présente les principes évoqués pour accomplir la synthèse.

## FONDEMENTS THÉORIQUES DE LA WFS

Tel que mentionné plus tôt, le fonctionnement de la WFS se base sur des notions physiques propres aux phénomènes ondulatoires pour atteindre le louable objectif de la reproduction spatiale du son. Soit le problème de la reproduction spatiale précisément défini comme suit pour la WFS (ou pour tout autre méthode de simulation de champs acoustiques) : l'objectif est de reproduire le détail (l'amplitude et la phase dans l'espace) d'un champ acoustique virtuel donné (il est habituellement défini virtuellement) sur un espace étendu.

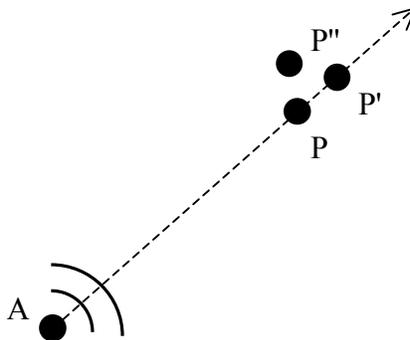
---

<sup>2</sup> La définition d'événement auditif correspond ici à celle introduite par Blauert [1]. L'événement auditif est la résultante sensorielle d'un événement acoustique.

<sup>3</sup> Une série cohérente de front d'ondes correspondrait par exemple à l'arrivée d'une première impulsion (directement issue de la source d'impulsion) et une série d'autres (décalées dans le temps et d'amplitude réduite) issues des réflexions de l'impulsion originale sur les murs et les objets. Cette suite de fronts formerait une série cohérente de front d'onde.

Voyons maintenant comment la WFS s'appuie sur des notions physiques pour réaliser la reconstruction d'un champ virtuel. Revisitons rapidement la notion physique du son. Le son se propage dans l'air comme des ondes, sonores dans ce cas. Ces ondes sont de petites oscillations du médium qui sont transmises de particules en particules jusqu'à ce que l'onde sonore rejoigne nos oreilles. Les ondes sonores se propagent en trois dimensions et non pas en deux dimensions comme les ondes à la surface de l'eau. Malgré le fait que les images présentées dans cet article sont essentiellement des sections en deux dimensions, il ne faut pas oublier cette réalité. Les ondes se propagent à la vitesse du son, ce qui n'est pas le cas de l'air vibrant en tant que tel. Les oscillations du milieu sont transmises de voisin en voisin sans qu'il n'y ait de transport de matière de la source acoustique à l'auditeur. Comme point de départ, c'est le principe de Huygens (ou le principe des sources secondaires) qui suggère que l'idée de reproduire un champ acoustique en entourant la zone d'écoute de sources est peut-être utile.

Le raisonnement de Huygens (qui ne date pas d'hier, mais bien de quelque part au 17<sup>e</sup> siècle [2]) ne concerne pas directement la reproduction sonore ou même la spatialisation, il stipule simplement qu'il est possible de recréer un champ ondulatoire en remplaçant un front d'onde à reproduire (originellement créé par une source primaire) par une suite continue de sources (les sources secondaires) sur le même front d'onde original. En agissant de la sorte, par le biais de cette expérience mentale, il serait possible de laisser tomber la source primaire puisque la contribution cumulative des sources secondaires recrée le front d'onde originellement attribuable à la source primaire. « Chaque particule (P) de la matière traversée par une onde communique son mouvement, non seulement à la particule voisine (P') qui est alignée avec (P) et la source [de l'onde] (A), mais également à toute autre particule (P'') qui touche (P) et qui s'oppose à son mouvement. De sorte qu'il faut qu'autour de chaque particule il se fasse une onde dont cette particule sera le centre [3]. » Telle était la proposition originale de Huygens qui, conjuguée avec les figures 2 et 3, se laisse plus facilement aborder. Dans la proposition de Huygens, les particules (P) correspondent aux sources secondaires et la source (A) est en fait la source primaire.



**Fig. 2 – L'illustration de l'énoncé de Huygens.**

Quoique le principe des sources secondaires laisse supposer que sa validité n'est vérifiable que lorsque les sources secondaires sont placées sur le front d'onde à

reproduire (voir la partie gauche de la figure 3), il est aussi possible de placer les sources secondaires sur une surface quelconque. Comme une ligne, telle que montrée à droite de la figure 3. Ce genre de configuration est évidemment plus pratique et flexible pour des applications technologiques réelles. En pratique, les sources secondaires continuellement distribuées sont remplacées par un nombre fini de sources réelles, des haut-parleurs électrodynamiques par exemple.

Sachant que la possibilité de remplacer une source réelle par un réseau de sources est valable, nous sommes en droit de nous demander comment, pratiquement, reproduire un enregistrement audionumérique en une telle application. Pour y répondre, il faudrait d'abord savoir comment alimenter les sources secondaires qui doivent remplacer la source primaire. On trouve l'essence de la réponse à cette problématique en utilisant la formulation mathématique du principe de Huygens. Le principe en question se manifeste en effet par le biais d'une formulation intégrale. C'est la fameuse formule de Kirchhoff-Helmholtz qui permet de trouver qu'elles doivent être les amplitudes et les phases des sources qui se trouveraient distribuées sur une surface de sources secondaires<sup>4</sup>. Connaissant ces amplitudes et ces phases en fonction de la fréquence pour la position d'une source virtuelle donnée et la position d'une source secondaire de reproduction donnée, ce couple amplitude/phase en fonction de la fréquence est ramené dans le domaine temporel pour finalement définir un filtre qui pourrait être réalisé par le biais du traitement des signaux numériques conventionnel pouvant être réalisé sur ordinateur. Un système complet de WFS pourrait ainsi inclure de 2 à une dizaine de PC normaux, selon la taille du réseau de haut-parleurs (jusqu'aux environs de 500 haut-parleurs). Entre chaque source virtuelle et chaque haut-parleur, un opérateur de synthèse doit être implanté tel qu'illustré à la figure 4 où  $WFS_i$  représente l'opérateur de synthèse (qui traite un flux audio issu d'un fichier définissant le son à reproduire en une position virtuelle prescrite) pour le haut-parleur  $i$ .

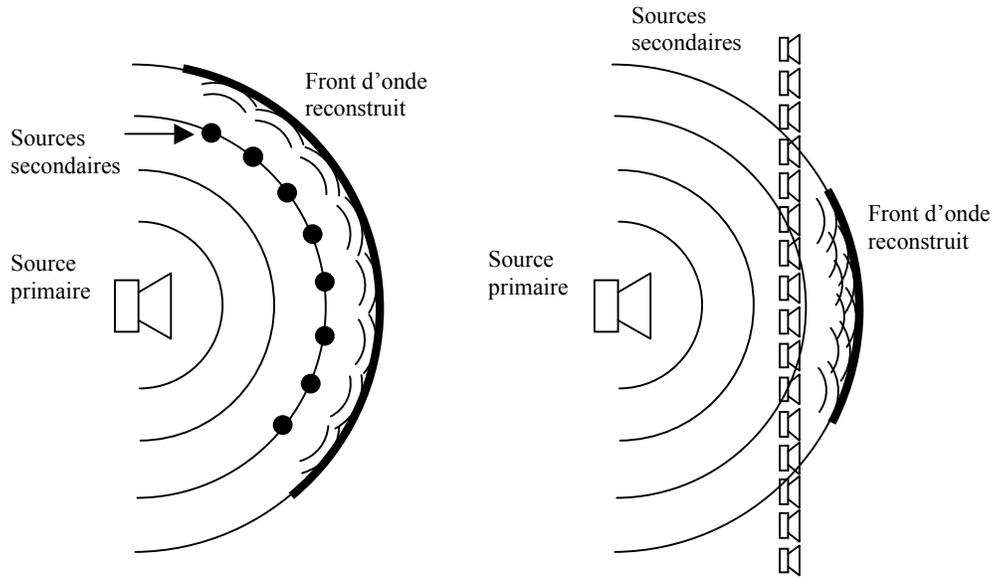
La WFS inclut quelques hypothèses supplémentaires dans ce développement bien simple en apparence. L'explication ici introduite est en fait bien sommaire, voire approximative, mais illustre judicieusement la physique derrière la WFS. Le lecteur intéressé par le détail du développement est invité à consulter les références.

Sans y accorder trop de détails, la WFS peut être mise en relation avec d'autres technologies. Historiquement, la WFS fut introduite par Berkhout vers la fin des années 80 et au début de la décennie suivante [4]. L'idée n'était pas nouvelle puisque le concept de l'holophonie avait été introduit par Jessel en 1973 [3]. La proposition de WFS par Berkhout est en fait une application spécifique de l'holophonie qui est un concept théorique que nous avons touché plus tôt lors de l'évocation de la formulation de Kirchhoff-Helmholtz. Qui plus est, il fut démontré vers la fin des années 90 que la reproduction dite « Ambisonic » et ses équations d'encodage et de décodage de base pouvaient être dérivées à partir du concept d'holophonie. Ce qui établit un lien notable

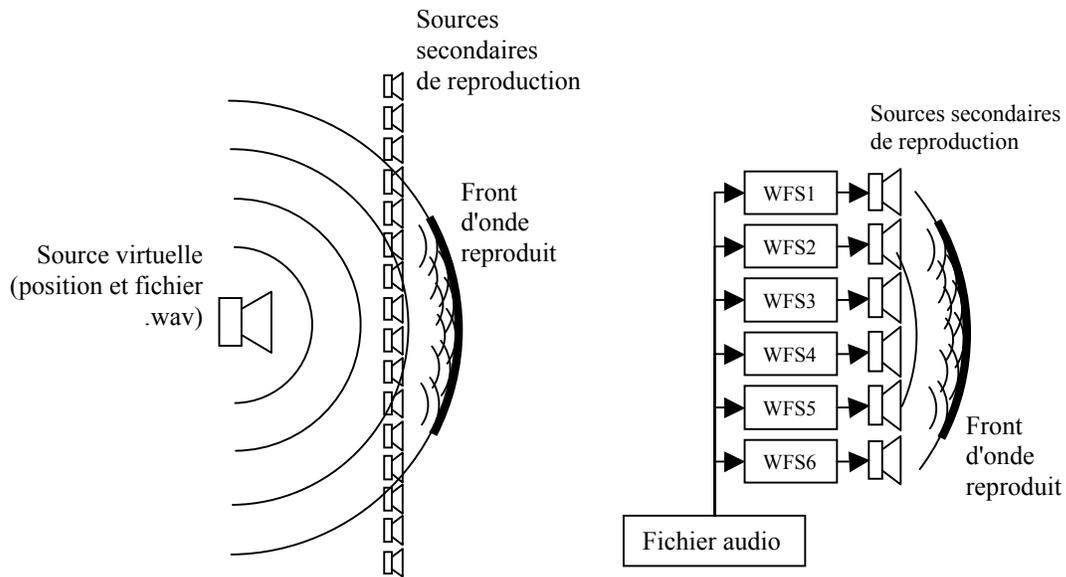
---

<sup>4</sup> En fait, Kirchhoff-Helmholtz permet de remplacer un univers acoustique au-delà d'une surface (S) fermée autour d'un auditeur par une surface couverte de sources au rayonnement omnidirectionnel et bidirectionnel. C'est précisément l'idée théorique de l'holophonie.

entre holophonie, WFS et « Ambisonic », puisque « Ambisonic » de base est un cas particulier de l'holophonie, de la même façon que la WFS est un autre cas particulier de l'holophonie [5].



**Fig. 3 – À gauche, le principe de construction de Huygens. À droite, le principe de construction appliqué à la reproduction de champs acoustiques à l'aide de sources pratiques et réelles.**

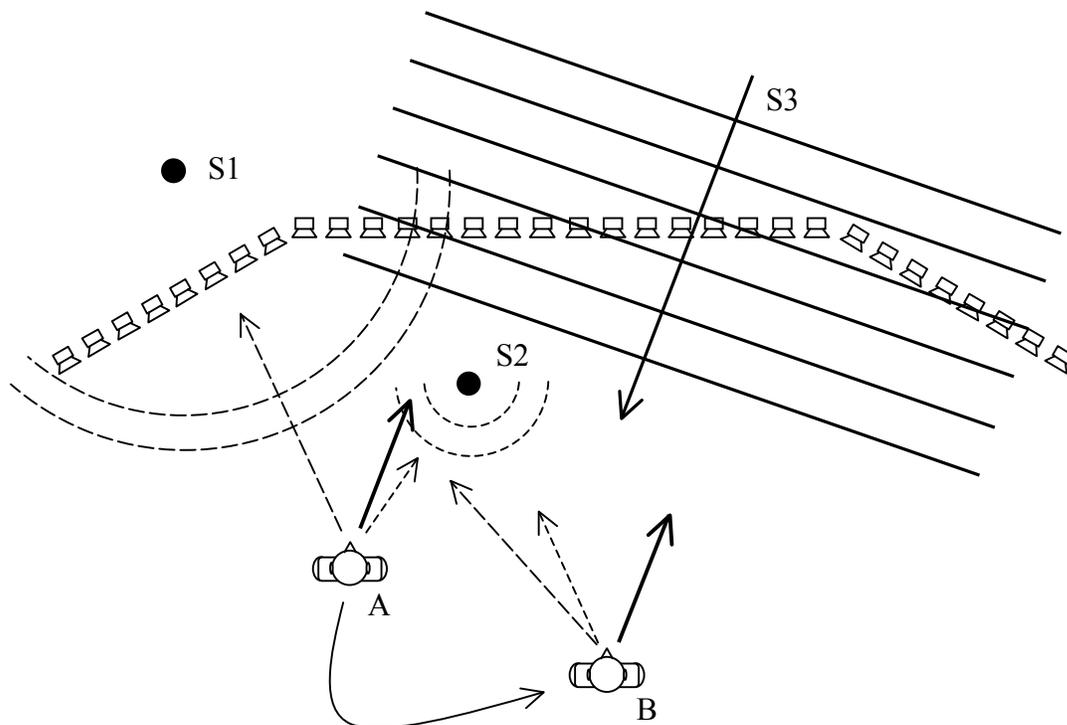


**Fig. 4 – À gauche, le principe de construction de Huygens appliqué à la reproduction. À droite, une schématisation de la réalisation de la WFS pour une source virtuelle seule.**

La stéréophonie conventionnelle peut même, sous l'exercice de quelques restrictions, être considérée comme une version de la WFS. Ainsi, en plaçant deux enceintes au rayonnement omnidirectionnel aux mêmes positions relatives qui furent utilisées lors de la prise de son avec des microphones bidirectionnels, une version clairsemée (pour « sparse » en anglais) de la WFS ou de l'holophonie serait atteinte. En poursuivant l'exercice en ajoutant de plus en plus de sources omnidirectionnelles et de microphones, l'expérience tendrait progressivement vers la WFS et l'holophonie.

## **REVUE DES APPLICATIONS ET DES USAGES POTENTIELS DE LA WFS**

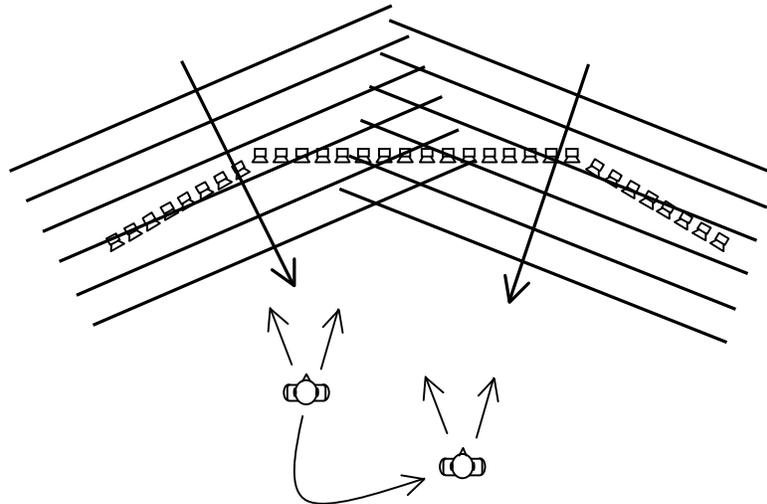
Maintenant que les notions physiques et technologiques supportant la WFS ont été sommairement revues, voyons quel est l'intérêt de reproduire des sources virtuelles simples qu'elles soient sphériques ou planes. La figure 5 propose un exemple de scène virtuelle formée de trois sources de type distinct. Comme on peut le voir, le déplacement de l'auditeur engendre un changement de perspective – de point de vue – sur la scène. Cette possibilité forme un des principaux intérêts de la reproduction de champs acoustiques puisque la scène virtuelle et les mouvements de l'auditeur sont maintenant cohérents. Il n'y a plus vraiment de région d'écoute optimale (« sweet spot ») et l'audience est dès lors soumise à un monde virtuel consistant et correctement distribué dans l'espace. Avec les sources sphériques (S1 et S2; S2 est une onde sphérique reconstruite à l'avant des sources réelles, c'est la focalisation de l'énergie acoustique), il est possible de créer des effets perceptibles de distance et de direction. L'option de pouvoir reconstruire une source virtuelle devant les sources réelles introduit un autre des aspects les plus attrayants pour la création sonore spatiale, notamment au sujet du dynamisme dans le temps qui peut être drastiquement modulé de l'intimité à l'éloignement. Avec une onde plane S3, seule la direction d'incidence est déterminée. Elle ne change pas avec les déplacements de l'auditeur. Ainsi, un déplacement de A vers B donnera toujours une source perçue à +30 degrés dans ce cas. Un exemple d'application typique requiert l'usage d'une source sphérique virtuelle à l'avant alimentée par un enregistrement de prise de son directe d'un instrument acoustique et de huit sources planes (s'étalant sur 360 degrés) alimentées par une réverbération artificielle ou naturelle pour créer un effet de salle virtuelle assurément dissemblable du caractère acoustique de la salle de reproduction.



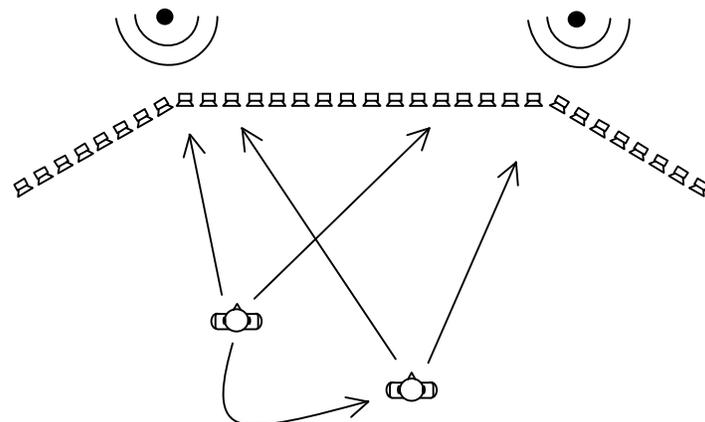
**Fig. 5 – Illustration d’une scène virtuelle formée d’une source sphérique S1, d’une source sphérique « focalisée » en S2 (à l’avant du réseau de haut-parleurs) et d’une source d’ondes planes S3. L’auditeur peut être en deux positions (A ou B) où les directions de localisation des sources virtuelles sont représentées par des flèches aux caractères graphiques correspondant aux sources virtuelles.**

Le fait que les ondes progressives planes n’impliquent pas de changement de direction de l’événement auditif avec le déplacement de l’auditeur révèle une application latente des plus nobles de la WFS : la récupération aisée et la diffusion améliorée de matériel stéréophonique (monophonique, « Surround », etc.) à l’aide de sources virtuelles planes représentant des haut-parleurs dans une configuration standard (2-0.0, 3-2.1, etc.)<sup>5</sup> pour laquelle une œuvre aurait été initialement élaborée. Les figures 6 et 7 présentent deux solutions pour une telle rediffusion de matériel multipiste. Alors que la première procure des sensations similaires pour l’ensemble de l’audience, la seconde propose sensiblement les mêmes effets qu’avec une configuration multipiste réelle (zone d’écoute optimale limitée ou « sweet spot »). Dans les deux cas, ce type d’approche introduit ce qui est parfois appelé le « Virtual PANning » ou VPAN. C’est-à-dire que le mixage audio classique (stéréophonie d’intensité, prises de son stéréo, etc.) est virtuellement applicable par le biais du VPAN avec la WFS. Les techniques de mixage conventionnelles, savamment maîtrisées depuis des années, ne sont conséquemment pas vétustes mais bien propulsées vers le renouveau de leurs usages actuels connus. Évidemment, il est tout aussi facile de représenter un système 3-2.1 ou même un autre système de type  $x$ - $y$ - $z$ . La compatibilité de la WFS avec d’autres formats multipistes est donc très flexible.

<sup>5</sup> On adhère ici à la notation suivante :  $x$ - $y$ - $z$  où  $x$  est le nombre de canaux directs (avant),  $y$  le nombre de canaux d’atmosphère (« Surround » ou arrière) et  $z$  le nombre de canaux d’extrêmes graves [6].



**Fig. 6 – Réalisation d’une scène virtuelle présentant une compatibilité avec la stéréophonie 2-0.0 à l’aide d’ondes planes virtuelles. Les deux auditeurs perçoivent les mêmes directions d’incidence.**



**Fig. 7 – Réalisation d’une scène virtuelle présentant une compatibilité avec la stéréophonie 2-0.0 à l’aide d’ondes sphériques. Le phénomène du « sweet spot » refait surface : selon la position de l’auditeur, l’effet perçu varie.**

Comme il est possible de l’imaginer, les applications potentielles de la WFS sont vastes. Que l’on parle de salle de concert, de cinéma, de présentation audio à plus ou moins grand déploiement et, évidemment ou même surtout, en diffusion de musique électroacoustique, la latitude acquise par l’emploi d’une telle technologie va bien au-delà du simple réalisme de la reproduction et de l’immersion : elle offre un accès à des nouveautés expressives comme des sources virtuelles mouvantes en deux dimensions, des sources qui sortent des sources réelles, des atmosphères virtuelles enveloppantes (un champ diffus ou réverbérant s’approche, ou du moins sa sensation, en superposant par exemple plus d’une dizaine d’ondes planes) et la recherche de la perspective spatiale ou du caractère relatif du point de vue de l’audience sur la scène. La WFS et l’ensemble des

méthodes de simulation de champs acoustiques vont ainsi, si ce n'est déjà amorcé, par les applications potentielles mentionnées, modifier profondément les bases de la composition sonore spatiale.

## **RECHERCHE, DÉVELOPPEMENT ET TECHNOLOGIE**

Depuis la formulation théorique du concept technologique de WFS, maints chercheurs se sont penchés, et se penchent toujours, sur la question de la reproduction de champs acoustiques. Il faut ainsi admettre que cette brève introduction au concept de WFS ne fut pas exhaustive : maints autres détails techniques caractérisent la WFS dont certains réduisent la qualité de la reproduction, notamment en ce qui est relié à l'effet de la salle de reproduction, au recouvrement spatial et à la distribution de la qualité de reproduction sur toute la zone d'écoute. C'est d'ailleurs pour palier à ces quelques lacunes que sont enclenchées une bonne part des recherches actuelles sur la WFS ou la reproduction de champs acoustiques. Ces quelques achoppements ne compromettent pas la qualité de la WFS classique, telle qu'elle est ici décrite, mais établissent les tangentes de recherche pour l'amélioration de la reproduction spatiale sur réseau de haut-parleurs.

Ainsi, à titre d'exemple, nous avons publié les premiers résultats de nos recherches sur la reproduction de champs acoustiques progressifs en salle [7]. Parmi les résultats importants, il fut démontré qu'il est théoriquement possible de recréer des champs acoustiques progressifs dans un lieu peu amorti qui présente donc une dynamique fortement modale en basses fréquences. Pour atteindre une telle reproduction, deux mécanismes physiques furent identifiés : l'absorption de puissance acoustique par certaines sources et le réarrangement de la réponse modale de la cavité. Ce type d'observations suggère de nouvelles pistes technologiques pour l'avenir.

Une grande part des recherches récentes en WFS ont été réalisées en Europe par le consortium de recherche CARROUSO qui incluait notamment l'Université Polytechnique de Delft (le lieu de naissance de la WFS), l'IRCAM, l'institut Fraunhofer et d'autres groupes peut-être moins connus en Amérique du nord [8]. Les recherches en question ont couvert un nombre considérable de sujets : traitement de signal, imposition d'une directivité de rayonnement aux sources virtuelles, mouvement des sources virtuelles, interfaces informatiques visuelles, type de sources de reproduction (électrodynamique, directivité, invisibilité, panneaux sources), évaluation subjective de la perception spatiale de la WFS et compensation de la salle de reproduction. Suivant l'ensemble des développements technologiques proposés par CARROUSO, les applications commerciales sont éminentes et seront assurément lancées par des protagonistes de l'ancien consortium. On pourra par exemple surveiller Fraunhofer qui propose une application commerciale de WFS baptisée IOSONO. Beaucoup d'énergies sont aussi déployées pour valoriser l'idée de la WFS, notamment par le biais de fréquentes démonstrations et présentations dans les conférences internationales et les conventions de l'AES (« Audio Engineering Society »).

Ce bref aperçu démontre l'activité et la vivacité de la recherche et du développement de la WFS ou de la simulation de champs acoustiques. Une pareille activité laisse espérer

une multitude de débats et d'applications réelles lors des années à venir. Il est ainsi probable que la technologie parvienne à s'installer fructueusement en quelques domaines d'application.

## **FUTUR DE LA DESCRIPTION SPATIALE ET CONSIDÉRATIONS POUR LA CRÉATION**

L'avènement et la standardisation de nouveaux formats de données multimédias structurés comme le MPEG-4 seront aussi une des pièces clés des plus opérantes dans le développement et l'épanouissement de la reproduction spatiale du son. Ce type de format, qui peut évidemment contenir des pistes audio sans compression de données, permet d'y inclure des instructions de traitement de signal et des descriptions de scènes virtuelles. Pour cette raison, les formats MPEG-4 ont retenu l'attention de ceux et celles qui sont concernés par l'interaction avec une virtualité [9-11], aussi variée qu'elle soit. Avec une telle approche on parle d'*encodage spatial* et non plus d'*encodage par canal*. Contrairement au temps où il employait des méthodes de stéréophonie traditionnelles (l'encodage par canal), lesquelles impliquaient des opérations de mixage sur un nombre variable de canaux pour atteindre manuellement la scène désirée, le compositeur qui travaillera dès lors sur l'encodage spatial sera beaucoup plus près de la scène virtuelle qu'il tente de décrire. Le contact entre l'œuvre et son créateur devrait ainsi être plus immédiat avec l'encodage spatial. En bout de ligne, c'est le système de reproduction qui sera en charge de décoder la scène virtuelle pour la présenter de son mieux en fonction des exigences de la scène virtuelle.

Avec le recul attribuable à cette brève présentation de la WFS, il est intéressant de revenir sur la figure 1 qui avait initialement posé un certain classement sur les méthodologies possibles pour la reproduction du caractère spatial de l'audition. Ainsi, au sujet de l'avenir de telles technologies, on peut se questionner au sujet de la compatibilité, de la complémentarité ou même du combat entre la simulation de champs acoustiques et la simulation de la perception. Pour l'instant les opinions sont partagées. Certains ne prêchent que par la simulation de la perception alors que d'autres ne croient qu'en la simulation de champs acoustiques. En gros, lorsqu'un système de diffusion s'adresse à un seul auditeur, la simulation de la perception est des plus séduisantes puisqu'elle demeure simple à réaliser avec un minimum d'équipement et de connaissances spécialisées. Évidemment, pour un système à grand déploiement pour une audience complète, il serait probablement plus efficace de se tourner vers la simulation de champs acoustiques. Sachant que l'argument en faveur de la simulation de la perception est essentiellement la simplicité de sa mise en œuvre et qu'en présence d'une audience complète un tel type de simulation devrait nécessairement inclure des systèmes d'adaptation aux mouvements de chacun des auditeurs (en faisant référence aux technologies binaurales dans ce cas), il est clair qu'un seuil de complexité équivalent à celui de la simulation de champs acoustiques est vite atteint. C'est un argument en faveur de la reproduction de champs acoustiques pour des situations de concerts par exemple.

## CONCLUSION

Il faut maintenant admettre que la WFS n'est pas nécessairement la solution miracle. Ce bref exposé n'avait pas comme objet l'aveugle ou partielle mise en valeur de cette technologie spécifique – qui se présente comme une des plus importantes pour le futur – mais bien sa simple présentation. Oublier l'intérêt et l'utilité des sources réelles (acoustiques ou électroacoustiques) et le pouvoir de l'installation sonore serait bien désolant puisque la figuration et même la signification des sources réelles peuvent jouer un rôle primordial pour certaines œuvres [12, 13]. Dans la grande famille de la diffusion sur système multi haut-parleurs, la WFS présente d'un autre côté des avantages des plus notables : d'abord la possibilité de créer des scènes virtuelles complexes par le biais de l'encodage spatial puis l'ouverture vers la compatibilité avec des formats multicanaux bien connus (2-0.0, 3-2.1) sur lesquels dépendent assurément déjà un nombre considérable d'œuvres électroacoustiques.

En fait, le détail des technologies importe peu dans un contexte de création. La flexibilité, l'expressivité, l'immatérialité ou la transparence du média et l'immortalité des œuvres sont peut-être les caractères les plus prisés. L'avenir devrait fournir de telles caractéristiques puisque les recherches abondent en ces voies, et la synthèse de champs dite WFS est un des moyens probants pour nous conduire vers ces caractéristiques.

Savoir prendre cette introduction comme une présentation très générale et approximative du sujet est nécessaire, le lecteur est ainsi prié de se rapporter aux références pour plus d'informations sur la WFS. La thèse de Verheijen est un excellent et complet ouvrage sur le sujet [14]. D'autres références sont introduites à titre de ressources complémentaires [15-18].

## RÉFÉRENCES

1. J. Blauert, *Spatial Hearing, The psychophysics of human sound localization*, The MIT Press, 1999.
2. W.W. Rouse Ball, *A Short Account of the History of Mathematics*, Dover publications, 1960.
3. M. Jessel, *Acoustique Théorique : Propagation et Holophonie*, Masson et cie, Paris, 1973.
4. A.J. Berkhout, D. de vries, P. Vogel, Acoustic control by wave field synthesis, *Journal of the Acoustical Society of America*, 1993, vol. 93, no 5, p. 2764-2778.
5. R. Nicol, M. Emerit, 3D-sound reproduction over an extensive listening area: A hybrid method derived from holophony and ambisonic, *Proceedings of the AES 16th international conference*, 1999, p. 436-453.
6. F. Rumsey, *Spatial Audio*, Focal Press, 2001.
7. P.-A. Gauthier, A. Berry, W. Woszczyk, Sound field reproduction using active control techniques: Simulations in the frequency domain, présentation invitée et cahier des actes, *International Congress on Acoustics*, Kyoto, Japon, 2004.
8. S. Brix, T. Sporer, J. Plogsties, CARROUSO – An European Approach to 3D audio, *AES Convention paper*, 2001.
9. P.R. Cook, *Real Sound Synthesis for Interactive Applications*, A K Peters, 2002.
10. J. Plogsties, O. Baum, B. Grill, Conveying spatial sound using MPEG-4, *Proceedings of the AES 24th international conference*, 2003, p. 58-65.

11. E.D. Scheirer, Structured audio and effects processing in the MPEG-4 multimedia standard, *Multimedia Systems*, 1999, vol. 7, p. 11-22.
12. M. Trochimczyk, From Circles to Nets: On the Signification of Spatial Sound Imagery in New Music, *Computer Music Journal*, 2001, vol. 25, no 4, p. 39-56.
13. S. Wilson, *Information Arts – Intersections of Art, Science, and Technology*, The MIT Press, 2002.
14. E.N.G. Verheijen, *Sound Reproduction by Wave Field Synthesis*, thèse de doctorat, Université Technique de Delft, 1998.
15. M.F. Davis, History of spatial coding, *Journal of the AES*, 2003, vol. 51, no 6, p. 554-569.
16. E. van der Heide, *A world beyond the loudspeaker*, [http://utopia.ision.nl/users/heide/world\\_beyond.html](http://utopia.ision.nl/users/heide/world_beyond.html).
17. G. Theile, H. Wittek, M. Reisinger, Potential wave field synthesis application in the multichannel stereophonic world, *Proceedings of the AES 24th international conference*, 2003, p. 43-57.
18. D. de Vries, R.Boone, Wave field synthesis: new improvements and extensions, *Proc. of the 18th International Congress on Acoustics*, 2004.