



Comparaison théorique et expérimentale de la pression générée par une matrice de haut-parleurs actionnés de façon analogique et digitale

R. Dejaeger^a, Q. Leclere^a, F. Casset^b, F. Ablitzer^c et S. Fanget^b

^aLaboratoire Vibrations Acoustique, INSA de Lyon Bâtiment St. Exupéry 25 bis av. Jean Capelle, 69621 Villeurbanne, France

^bCEA, LETI, MINATEC Campus, 17 rue des Martyrs, 38054 Grenoble, France

^cLAUM - Université du Maine, avenue Olivier Messiaen, 72085 Le Mans Cedex 9, France
remy.dejaeger@gmail.com

Le but du papier est de comparer la pression générée par une matrice composée de transducteurs électroacoustiques MEMS actionnés de façon analogique puis digitale. On montrera alors que pour qu'il y ait un intérêt à jouer en digital du point de vue acoustique, les transducteurs doivent fonctionner de manière non linéaire. Dans une première partie, la théorie de rayonnement d'un Haut-Parleur Digital Matriciel (HPDM) idéal sera comparée au rayonnement d'un Haut-Parleur Analogique (HPA) de même surface. La deuxième partie présentera le principe de fonctionnement des transducteurs électroacoustiques utilisés pour réaliser un démonstrateur, ainsi que leurs dimensions et l'empilement technologique MEMS associé. La troisième partie présentera les résultats issus d'enregistrements d'un même son joué en analogique puis en digital. On conclura en comparant la théorie et les expérimentations, puis les perspectives à court et long termes sur le sujet seront évoquées.

1 Introduction

Le Haut-Parleur Digital Matriciel (HPDM) est un moyen de transduction électroacoustique innovant qui est composé de plusieurs éléments rayonnants identiques, disposés en matrice [1]. Ces éléments, désignés par le terme *speaklet* lorsqu'ils sont de tailles réduites [2], sont capables d'émettre de brèves impulsions de pression. Le HPDM reçoit alors directement comme consigne d'entrée le signal électrique numérique. A chaque période d'horloge, un nombre de speaklets proportionnel à l'amplitude du signal est activé. La conversion du numérique vers l'analogique se fait alors directement dans l'air, grâce à l'additivité linéaire des impulsions de pression [3]. Il a été démontré que ce principe fonctionne d'autant mieux lorsque la taille de la matrice est faible, afin de limiter l'effet des différences de marches qui est la principale cause de distorsions [4]. La technologie MEMS apparaît donc comme étant particulièrement adaptée à la fabrication de HPDM [2, 5].

2 Rayonnement d'un HPDM idéal

Soit un HPDM constitué de N_0 speaklets tous identiques, positionnés au même point, et dont l'impulsion de pression $s(t)$ est répétable dans le temps. On cherche à reconstruire un signal $g(t) \in [-1; 1]$ avec une fréquence de reconstruction f_r (période T_r). En faisant l'hypothèse que chaque speaklet peut émettre une impulsion positive ou négative, la pression totale générée par le HPDM peut s'écrire sous la forme de l'équation 1.

$$P(t) = \sum_{i=0}^{E(t/T_r)} \text{round}(N_0 \cdot g(i \cdot T_r)) s(t - i \cdot T_r) \quad (1)$$

E désigne la partie entière et round l'arrondi. Cette équation se reformule en faisant apparaître le produit de convolution :

$$P(t) = (\text{round}(N_0 \cdot g(t)) \cdot \text{III}_{T_r}) * s(t), \quad (2)$$

où III_{T_r} désigne un peigne de Dirac espacé de T_r . En considérant un grand nombre de speaklets et en passant dans le domaine fréquentiel, l'équation 3 est obtenue.

$$P(f) = N_0 \cdot f_r (G(f) * \text{III}_{f_r}) \cdot S(f) \quad (3)$$

On constate que la pression générée par le HPDM dépend directement des caractéristiques de l'impulsion de pression émise par les speaklets. Le spectre de la pression reconstruite est en effet proportionnel au contenu spectral $S(f)$ d'une impulsion. De plus si $f_r > 40$ kHz et si $g(t)$ n'a pas d'énergie au delà de 20 kHz, $G(f) * \text{III}_{f_r} \approx G(f)$ pour l'oreille humaine. Dans ce cas, en revenant dans le domaine temporel, la pression s'exprime selon l'équation 4.

$$P(t) = N_0 \cdot f_r \cdot g(t) * s(t) \quad (4)$$

Finalement, pour $f_r > 40$ kHz, la pression d'un HPDM idéal se calcule en faisant la convolution entre une impulsion de pression et le signal que l'on cherche à reconstruire. Ce résultat est important car cette équation 4 rappelle l'expression de la pression générée par un Haut-Parleur Analogique (HPA). En effet, cette dernière s'obtient en convoluant la réponse impulsionnelle $h(t)$ du HPA avec le signal de commande. Dans le cas où on fait jouer les N_0 speaklets en parallèle, elle s'exprime selon l'équation 5.

$$P_{HPA}(t) = N_0 \cdot g(t) * h(t) \quad (5)$$

La différence entre le HPD et le HPA se situe donc entre $h(t)$ et $s(t)$. L'impulsion de pression $s(t)$ peut être d'amplitude assez importante. Si le speaklet est par exemple prévu pour interrompre sa course sur une butée, il est alors possible d'obtenir de fortes accélérations (ou de fortes décélérations) qui sont directement proportionnelles à la pression générée. Plus généralement, le comportement des speaklets est linéaire en fonction de la tension. Il est alors possible de calculer l'impulsion de pression $s(t)$ en fonction de la réponse impulsionnelle $h(t)$ et de l'impulsion de tension $u(t)$ (voir équation 6).

$$s(t) = h(t) * u(t) \quad (6)$$

En combinant les équations 4 et 6 puis en comparant l'équation obtenue avec l'équation 5, l'équation 7 est obtenue.

$$\begin{aligned} P(t) &= N_0 \cdot f_r \cdot g(t) * h(t) * u(t) \\ &= f_r \cdot P_{HPA}(t) * u(t) \end{aligned} \quad (7)$$

En utilisant les speaklets d'un HPDM dans leur domaine linéaire, l'équation 7 montre que la pression d'un HPDM idéal est proportionnelle à la pression que générerait les speaklets pilotés de façon analogique, convoluée avec l'impulsion de tension $u(t)$. Pour effectuer une comparaison pertinente, les limitations imposées sur l'amplitude du signal $g(t)$ se répercutent sur le signal de tension $u(t)$ dont l'amplitude maximale vaut alors 1. En considérant que ce signal $u(t)$ soit une impulsion de tension rectangulaire de largeur T_r , on a alors $u(t) \approx \delta(t) \cdot T_r$, avec $\delta(t)$ la fonction de Dirac, élément neutre de la convolution dont l'aire vaut 1. En injectant cette forme de tension dans l'équation 7, on obtient alors $P(t) = P_{HPA}(t)$.

En utilisant une matrice de speaklets de façon linéaire, il n'y a donc pas d'intérêt à reconstruire le son en digital du point de vue de la qualité et du niveau sonore. Il peut par contre exister un intérêt en terme de directivité et de consommation électrique.

Revenons à l'équation 4 et passons au cas où $f_r < 40$ kHz. Le peigne de Dirac provoque alors l'apparition de pics indésirables dans le spectre audible qui déforment le signal attendu. Ainsi si $g(t)$ est un signal sinusoïdal de fréquence f_0 , les pics apparaissent pour les fréquences f_n données par l'équation 8.

$$f_n = |f_0 - n.f_r|; n \in \mathbb{Z} \quad (8)$$

Cette équation 8 traduit le phénomène de repliement spectral lorsque le théorème de Shannon n'est pas respecté.

3 Présentation du prototype de HPDM

3.1 Principe de fonctionnement

Le prototype de HPDM MEMS étudié est composé de speaklets dont le fonctionnement repose sur l'effet piézoélectrique et sur l'effet bilame. Parmi les matériaux piézoélectriques existants, le Titano-Zirconate de Plomb (PZT) a été choisi car il présente des coefficients piézoélectriques élevés et produit donc de plus grands déplacements à champ électrique égal. Les speaklets sont composés d'une membrane circulaire de 800 μm de rayon encadrée à sa périphérie, qui joue le rôle d'élément mobile, sur laquelle est déposée la couche active en PZT. Lorsque le PZT se contracte dans le plan du substrat sous l'effet d'un champ électrique externe, la membrane fléchit par effet bilame. En réalité, le terme *membrane* est un abus de langage puisque le comportement mécanique de cet élément mobile est beaucoup plus proche de celui d'une plaque circulaire dont la fréquence propre vaut 40 kHz.

Il a été choisi de fabriquer un HPDM composé de speaklets possédant 3 états possibles à chaque période d'horloge afin d'augmenter la résolution : l'absence d'impulsion, la génération d'une impulsion de pression positive et la génération d'une impulsion de pression négative. Etant donné la nature ferroélectrique du PZT, il est donc nécessaire de disposer de 2 actionneurs piézoélectriques par membrane pour obtenir des moments induits sur la ligne neutre de sens opposés [7]. Le premier actionneur en forme de disque permet d'obtenir une déflexion vers le bas et l'actionneur externe en forme d'anneau permet d'obtenir une déflexion vers le haut. L'effet bilame associé aux 2 types d'actionneurs est illustré figure 1.

En fonction de la tension appliquée, ces speaklets se déforment statiquement de façon non-linéaire mais possèdent un comportement dynamique linéaire (au moins jusqu'à une tension de 15 V), ce qui permet de les caractériser à l'aide de fonctions de transfert.

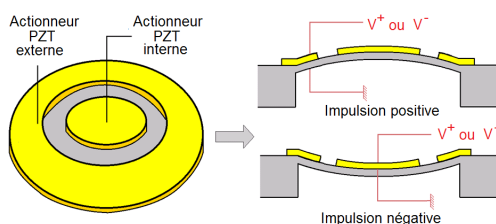


FIGURE 1 – Principe de fonctionnement des speaklets MEMS.

3.2 Empilement technologique

L'empilement technologique peut être résumé par une membrane en polysilicium de 4 μm d'épaisseur sur laquelle est déposé le matériau piézoélectrique en étapes successives selon la méthode Sol-Gel pour obtenir une épaisseur de 0.54 μm . Deux électrodes inférieure et supérieure, respectivement en platine et en ruthénium/or, sont utilisées pour appliquer un champ électrique. La membrane est libérée en fin de fabrication par une gravure profonde réalisée en face arrière.

3.3 Puce MEMS et carte de pilotage

Pour piloter le HPDM, un prototype de carte électronique a été fabriqué. La carte mère est un FPGA qui contient un support à levier dans lequel s'insère un PCB. Sur ce PCB, le HPDM de 28x28 mm² est collé et du bonding est réalisé avec des micro-fils d'or pour relier les pads de la puce MEMS et ceux du PCB. L'électronique de pilotage ainsi que le HPDM composé de 64 speaklets sont visibles figure 2. Les fils d'or ont été protégés par une résine.

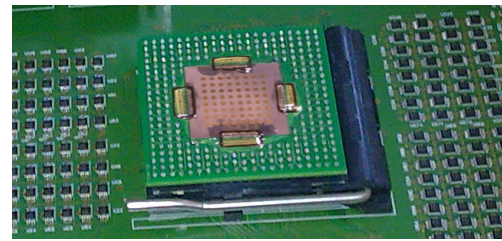


FIGURE 2 – Photographie du prototype de carte électronique de pilotage pour HPDM MEMS.

4 Expérimentations

4.1 Mesure d'uniformité des speaklets

Après la taille de la matrice, le second facteur ayant le plus d'impact négatif sur les performances d'un HPDM sont les non-uniformités entre les réponses fréquentielles des transducteurs électroacoustiques élémentaires [6]. L'uniformité des speaklets de notre prototype a donc fait l'objet d'une mesure. Pour cela, chacun des 64 speaklets a été successivement alimenté en tension par un bruit blanc gaussien et un microphone a été placé à 5 cm en face du HPDM. Les enregistrements, obtenus en activant les actionneurs internes puis externes, sont visibles figure 3. On constate alors un manque d'uniformité entre les speaklets : une sorte de modulation apparaît, mise en évidence sur la figure par 2 courbes en pointillé. Cette modulation est liée à la longueur des pistes d'accès aux actionneurs, qui sont planaires et qui se fauilte entre les speaklets. Les largeurs et les épaisseurs de ces pistes sont en effet constantes. Par conséquent, leurs résistances électriques varient en fonction de leurs longueurs. Plus la résistance des pistes est grande, plus l'amplitude de la tension électrique entre les électrodes est diminuée par rapport à celle de la tension appliquée au niveau des pads. C'est ce qui explique les différences d'amplitudes des sons émis par chacun des speaklets. Néanmoins certains speaklets semblent ne pas du tout fonctionner. La cause exacte n'est pas connue mais

deux possibilités principales ont été identifiées. La première concerne les fils de bonding, qui peuvent se décoller ou se casser. La seconde concerne des défaillances au niveau du PZT. De récentes expériences ont en effet montrées qu'en dessous d'une épaisseur de 2 μm de PZT, des décollements à l'interface avec le ruthénium peuvent se produire.

Afin de quantifier les non-uniformités entre les speaklets en état de marche, les fonctions de transfert *vitesse/tension* du centre de plusieurs speaklets ont été mesurées. Jusqu'au premier mode de résonance, les speaklets se comportent en effet comme des systèmes à un degré de liberté. Une étude statistique a donc été menée sur les mesures afin d'obtenir les écarts-types relatifs des paramètres mécaniques suivants : la fréquence propre, le taux d'amortissement et l'amplitude de vibration. L'écart-type relatif, sans unité, permet de comparer des séries de valeurs exprimées dans des unités différentes. Les chiffres obtenus ont alors confirmés que la plus grande source de dispersion entre les speaklets se situe au niveau de l'amplitude de vibration.

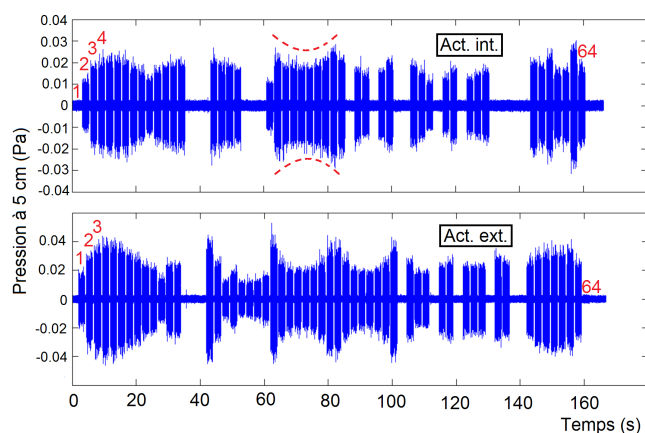


FIGURE 3 – Enregistrements des pressions générées par chaque speaklet de la matrice soumis à un bruit blanc gaussien.

4.2 Enregistrement en pilotage analogique

En pilotage analogique, un signal alternatif d'amplitude 7.5 V ainsi qu'un signal continu de même amplitude sont envoyés sur un seul type d'actionneur, afin de rester dans une gamme de tension positive. Un signal cible constitué des fréquences 3 et 7 kHz a alors été joué de façon analogique. La figure 4 présente le spectre des enregistrements en utilisant les actionneurs internes puis externes de tous les speaklets. Le microphone est placé à 8 cm en face de la matrice. Les fondamentaux ainsi que les harmoniques et les intermodulations (de rang 2 uniquement) sont repérés. Le pic obtenu vers 5800 Hz doit provenir du premier mode de résonance de la puce en silicium de 725 μm d'épaisseur. En effet, en utilisant des conditions aux limites libres aux extrémités de la puce (qui est collée par le dessous), la théorie [8] prédit l'apparition du premier mode aux alentours de cette fréquence.

En utilisant les actionneurs internes, les amplitudes des 2 fréquences d'intérêt 3 et 7 kHz sont respectivement de 61.2 et 64.5 dB SPL en effectuant une intégrale sur la largeur des pics. L'utilisation des actionneurs externes donne respectivement 52.1 et 75.8 dB SPL. Ceci indique que les deux types d'actionneurs ne mettent pas en mouvement

les membranes de façon parfaitement identiques : les actionneurs externes semblent avoir plus de difficultés à générer les basses fréquences que les actionneurs internes. Ce manque de symétrie dans le comportement des 2 types d'actionneurs sera une source de distortions supplémentaire lors de la reconstruction digitale.

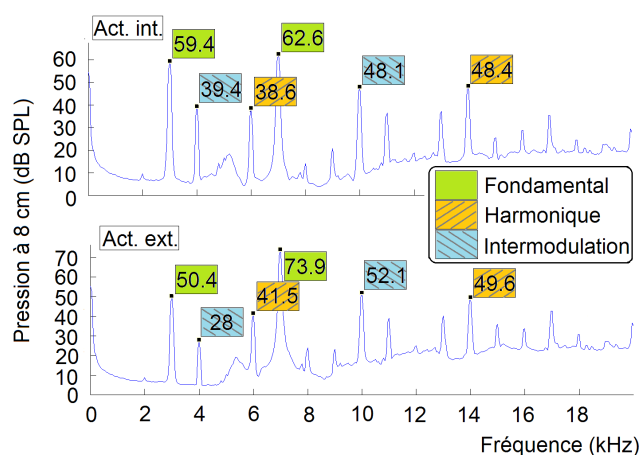


FIGURE 4 – Spectres des enregistrements du son joué en analogique.

4.3 Enregistrement en pilotage digital

La figure 5 présente le spectre de l'enregistrement du même signal cible joué en digital. Pour décorréler l'erreur de quantification, un signal aléatoire de densité de probabilité gaussienne, appelé signal de dither, a été appliqué avant le calcul du nombre de speaklets à activer. L'impulsion de tension utilisée est théoriquement rectangulaire, d'une amplitude de 15 V et d'une durée de 22.68 μs . Elle possède en réalité des temps de montée et de descente de 2 μs . Le microphone est de nouveau placé à 8 cm en face de la matrice et la fréquence d'horloge est fixée à 44.1 kHz. La stratégie de pilotage n'est pas optimisée et les speaklets sont activés de gauche à droite et de bas en haut.

Les amplitudes des deux fréquences d'intérêt mesurées sont respectivement 64 et 77.2 dB SPL. Un faible gain d'amplitude a donc été obtenu aux fréquences d'intérêt par rapport aux meilleurs résultats obtenus en analogique mais on constate également plus de bruit. Les harmoniques et intermodulations sont toujours présents mais plus difficiles à repérer en haute fréquence justement à cause de ce bruit. En gardant à l'esprit que l'analogique possède une dynamique 2 fois plus petite que le digital à cause de l'utilisation d'un seul type d'actionneur, on conclut donc que les 2 approches fournissent des niveaux de pression semblables.

4.4 Indicateurs de qualité

Le THD+N, pour *Total Harmonic Distortion + Noise*, est un indicateur de qualité du son, exprimé en pourcentage. Il est défini comme le ratio de la puissance de tous les harmoniques, intermodulations et bruits présents dans une certaine gamme de fréquences sur la puissance du fondamental.

En pilotage analogique, les THD+N valent 27 et 12 % entre 500 et 20000 Hz, en utilisant respectivement les actionneurs internes et externes. En pilotage digital, cet

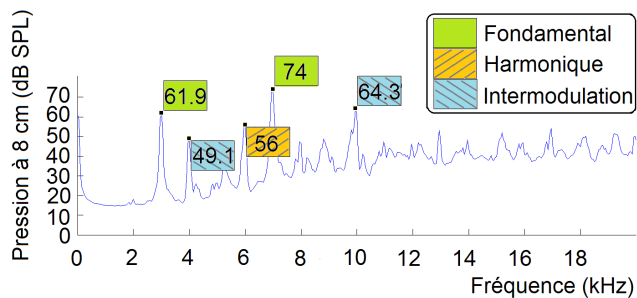


FIGURE 5 – Spectre de l'enregistrement du son joué en digital.

indicateur monte à 43 % sur la même gamme de fréquence. Ceci s'explique par le manque d'uniformité entre les comportements des speaklets, le manque de symétrie entre les impulsions positives et négatives, le nombre de speaklets relativement faible et la stratégie de pilotage non optimisée.

5 Conclusions et perspectives

Les expérimentations, en accord avec la théorie, confirment qu'il n'y a pas de gain d'amplitude en digital par rapport à l'analogique si les speaklets sont utilisés dans leur domaine linéaire. Les THD+N peuvent cependant être théoriquement très proches lorsque suffisamment de speaklets sont utilisés en pilotage digital. Dans ce dernier cas, une augmentation du THD+N a été expérimentalement constaté, à cause de 3 points faibles principaux :

- Les résistances d'accès des pistes en platine menant aux actionneurs peuvent varier de 200 à 800 Ω , ce qui fait varier les amplitudes des tensions appliquées sur les électrodes.
- Un manque de fiabilité des speaklets conduit à une matrice comportant environ 20 % d'actionneurs défectueux.
- Un manque de symétrie existe entre les impulsions positives et négatives.

Les perspectives à court terme visent donc à régler les trois points faibles évoqués. Ainsi, en baissant et en égalisant toutes les résistances d'accès à 20 Ω , les problèmes de non-uniformité entre les speaklets seront largement atténués. Pour cela, une solution utilisant une électrode inférieure en platine non gravée et donc commune à tous les speaklets d'une matrice sera utilisée. Les pistes d'accès en or reliées aux électrodes supérieures verront quant à elles leurs largeurs ajustées en fonction de leurs longueurs. Afin d'augmenter la fiabilité des speaklets, une épaisseur de 2.1 μm de PZT sera de plus utilisée. Enfin, les dimensions des actionneurs internes et externes seront optimisés afin d'obtenir un comportement plus symétrique. En augmentant alors le nombre de speaklets de 64 à 256 sur une puce de 40x40 mm^2 , la qualité du son joué en digital devrait alors être assez proche du son joué en analogique. La diminution des résistances d'accès R_s et de la capacité C_s des actionneurs en PZT conduira également à une baisse de la consommation électrique. Le temps caractéristique $\tau = R_s \cdot C_s$ étant énormément diminué, il sera de plus intéressant de tester des méthodes de charges adiabatiques par l'utilisation d'impulsions de tension trapézoïdales afin

de diminuer davantage la consommation par rapport à l'utilisation d'impulsions de tension proches d'une forme rectangulaire. En effet, la méthode de charge adiabatique consiste à délivrer un flux de charges constant et adapté à la constante de temps du circuit afin de limiter les pertes par effet Joule [11]. Si le temps de montée T_m de la tension est exprimé en fonction de la constante de temps et d'un coefficient multiplicateur α selon $T_m = \alpha \cdot \tau$, la méthode de la charge adiabatique devient intéressante à partir de $\alpha > 2$. L'équation 9 indique le pourcentage d'énergie économisée en charge adiabatique par rapport à l'utilisation d'un échelon de tension, en fonction de α . Théoriquement, avec α suffisamment grand, les pertes par effet Joule deviennent négligeables.

$$y(\%) = \frac{\alpha - 2}{\alpha} 100 \quad (9)$$

Le cas de la décharge adiabatique est un peu différent. En effet, l'énergie dissipée dans le cas de la décharge est égale à celle dissipée dans le cas de la charge seulement si le circuit est prévu pour recycler l'énergie en utilisant par exemple un type de générateur particulier capable de stocker les charges électriques retournées [9, 10]. Si une telle source est utilisée et qu'on suppose qu'elle stocke et réutilise l'énergie sans perte, le pourcentage d'énergie économisée en un cycle de charge et décharge adiabatique par rapport à une impulsion rectangulaire reste celui qui est décrit par l'équation 9. Avec le futur prototype et en utilisant des impulsions de tension trapézoïdales de durée totale $T=22.68 \mu\text{s}$ divisé en 3 parties égales de 7.6 μs , le paramètre α vaudra plus de 250, ce qui conduira à un gain en consommation électrique de plus de 99%. Il sera ainsi théoriquement possible d'obtenir un HPDM très basse consommation tout à fait adapté aux appareils nomades tels que les téléphones portables.

Pour obtenir en plus une augmentation du niveau de pression, les perspectives à plus long terme doivent ensuite se concentrer sur la conception d'un speaklet au comportement intrinsèquement non-linéaire produisant des impulsions de pression de fortes amplitudes. Le fonctionnement de ces speaklets pourrait par exemple être basé sur un système de butées. Ceci implique l'étude d'un nouvel empilement technologique MEMS, certainement assez complexe. Il serait alors très intéressant de modéliser et de maîtriser l'effet de la butée, afin de jouer sur la forme et le contenu fréquentiel de l'impulsion de pression générée par un speaklet. Pour obtenir une forte densité de speaklets, l'utilisation de pistes 3D traversant le substrat serait bénéfique. Avec cette technique, de faibles résistances d'accès seraient également obtenues étant donné la faible épaisseur du substrat. Enfin, la dernière étape consisterait à intégrer l'électronique de pilotage au processus de fabrication en utilisant la technologie CMOS.

Remerciements

Les auteurs remercient Brigitte Desloges, du CEA-LETI, pour le suivi de la procédure de fabrication MEMS ainsi que Christophe Leblanc et Nicolas Delorme, de la société Asygn, pour la conception et la fabrication du FPGA.

Ce travail a été réalisé au sein du Labex CeLyA de l'Université de Lyon, géré par l'Agence Nationale de la Recherche (ANR-10-LABX-0060/ ANR-11-IDEX-0007)

Références

- [1] W. E. Stinger, Direct digital loudspeaker, *US Patent 4,515,997*, (1985).
- [2] B. M. Diamond, J. J. Neumann, K. J. Gabriel, Digital sound reconstruction using arrays of CMOS-MEMS microspeakers, *The Fifteenth IEEE International Conference on Micro Electro Mechanical Systems*, 292-295 (2002).
- [3] J. L. Flanagan, Direct acoustic digital-to-analog conversion with digital transducer array loudspeakers, *The Bell System Technical Journal*, 1693-1719 (1980).
- [4] Y. Huang, S. C. Busbridge, D. S. Gill, Distortion and directivity in a digital transducer array loudspeaker, *Journal of the Audio Engineering Society*, 337-352 (2001).
- [5] Y. Cohen, Digital Loudspeakers, *Journée Acoustique et Microsystèmes du Mans (JAM)*, (2011).
- [6] J. Mendoza-Lopez, S. C. Busbridge, P. A. Fryer, Direct acoustic digital-to-analog conversion with digital transducer array loudspeakers, *Journal of the Audio Engineering Society* **55**, 472-502 (2007).
- [7] T. Bakke, A. Vogl, O. Zero, F. Tyholdt, I-R. Johansen, D. Wang, A novel ultra-planar, long-stroke and low-voltage piezoelectric micromirror, *Journal of Micromechanics and Microengineering* **20**, 1-7 (2010).
- [8] R. D. Blevins, *Formulas for Natural Frequency and Mode Shapes*, Krieger Publishing Company (1979).
- [9] M. Arsalan, M. Shams, Charge-recovery power clock generators for adiabatic logic circuits, *Proceedings of the IEEE International Conference on VLSI Design*, 171-174 (2005).
- [10] J. G. Koller, W. C. Athas, Adiabatic switching, low energy computing, and the physics of storing and erasing information, *Proceedings of the Workshop on Physics and Computation*, 267-270 (1992).
- [11] W. C. Athas, L. Svensson, J. G. Koller, N. Tzartzanis, E. Y-C. Chou, Adiabatic switching, Low-power digital systems based on adiabatic-switching principles, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 398-407 (1994).