

Chers amis, chers amateurs de Musique,

L'idée d'écrire ce texte m'est venue à la suite d'une discussion par courriel avec un audiophile, au cours de laquelle j'ai découvert que celui-ci ne comprenait pas vraiment la *nature réelle* de l'enregistrement. De surcroît, il refusait de considérer à quel point c'est un élément critique dont tout le reste dépend.

M'étant penché un peu plus sérieusement sur la question, je me suis alors aperçu que cette incompréhension était parfois partagée, et que nombreux étaient ceux, y compris certains concepteurs, qui n'avaient pas les idées très claires sur le sujet.

Comme la prise de son a occupé une place importante dans mon travail, je me suis décidé à produire ce document, dont le titre s'est naturellement imposé de lui-même, clin d'œil à Woody Allen.

Bonne lecture.

Tout ce que vous avez toujours voulu savoir sur l'enregistrement, sans jamais oser le demander.

Ce qui m'a alerté chez cet audiophile c'est qu'il défendait l'idée que l'enregistrement **est une captation de la réalité**. Malheureusement, c'est un point de vue inexact car, au mieux, un enregistrement ne peut être qu'une simple **représentation de celle-ci**.

Ok, ce n'est pas évident, cela demande des explications :

Représenter l'espace : une question centrale.

Je n'aime pas cela, mais pour rendre les choses plus claires, je vais avoir recours à la comparaison avec l'image :

Supposons que vous preniez une photo d'une rue. En faisant cela vous n'avez pas capté la réalité, vous l'avez seulement *représenté*. En effet, la rue est un objet à trois dimensions, dans lequel vous pouvez déambuler. Pas la photo, qui est réduite à *un simple plan*. Visuellement, elle **suggère** bien l'existence d'une troisième dimension, la profondeur, au moyen de la perspective, mais vous ne pouvez pas vous promener physiquement dedans. De surcroît, l'espace est *continu* dans le réel, mais *discret* dans la photo, car elle a découpé, séparé l'espace en imposant *un cadre*.

Ainsi, comme nous le voyons, dans le cas d'une représentation, ce n'est plus l'objet qui définit l'espace, c'est la méthode de prise de vue.

Mais attention, en imaginant que vous puissiez prendre une photo véritablement en trois dimensions, on pourrait croire qu'il va y avoir une captation exacte du réel. Mais non, c'est toujours une **représentation**, même si elle est évidemment **beaucoup plus proche de la réalité**. Pourquoi ? Parce que la photo 3D impose tout de même **son cadre, ses limites physiques et son point de vue**. C'est toujours elle qui définit l'espace !

Et c'est exactement la même chose pour l'enregistrement ! Quand on enregistre un quatuor à cordes, ce ne sont pas ses caractéristiques spatiales *réelles* qui sont exhaustivement captées, mais seulement celles permises par la méthode de prise de son. Ainsi, comme pour la photo, l'enregistrement va **fixer définitivement** une certaine *représentation* de l'espace.

Compte tenu de ce que j'ai lu et entendu, cela ne semble pas aller de soi pour tout le monde.

De l'importance de l'espace dans la perception auditive.

J'ai déjà plusieurs fois souligné combien la perception auditive est souvent mal comprise, voire déformée, avec des conséquences néfastes sur l'élaboration des systèmes : *comment répondre exactement à un besoin quand on a une image inexacte de ce qu'il est réellement?*

Je rappelle au préalable que l'audition humaine est fondamentalement binaurale, ce qui signifie simplement qu'on a besoin de nos deux oreilles pour percevoir l'espace. Notre perception spatiale ressort d'un fonctionnement cognitif complexe, basé principalement sur la différence de temps *interaurale*, mais aussi sur la différence d'intensité *interaurale*. S'y ajoute un travail cognitif de mémorisation et d'analyse nous permettant de lever certaines ambiguïtés de localisation.

Je voudrais une nouvelle fois insister sur la dimension absolument fondamentale de la perception spatiale dans l'audition humaine. Elle fut la condition de notre survie en tant qu'espèce au sein d'un environnement hostile. Au travers de *l'effet de précedence*, - l'absolue prééminence cognitive de l'onde directe, celle qui nous parvient en premier - nous repérons non seulement la **direction** de la source sonore, mais aussi sa **distance** et sa **qualité**.

On notera à ce sujet une confusion, très répandue, qui explique un certain nombre de croyances : en effet nous ne sommes plus capables *consciemment* de séparer les sons *lorsque la différence temporelle entre onde directe et onde réfléchi est inférieure à environ 30ms*. Néanmoins cela n'empêche pas notre cerveau de continuer imperturbablement à discriminer et privilégier l'onde directe, *sans que nous en ayons conscience, afin de localiser sa source*.

Pour résumer, il y a traditionnellement focalisation sur la seule dimension *perceptive* de l'audition au détriment de sa dimension *cognitive*.

Une conséquence, par exemple, c'est que la doctrine habituelle qui tient l'acoustique environnante pour prépondérante est basée sur *l'apparence*, ce que nous percevons (saupoudré de quelques fantasmes*), mais pas sur le **fonctionnement réel de notre audition**.

Un moyen simple de sentir l'importance de *l'effet de précedence* est de s'enfermer dans la douche et de laisser tomber une petite pièce de monnaie, tout en gardant les yeux fermés. Si l'on considère l'exiguïté, le caractère ultra réfléchissant des surfaces carrelées, l'extrême concentration d'ondes réfléchies dont les différences temporelles avec l'onde directe sont exceptionnellement faibles (quelques centaines de μ s), un tel *brouillage temporel* devrait rendre très difficile la localisation de la pièce de monnaie, d'autant que s'y ajoute la nécessité d'une localisation verticale.

Or, non seulement son point de chute et son éventuel trajet sur le sol sont analysés - vous pouvez assez facilement la ramasser sans avoir besoin d'ouvrir les yeux - mais la nature des matériaux - métal, carrelage - l'est tout aussi nettement.

* Il s'agit souvent, me semble-t-il, d'imiter un peu les musiciens, toujours prompts à discuter et critiquer l'acoustique d'une salle. Une manière de dire " mon ouïe est si *fine* que le moindre défaut me perturbe". En quelque sorte la réplique *audio* de la [princesse au petit pois](#) d'Andersen ;-)

A chaque méthode, son espace.

- **Monophonie** : si l'on oublie son ancienneté, un enregistrement mono peut présenter des timbres parfaitement respectés, mais il ne contient par contre **aucune information spatiale**. Il est donc incapable de *représenter* l'espace, ni en largeur ni en profondeur.

Ce qui est intéressant de remarquer, c'est ce qui advient lorsqu'un tel enregistrement est reproduit sur un système stéréophonique. Puisqu'il ne contient aucune représentation spatiale en lui-même, en toute logique, on ne peut aucunement **préjuger du résultat** : ainsi la restitution qui s'étale largement entre les enceintes est une reproduction tout aussi valable que lorsque la présentation est ponctuelle ! Comme il n'y a pas d'image à restituer, ce sont par conséquent *les enceintes qui la fabriquent*, du fait de leur *mode de propagation*.

Néanmoins, j'ai appris que des *audiophiles* utilisaient un signal mono pour tester l'image stéréo de leur système. **Je suis surpris qu'aucun n'ait remarqué qu'évaluer l'image avec un signal ne possédant aucune image était totalement absurde !**

En procédant ainsi, ce qu'ils observent est en réalité la capacité des enceintes à *focaliser l'énergie*. Et dans ces conditions, les pavillons ou guides d'onde, par leur principe même de fonctionnement, vont fournir une image plus ponctuelle. Autrement dit, plus l'enceinte est directive, plus l'image est ponctuelle et resserrée, mais cela n'a *rien à voir* avec la capacité à fournir une *véritable* image spatiale, qui est un phénomène bien plus complexe !

En effet, dans cette situation, imaginer que la reproduction *doit obligatoirement* être ponctuelle est abusif : c'est projeter une observation faite dans la *réalité* sur une *représentation* qui en diffère radicalement par sa nature même, **puisqu'elle ne contient pas d'espace**.

On remarquera, au passage, que c'est un à priori d'autant plus facile à avoir quand on ne saisit pas que c'est la méthode de prise de son qui détermine l'espace, et non la source enregistrée...

- **Stéréophonie d'intensité** : c'est LA technique de prise de son que nous écoutons tous en permanence, puisqu'elle représente au pifomètre 95% des enregistrements, chiffre certainement en dessous de la vérité. En simplifiant, dès qu'il y a une prise multi micros, souvent de proximité, c'est de la *stéréophonie d'intensité*.

Ce terme signifie que le placement dans la largeur de l'espace stéréophonique – déterminé lors de la restitution par l'espacement entre les enceintes – est fonction de l'amplitude : si un instrument a un niveau sonore supérieur dans le transducteur de gauche par rapport à celui de droite, son placement sera décalé vers la gauche.

Sur le plan de la représentation de l'espace, la stéréophonie d'intensité ne contient **aucune information de hauteur, ni de profondeur**, seulement de largeur.

De la même façon que la photo **suggère** la profondeur au moyen de la *perspective*, la stéréophonie d'intensité suggère la profondeur avec la *réverbération* : un son fortement réverbéré peut évoquer l'éloignement, en fonction soit de la distance réelle du micro à la source, ou le plus souvent, des réglages effectués sur un [périphérique de traitement](#).

J'ajoute trois remarques à cette description :



Comme le montrent les deux images ci-dessus, la prise d'une source sonore en stéréophonie d'intensité se fait la plupart du temps avec un seul micro.

Si cela simplifie le travail de captation - un micro = une piste sur la console, pas de problèmes de phase (voir le paragraphe sur la batterie) – cela a une conséquence : si la source est placée dans la largeur de l'espace stéréo avec le *panoramique*, d'origine, il n'y a **aucune information** relative à sa **dimension**. Pour que la **largeur propre de la source puisse être captée**, il faut avoir recours à deux micros, comme dans l'exemple ci-dessous.

A noter qu'une fois captée, la *taille* de la source au sein de l'image stéréo peut ensuite être manipulée à loisir¹.



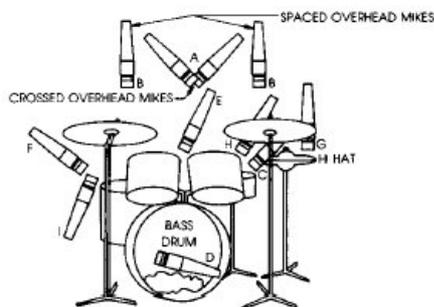
- Ainsi, en stéréophonie d'intensité, l'image spatiale est une construction (ou une reconstruction) totalement fabriquée dans presque tous les cas. En conséquence, la notion même de *représentation de la réalité* devient évidemment problématique.

Pour l'anecdote, il est toujours amusant de lire des commentaires concernant l'étagement des plans sonores d'une enceinte, y compris dans de nombreux magazines ayant pignon sur rue, alors que les enregistrements écoutés sont tous du type *stéréophonie d'intensité*, c'est à dire ne contenant généralement aucune information de profondeur...

Plus sérieusement, la généralisation d'enregistrements sans profondeur explique la résurgence régulière, presque obsessionnelle, de techniques susceptibles de compenser cette lacune : quadraphonie, surround, procédé ambisonic, etc.

Pour ceux qui découvrent ces questions de prise de son, une conséquence tout à fait inattendue de la méthode concerne la *restitution des timbres*. En effet lorsqu'il y a plusieurs micros, les signaux qu'ils captent sont susceptibles d'interférer les uns avec les autres, et de déformer les timbres.

Un parfait exemple pour démontrer l'importance du phénomène est la prise de son de la batterie. Voici ci-dessous le dispositif traditionnel pour capter un " kit " sous la forme d'un schéma, puis d'une illustration pratique :



On trouve donc deux micros (overhead, séparés ou en couple) chargés de capter les cymbales et accessoirement la batterie dans son ensemble. Ensuite chaque fût dispose de son propre capteur, sans oublier la hi-hat (pédale *charleston* en français). Imaginez maintenant que le batteur frappe la caisse claire. Le micro dédié situé juste à côté (~ 25 cm) enregistre la frappe. Mais celui de la hi-hat (~ 70 cm) le fait aussi, tout

comme les différents micros des toms (~ 40cm + ~ 50cm + ~ 80cm). Ensuite, il ne faut pas oublier les deux overhead qui captent aussi le même événement (~ 1m + ~ 1,1m).

En conséquence la même frappe se trouve enregistrée par plusieurs capteurs avec des décalages temporels successifs (soit ~ 0,44ms + 0,7ms + 1,6ms + 2,9ms + 3,2ms). Bien sûr, ces délais sont trop faibles pour que l'on puisse *percevoir* quoi que ce soit, et on entend bien qu'un seul coup. Cependant, lorsqu'ils sont mélangés ensemble pour restituer l'instrument en entier, ces sons identiques, **mais décalés en phase**, créent de violents accidents d'amplitude² : à cause de cela, le timbre est transformé, il ne **peut pas** être *fidèle*.

C'est pourquoi l'enregistrement d'une batterie est une sorte de *bricolage* personnel du producteur et/ou de l'ingénieur du son, où timbre et espace sont recréés : en quelque sorte une *fiction sonore*. Un instrument *artificiel* auquel nous sommes habitués depuis longtemps, certes, mais néanmoins **irréel**. Et les batteurs ont raison de sempiternellement se plaindre du fait qu'ils ne retrouvent jamais sur l'enregistrement le timbre exact de leur " kit ", qu'ils ont pourtant si soigneusement travaillé et mis au point...

Ce problème de la *reprise* (d'une *élégance douteuse*, le terme décrit le fait qu'un instrument est capté par le micro dédié à un autre) va devenir crucial avec le multi pistes. En effet ce type d'enregistreur permet à un musicien de refaire (*re-recording* ou *dubbing*) individuellement et autant de fois que nécessaire, certaines parties de sa prestation qu'il veut améliorer. Or, la présence sur sa piste du son d'un autre instrument, même à bas niveau, constitue un obstacle à cette pratique.

Il devient alors capital d'isoler acoustiquement chaque instrument le mieux possible. Ce qui explique l'apparition dans les studios de cabines insonorisées, l'usage de micros directifs placés très près, de casques et de la *captation directe* (l'instrument électrique est raccordé à la console via une *boîte de direct*, il est enregistré sans micro), transformant ceux-ci en véritables *tombeaux acoustiques*, pour reprendre l'expression de Joe Jackson.

En conséquence, le *mixage* devient la phase essentielle, où chaque partie va être individuellement transformée (égalisations, spatialisation, contrôle de la dynamique, effets divers, etc.) pour s'intégrer dans l'ensemble. Le degré considérable de modification aboutissant la plupart du temps à un résultat sonore n'ayant plus **rien à voir avec la réalité acoustique initiale**, comme le démontre par exemple l'usage immodéré de la *compression dynamique*.

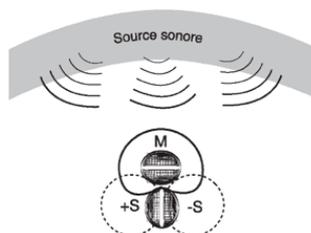
J'ai une fois associé la fameuse *malbouffe* – des produits artificiels, entièrement produits industriellement - à l'existence d'un *malson* du même acabit.

Compte tenu de ce que je viens de décrire, et qui est devenu la règle, peut-être ne suis-je pas si loin de la vérité, non ?

Quant à ceux, probablement nombreux, qui trouveront que je grossis le trait, qu'ils songent un instant aux innombrables personnes qui n'ont malheureusement comme unique expérience du **concert** que le *brouet* qui s'échappe d'une sonorisation acoustiquement infâme, et avec un niveau sonore tel qu'il faut se munir de bouchons d'oreille.

Se boucher les oreilles pour écouter de la musique ! Est-il besoin d'en dire plus ?

1/ Cette caractéristique s'avère particulièrement intéressante pour la prise de son cinéma, avec notamment la technique dite MS en cas de captation du *son direct*. Elle permet de faire correspondre facilement la taille de l'image stéréo au champ de la prise de vue.



2/ Rappelons que *phase* et *amplitude* sont étroitement liés et interagissent sur le signal.

- **Stéréophonie de phase** : sur le plan de la capacité à représenter l'espace, c'est de loin la méthode la plus performante, la plus *fidèle*, puisque sont captées les informations temporelles, se rapprochant ainsi de l'audition humaine. Lorsqu'elle est bien maîtrisée, cette méthode permet de donner une image plausible des dimensions de la source, tant en largeur qu'en profondeur*.

Cependant, un certain nombre de raisons expliquent le faible nombre d'enregistrements effectués avec cette technique, pourtant supérieure qualitativement.

- Elle est incompatible avec la monophonie. Or, souvenez vous que la technique dite *gravure universelle* des disques vinyles nécessitait que le grave soit mono. Cela n'a plus beaucoup de sens désormais, mais l'habitude en a été prise à cette période.

- Elle interdit le recours à l'enregistrement multi-pistes, et en conséquence oblige les musiciens à jouer parfaitement, **simultanément**. Elle contraint donc ceux ci à se placer et jouer en "situation de concert". C'est facile pour des instruments acoustiques, mais quid des instruments électriques et/ou électroniques ? D'autre part cela oblige les instrumentistes à être meilleurs puisqu'ils n'ont plus le droit à l'erreur et au re-recording.

- Elle tend à placer l'auditeur à *distance de la source sonore*, comme s'il se trouvait lui aussi en "situation de concert". Or, nous allons voir plus loin que cette caractéristique n'est pas anodine du tout, ni souhaitable du point de vue de *l'industrie culturelle*.

- Comme la stéréophonie de phase permet de capter la profondeur, elle excelle pour l'enregistrement des petites formations, de l'orgue, ou lorsqu'il est essentiel de restituer l'acoustique environnante, comme dans le *chant grégorien* par exemple.

On peut en voir un exemple sur la photo ci dessous, prise lors de l'enregistrement par Jean-Guihen Queyras des [6 suites pour violoncelle](#) de JS Bach chez Harmonia Mundi. On aperçoit le couple AB de micros DPA 4007 à environ 2 m devant l'instrumentiste, et l'église romane qui apporte son acoustique naturelle.



Cependant cette capacité à restituer la distance peut devenir un inconvénient dans le cas d'une source sonore de grande dimension, telle qu'un orchestre symphonique. En effet les micros, placés au niveau du chef d'orchestre, se trouvent proches des premiers violons et des violoncelles, mais à distance importante des percussions !



Or, à titre d'exemple, les timbales ou la grosse caisse demandent à être enregistrées d'assez près, sous peine de paraître floues dans le grave. Une amélioration relative consiste traditionnellement à surélever les micros afin de diminuer l'écart de distance entre l'avant et l'arrière de l'orchestre, mais on peut aussi, si on utilise un couple AB, *tricher* en plaçant, de façon moins conventionnelle, les percussions à proximité immédiate, dans le dos du chef d'orchestre !

* Mais pas en hauteur. Pour pouvoir restituer une telle information, il faut avoir recours à une tête artificielle pour l'enregistrement et écouter exclusivement au casque... A part ce cas très spécifique, les prises de son ne permettent pas de capter d'informations relatives à la hauteur. Si on observe des variations dans ce domaine, c'est exclusivement dû à des défauts des enceintes, en particulier à cause d'irrégularités des diagrammes polaires.

Quels enseignements?

Le premier, c'est qu'il n'existe pas de méthode parfaite, qui puisse répondre à toutes les situations. Tout est une question d'objectif, de ce que l'enregistrement cherche à privilégier dans la musique.

Bien sûr la *stéréophonie de phase* est intrinsèquement plus *fidèle*, plus proche de la réalité en raison de sa capacité à mieux restituer *l'espace* et à capter l'acoustique du lieu. Mais elle est aussi plus délicate à maîtriser et, surtout, elle est considérée comme inadaptée aux pratiques de la musique populaire et des instruments électriques.

Ce dernier point explique pour une large part l'hégémonie extrême exercée par la *stéréophonie d'intensité*. Néanmoins, cela se fait incontestablement au détriment de la faculté de restituer convenablement l'espace. Mais pas seulement, car, au-delà, c'est la volonté de représenter une certaine *réalité sonore* qui se trouve fondamentalement questionnée, au profit de ce qu'il faut bien appeler une *fiction, une production artificielle*.

Je voudrais maintenant essayer d'expliquer succinctement pourquoi et comment notre paysage musical et sonore a évolué dans cette direction.

Le son, l'intime et la massification.

La première remarque qu'il faut faire, c'est indiquer que, contrairement à ce que l'on entend souvent, l'évolution des techniques n'est pas la principale source du changement de conception de l'enregistrement que nous avons connu.

Pour la trouver, il faut plonger loin dans le passé et l'histoire de la Musique, il y a plus de 100 ans, au tout début du XX^{ème} siècle.

Le son, un matériau nouveau.

Émerge en Italie, à partir de 1909, un nouveau mouvement artistique, le [futurisme](#), qui exalte le monde moderne, l'industrie, la puissance, la vitesse, et bouscule l'académisme.

Parmi ces futuristes, un compositeur et peintre, [Luigi Russolo](#), publie en 1913 un livre intitulé "[l'Art des bruits](#)", qui proclame la nécessité d'un *son nouveau*. Ce mouvement, ces idées, vont exercer une influence considérable sur un musicien, qui lui aussi se heurte au conformisme des institutions, [Edgar Varèse](#).

Durant toute sa carrière, Varèse va avoir l'obsession **du son** : depuis l'introduction de sirènes au sein de l'orchestre, en passant par l'écriture d'une des premières pièces pour percussions seules, [lonisation](#), jusqu'à son intérêt permanent pour les balbutiements de la lutherie électronique ([Ondes Martenot](#), [Theremin](#), etc). Tout ceci aboutissant au mémorable scandale de [Déserts](#), puis à *Poème électronique*.

C'est ainsi que s'impose pour la première fois, un idée fondamentale pour notre propos : **le son, considéré comme un matériau pur, en soi, une glaise modelable et manipulable à loisir.**

La suite s'écrit en France puis en Allemagne, dès la fin de la 2^{ème} guerre mondiale. [Pierre Schaeffer](#) et [Pierre Henry](#), partent de sons réels enregistrés, *concrets*, qu'ils manipulent et transforment ensuite. [Karlheinz Stockhausen](#), qui travaille un moment en 1951 avec les deux Pierre, ne procède pas de la même manière : l'allemand crée des sons nouveaux en s'appuyant sur la série de Fourier et avec des générateurs. Mais tous partagent le même outil : le studio (WDR à Cologne et GRM à l'ORTF), qui devient le lieu alchimique, le territoire de la création, où l'on *forge* un **son nouveau**, proprement *inouï*.



Pierre Henry, Pierre Schaeffer



Karlheinz Stockhausen

Une vingtaine d'années plus tard, cette idée va "sauver" les Beatles. Ceux-ci ne supportent plus ces concerts où 20 000 teenagers hurlant couvrent leurs instruments et les empêchent de s'entendre quand ils jouent. Ils veulent fuir la scène.

Ils rencontrent un producteur dans la vénérable maison EMI, Georges Martin, qui s'intéresse à ce qui se passe dans l'avant garde, et est prêt pour l'aventure. Un monde nouveau s'ouvre devant eux qu'ils ignoraient. L'album *Revolver* (1966) est le premier fruit (écoutez *Tomorrow Never Knows*), mais c'est avec l'album *Sgt Pepper's Lonely Heart Club band* - Karlheinz Stockhausen est en photo sur la pochette - que le basculement devient définitif : **le studio d'enregistrement est devenu un lieu de création, et sa raison d'être n'est plus de tenter de capturer une certaine réalité sonore, mais de devenir une *manufacture* du son.**

A partir de là, rien ne sera plus comme avant, musiciens et technique vont suivre ensemble ce chemin tout tracé. La puissance du studio, devenu laboratoire et usine, va accroître son emprise sur la musique, en particulier populaire, mais pas que : multiplications du nombre des pistes, généralisation des traitements du son, lutherie électronique, etc.

La signature sonore.

De surcroît, la possibilité de modeler le son, le timbre, presque sans limite, va offrir une solution pertinente à un potentiel problème.

Comme je l'ai déjà évoqué, le *langage musical* de la musique populaire est généralement simple, souvent rudimentaire (harmonie *tonale* basique, structure couplet/refrain presque immuable, signature rythmique systématique) lorsqu'on le compare, par exemple, à l'extrême complexité de celui de la musique savante dite *contemporaine* qui s'élabore durant la même période.

Or, avec l'explosion du *marché de la musique*, la *massification culturelle* postérieure aux années 60, il devient de plus en plus difficile de **se différencier** des autres quand on partage fondamentalement le même langage et que celui-ci offre relativement peu de possibilités de variations.

C'est donc le son qui va endosser en grande partie le rôle majeur de *personnification* de l'artiste, autorisant ainsi la reconnaissance immédiate de son *identité musicale*.

En conséquence, le *producteur* voit sa responsabilité devenir centrale dans la réalisation d'un album, en ce qu'il est en charge de sa *personnalité*, et le studio se trouve renforcé dans son statut de *forge sonore*.



L'intime.

Dans la première partie de ce document, j'ai indiqué la différence fondamentale de *représentation spatiale* entre stéréophonie de phase et d'intensité. Cependant si l'absence de vraie restitution de la profondeur des prises de son multi micros constitue un déficit évident pour les *mélomanes*, **en s'éloignant de la réalité**, elle offre paradoxalement un avantage important pour *l'industrie culturelle* lorsqu'elle vise le *grand public*.

C'est d'autant plus notable que la musique est devenue omniprésente - de l'ascenseur à la voiture, en passant par le smartphone ou les magasins - dans des lieux où, à priori, elle n'avait auparavant pas sa place.

Or, justement, la *stéréophonie de phase* - sa restitution de la profondeur et de l'acoustique - tend à placer l'auditeur en position de **spectateur**, *comme dans un concert*. Il est en quelque sorte *extérieur* à l'événement, qui se produit à distance de lui : c'est la Musique qui est au centre, et c'est en conséquence **à lui d'aller vers elle**.

Tout au contraire, la *stéréophonie d'intensité* supprime toute distance entre l'auditeur et la musique, puisqu'il n'y a pas de profondeur. **C'est elle qui vient vers lui, établissant nécessairement un rapport intime**. Qu'importe si les musiciens donnent l'impression de jouer sur ses genoux, l'important c'est que l'auditeur ait la sensation qu'il est au centre du dispositif et que l'événement n'a lieu que pour *lui*.

Bien évidemment ce lien *intime*, fortement teinté de narcissisme, constitue un facteur fondamental dans la relation commerciale, pour *impliquer et séduire le consommateur*.

En conséquence, il est parfaitement logique que l'hégémonie de la stéréophonie d'intensité ait accompagné la société de consommation et la montée en puissance de l'industrie culturelle.

Malheureusement, quand on essaie de le faire sortir par la porte, il arrive fréquemment que le *consumérisme* revienne par la fenêtre...

Autre remarque, surprenante à première vue, c'est le fait que la *culture audiophile* ait intégré les limitations intrinsèques de la *stéréophonie d'intensité*, au point de faire l'apologie de systèmes particulièrement inadaptés à la reproduction correcte de la profondeur.



Cependant, cette situation n'est paradoxale qu'en apparence et en réalité plutôt logique : j'ai déjà souligné plusieurs fois que *l'audiophilie*, en plaçant systématiquement l'auditeur *au centre*, considérait sa *subjectivité comme un critère absolu* (y compris, souvent, au dessus des mesures scientifiques objectives). Ce qui revient à dire que **le rapport à la réalité sonore** n'est pas sa préoccupation première.

On perçoit alors la concordance évidente qui s'établit entre la pratique de la *stéréophonie d'intensité* telle que je viens de la décrire, *artificialité comprise*, et la philosophie globale de *l'audiophilie*. Ceci se concrétisant sans ambiguïté dans la déclaration d'un membre très célèbre et éminent de cette communauté, qui affirmait qu'en terme de *qualité sonore*, il préférerait l'écoute d'un enregistrement dans son salon à celle du concert *acoustique naturel*.

Ce dernier point souligne l'évidence qu'il y a à indiquer, comme je l'ai fait à maintes reprises, et malgré certaines dénégations véhémentes mais infondées, à quel point on peut considérer *l'audiophile* comme un acteur bienveillant de la *société de consommation* et de *l'industrie culturelle*.

Impacts sur la reproduction sonore

Disons le tout net : un enregistrement en stéréo d'intensité, en particulier ceux qui sont d'origine multi pistes, est **plus facile** à reproduire correctement qu'un enregistrement de phase. Dis autrement, cela signifie que lorsqu'on se trouve en situation d'évaluer la *fidélité* de reproduction d'un système, il constitue un outil moins révélateur, donc moins *pertinent*.

Il y a plusieurs raisons à cela :

- Il est beaucoup plus éloigné de la réalité qu'un enregistrement de phase. Bien sûr, *l'artificialité* n'est pas un problème en soi - les musiciens créent les sons qu'ils veulent comme ils veulent - mais il demeure plus délicat de discourir du goût *véritable* des haricots, quand on ne déguste que ceux en boîte. ;-)
- Il contient moins d'informations spatiales.
- Il contient moins d'informations de bas niveau.

Je ne reviens pas sur les explications concernant les deux premiers points, amplement développés. Mais j'aimerais m'arrêter sur le troisième, parce qu'il n'est pas évident de prime abord.

- Premièrement, ce type d'enregistrement contient très peu d'informations relatives à *l'acoustique environnante*, soit parce qu'elle est pratiquement inexistante (studio), soit parce qu'elle est susceptible d'être une perturbation : capter l'acoustique, c'est aussi

favoriser la fameuse *repisse* dont j'ai déjà parlé. Enfin, elle doit parfois être soigneusement évitée parce que mauvaise, ou parce qu'on ne souhaite simplement pas l'entendre, ce qui peut se produire dans le cas d'enregistrements *live*. Or, justement, les informations relatives à l'acoustique environnante sont de bas niveau...

- Deuxièmement, comme je l'ai expliqué, le signal, une fois enregistré, va être modifié et va traverser souvent une grande quantité de [périphériques et d'effets divers](#) en perdant, à chaque fois, quelques micro informations et un peu de subtilité.

Certes, c'est moins vrai maintenant que les effets sont presque tous *informatiques* et que le signal peut rester numérique tout au long du traitement, sans subir de pertes. Néanmoins, chez certains, la mode des effets *vintage* conduit le signal à continuer à traverser des étages analogiques, parfois nombreux et à lampes. Avec les conséquences que l'on sait.

- Troisièmement, l'usage de micros directifs placés à proximité immédiate des instruments (toujours la *repisse*) peut aboutir à limiter les informations recueillies.

Un bon exemple est le cas de la famille des *Bois*. Si placer un micro à 10 cm du pavillon d'un cuivre n'a pas d'inconvénient, ce n'est pas du tout le cas d'une clarinette par exemple, puisque c'est l'ensemble du corps de l'instrument qui rayonne, en fonction de l'ouverture des clefs.

Pourtant, je ne compte plus les fois où j'ai vu capter un saxophone ténor avec un micro cardioïde presque enfoncé dans le pavillon. Négliger le mode de propagation acoustique de l'instrument conduit à perdre une partie des informations...et à déformer le timbre.

A noter, au passage, que les micros directifs (cardioïdes, hyper cardioïdes, etc.) souffrent particulièrement de l'*effet de proximité* : lorsqu'ils se trouvent trop près de la source sonore, ils augmentent alors nettement leur niveau dans le grave.

CHARACTERISTIC	OMNI-DIRECTIONAL	CARDIOID	SUPER-CARDIOID	HYPER-CARDIOID	BI-DIRECTIONAL
POLAR RESPONSE PATTERN					
COVERAGE ANGLE	360°	131°	115°	105°	90°
ANGLE OF MAXIMUM REJECTION (null angle)	—	180°	126°	110°	90°
REAR REJECTION (relative to front)	0	∞	12 dB	6 dB	0
AMBIENT SOUND SENSITIVITY (relative to omni)	100%	33%	27%	25%	33%
DISTANCE FACTOR (relative to omni)	1	1.7	1.9	2	1.7

Microphone polar patterns compared

Si on raisonne sans tenir compte des contraintes pratiques que j'ai décrites précédemment, l'*enregistrement de phase* est globalement plus *fidèle*. Ce qui explique qu'il soit presque exclusivement réservé à certains domaines de la musique dite *classique*, où l'objectif n'est pas de *fabriquer* un son original, mais de respecter le mieux possible celui produit naturellement par les musiciens*.

C'est pourquoi, comme elle contient plus d'informations à retranscrire correctement, cette technique impose aux transducteurs d'être eux même plus *fidèles*, y compris sur le plan de la linéarité de l'**amplitude, et surtout, de la phase**. Enfin, comme l'enregistrement contient des informations relatives aux dimensions de la source sonore, *les enceintes n'ont plus à compenser un déficit à ce niveau en fabricant elles-mêmes une image ultra focalisée*, celle-ci peut imposer ses propres caractéristiques, qu'elle soit ponctuelle ou large, et les transducteurs doivent avoir en conséquence une propagation moins directive... ;-)

*En matière d'enregistrement, le cas du Jazz est assez particulier et il justifierait amplement une étude qui lui soit entièrement consacrée. Disons, pour faire scandaleusement court, que l'album [In a silent way](#) (1969) de

Miles Davis constitue un point de bascule de ce point de vue, en passant de l'approche habituelle type [Rudy Van Gelder](#), avec sur-signalisation du rapport *intime* tel que je l'ai décrit précédemment, vers un usage créatif du studio. Une intéressante illustration de ce à quoi cette double tradition a abouti, est constituée par le double album du pianiste Laurent De Wilde [Organics / The present](#) qui propose les deux aspects.

Pour conclure, provisoirement.

Avec ce document, j'aimerais initier une réflexion plus globale, s'inscrivant dans les nécessités écologiques de notre temps, en proposant ce que j'appelle la **décroissance sonore**.

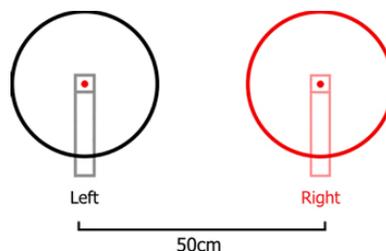
Vaste programme, cohérent avec ce qu'est ma démarche depuis longtemps : il s'agit à la fois de valoriser la *rationalité* dans les pratiques de l'audio (*convaincre* plutôt que *séduire*), de rendre le **pouvoir sur le son aux musiciens**, mais aussi d'une démarche qui entend interroger en permanence nos pratiques afin de les rendre plus respectueuses de la Musique.

C'est tout particulièrement important à propos des concerts ou des enregistrements, où un certain nombre d'acteurs (producteurs, techniciens, sonorisateurs, etc.) ont finalement dessaisi les musiciens du contrôle du son et installé des habitudes discutables voire absurdes.

Par exemple, un violoniste peut jouer en solo les [24 caprices de Paganini](#) devant 800 personnes et tout le monde l'entendra parfaitement. Par conséquent, il n'y a aucune **raison** de voir le public d'un groupe de musique électrifiée, quelle qu'elle soit, terrassé par une absurde débauche de décibels, doublée d'une qualité sonore scandaleuse.

Il ne s'agit pas là d'un simple vœu pieux. Une expérience extrêmement satisfaisante a été menée récemment avec le groupe [Guacamole Green](#) lors d'un concert, qui a montré tous les avantages que musiciens et public peuvent tirer d'une telle approche, résolument innovante.

Certes, ce n'est pas transposable dans un stade. Mais qu'est-ce qui justifie d'utiliser un stade pour un concert, le pire endroit possible, si ce n'est la seule motivation commerciale...



Enfin, voici un petit cadeau destiné à tous les lecteurs, afin d'utiliser à bon escient les performances de la *stéréophonie de phase* pour évaluer la qualité de l'image spatiale d'un système. Il suffit de télécharger le fichier Flac en suivant le lien ci-dessous, puis de cliquer avec le bouton droit de la souris pour l'enregistrer sur votre ordinateur :

[Signal de test](#)

J'ai utilisé un couple stéréophonique de type AB. Les deux micros *omnidirectionnels* ont été espacés de 50 cm et placés à environ 1m60 du sol sur un pied. Le lieu choisi est un simple préau d'école en bois, maçonné à mi hauteur sur trois côtés, et suffisamment vaste pour assurer une réverbération naturelle nettement perceptible.

Je me suis ensuite placé assez loin, à une distance d'environ 6m du couple de micros, complètement décalé vers la droite, approximativement à 60° par rapport à l'axe de celui-ci.

Le signal est constitué de simples claquements de mains qui doivent être reproduits à un volume sonore raisonnable, comme pour une écoute. Cet enregistrement n'a évidemment aucun intérêt sur le plan musical, ce n'est pas l'objectif. Néanmoins, l'espace est très correctement capté, en particulier la profondeur (ainsi que quelques bruits...).

Le test est très facile à utiliser : lors de la reproduction, si on a l'impression que le son **sort directement de l'enceinte droite**, c'est que l'image spatiale du système est médiocre.

Si, par contre, on a la sensation nette **que l'enceinte disparaît**, comme si elle s'était dématérialisée, et que la source est reproduite à la bonne distance **à travers elle**, alors, bingo, c'est que l'image spatiale est bonne !

Bon amusement !

Musicalement vôtre,
Franck Mounier