



Contributions au traitement des signaux à valeurs sur des structures algébriques non-commutatives

Nicolas Le Bihan

► To cite this version:

Nicolas Le Bihan. Contributions au traitement des signaux à valeurs sur des structures algébriques non-commutatives. Traitement du signal et de l'image [eess.SP]. Université de Grenoble, 2011. <tel-00606665>

HAL Id: tel-00606665

<https://theses.hal.science/tel-00606665v1>

Submitted on 7 Jul 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

UNIVERSITÉ DE GRENOBLE
ÉCOLE DOCTORALE EEATS
SIGNAL, IMAGE, PAROLE ET TÉLÉCOMMUNICATOINS

HABILITATION À DIRIGER LES RECHERCHES

Présentée et soutenue par
Nicolas LE BIHAN

Contributions au traitement des signaux à valeurs sur des structures algébriques non-commutatives

préparée au GIPSA-Lab, Grenoble
soutenue le 20 Juin 2011

Jury :

<i>Rapporteurs :</i>	Alfred O. HERO III	-	University of Michigan, Ann Arbor, USA
	Éric MOULINES	-	Télécom ParisTech (ENST), Paris
	Philippe RÉFRÉGIER	-	Institut Fresnel, Marseille
<i>Président :</i>	Olivier J.J. MICHEL	-	Grenoble INP, Grenoble
<i>Examineur :</i>	Bernard CASTAING	-	ENS, Lyon

Table des matières

I	Quaternions et traitement du signal	3
1	Quaternions $\mathbb{H}$	5
1.1	Historique	6
1.2	Définitions et propriétés	7
1.2.1	Théorèmes de Frobenius et Hurwitz	7
1.2.2	Représentations	8
1.2.3	Géométrie	9
1.2.4	$\mathbb{H}$ et les groupes $SU(2)$ et $SO(3)$	10
1.2.5	Quaternions complexes, octonions et algèbres de Clifford	10
1.3	Matrices de quaternions	11
1.4	Transformation de Fourier quaternionique	12
1.5	Circularité	15
1.6	Conclusion	16
1.7	Publications annexées en lien avec ce chapitre	17
1.7.1	"Fundamental representation and algebraic properties of biquaternions or complexified quaternions" AACA 2010	18
1.7.2	"Fast Complexified Quaternion Fourier Transform" IEEE TSP 2008	48
1.7.3	"On Properness Of Quaternion Random Variables" IMA 2004	58
2	Signaux quaternioniques	63
2.1	Signaux quaternioniques et polarisation	63
2.2	Méthodes par sous-espaces et traitement d'antenne	65
2.3	Signal hyperanalytique	70
2.4	Conclusion	73
2.5	Publications annexées en lien avec ce chapitre	74
2.5.1	"MUSIC algorithm for vector-sensors array using biquaternions" IEEE TSP 2007	75
2.5.2	"The $\mathbb{H}$ -analytic signal" EUSIPCO 2008	86
2.5.3	"Quaternionic independent component analysis using hypercomplex nonlinearities" IMA 2006	91
II	Signaux à valeurs sur $SO(3)$ et $\mathcal{S}^2$	97
3	Signaux à valeurs sur $SO(3)$ et $\mathcal{S}^2$	99
3.1	Polarisation, sphère de Poincaré et $SO(3)$	99
3.2	Variables aléatoires sur $SO(3)$	101
3.2.1	Théorème de Peter-Weyl	102
3.2.2	Fonctions caractéristiques	103
3.3	Processus de Lévy sur $SO(3)$	104
3.3.1	Dépolarisation	104
3.3.2	Diffusion multiple et Processus de Poisson composé sur $SO(3)$	105
3.4	Phase géométrique des ondes polarisées	109
3.4.1	Phase géométrique non-adiabatique : mise en évidence expérimentale pour les ondes élastiques	111

3.4.2	Phase géométrique des ondes élastiques en diffusion multiple	113
3.5	Conclusion	115
3.6	Publications annexées en lien avec ce chapitre	116
3.6.1	"Higher-order statistics of Stokes parameters in a random birefringent medium" Waves in random and Complex media 2008	117
3.6.2	"Decompounding on compact Lie groups" IEEE Transactions on Information Theory 2010	136
3.6.3	"Non-adiabatic geometric phase of elastic waves" Soumis à Journal of the Acoustical Society of America 2011	148
4	Travaux en cours et perspectives	155
4.1	Quaternions et signal	155
4.2	Processus de Poisson : extensions	156
4.2.1	Processus sur $SE(3)$, ondes P et S	156
4.2.2	Taches solaires	157
4.3	Processus non-holonomes	157
4.3.1	Chaînes de Markov	157
4.3.2	Phase géométriques, processus et analyse de Fourier géométrique . .	158
III	Curriculum Vitae	161
5	Curriculum Vitae	163
	Bibliographie	171

Introduction

Ce manuscrit présente les travaux de recherche que j'ai menés depuis une dizaine d'années. Il se compose de deux parties distinctes qui possèdent néanmoins un point commun : la non-commutativité. Les signaux qui sont rencontrés dans ce manuscrit ont tous en commun la spécificité de prendre leurs valeurs sur des structures algébriques non-commutatives : quaternions, biquaternions, groupes de Lie (groupes matriciels, groupe des rotations), variétés différentiables.

Dans une première partie, je présente les outils de traitement du signal quaternionique qui vont du traitement d'antenne, à la transformation de Fourier, en passant par l'extension de la notion de circularité. Il est montré comment l'extension quaternionique de ces concepts et techniques permet de proposer de nouveaux algorithmes de traitement des signaux à valeurs quaternioniques ou complexes. La mise au point de ces algorithmes a nécessité le développement d'outils théoriques nouveaux que nous présentons dans le chapitre 1 : décomposition matricielle, circularité des variables aléatoires quaternioniques, représentations polaires des quaternions. Dans le chapitre 2, il est montré comment on en arrive à considérer des signaux quaternioniques : soit par une modélisation particulière de signaux issus de capteurs vectoriels, soit par transformation de signaux à valeurs complexes. Nous illustrons quelques-uns des résultats obtenus dans le domaine du traitement des signaux quaternioniques, en particulier en traitement d'antenne sismologique et pour l'extension du signal analytique dans le chapitre 2.

Dans la seconde partie, nous nous intéressons aux signaux qui prennent leurs valeurs sur la sphère unité dans $\mathbb{R}^3$, *i.e.* S^2 , et le groupe des rotations dans l'espace, *i.e.* $SO(3)$. Les signaux considérés dans cette partie sont les signaux polarisés, et en particulier les signaux liés aux ondes élastiques. Nous proposons une nouvelle approche pour l'étude de la propagation de ces signaux dans les milieux aléatoires fondée sur les processus de Lévy à valeurs sur le groupe des rotations. Ce modèle nous permet d'étudier les phénomènes de dépolarisation des ondes dans les fibres optiques ou les milieux hétérogènes. L'originalité de l'approche proposée permet également un éclairage nouveau sur l'étude de la polarisation, en particulier par une nouvelle définition du degré de polarisation, dite d'ordre supérieur.

Enfin, je présente mes travaux récents sur l'étude de la phase géométrique des ondes élastiques, phénomène dont nous avons rapporté l'observation récemment. Je présente également un modèle statistique qui prédit l'apparition de cette phase lors de la propagation des ondes polarisées dans les milieux aléatoires.

Les travaux présentés dans ce manuscrit sont donc de natures diverses, mais ont tous en commun de considérer des signaux et processus dont les échantillons sont à valeurs sur des structures algébriques non-commutatives. La spécificité de ces signaux fait que leur étude n'est pas très répandue, souvent dispersée, mais leurs champs d'application sont de plus en plus importants. Je présente principalement des techniques de traitement du signal polarisé (sans doute par affinité personnelle), mais les algorithmes qui sont exposés ont un spectre potentiel d'applications plus large que les seuls signaux polarisés.

Première partie

Quaternions et traitement du signal

Quaternions $\mathbb{H}$

Sommaire

1.1	Historique	6
1.2	Définitions et propriétés	7
1.2.1	Théorèmes de Frobenius et Hurwitz	7
1.2.2	Représentations	8
1.2.3	Géométrie	9
1.2.4	$\mathbb{H}$ et les groupes $SU(2)$ et $SO(3)$	10
1.2.5	Quaternions complexes, octonions et algèbres de Clifford	10
1.3	Matrices de quaternions	11
1.4	Transformation de Fourier quaternionique	12
1.5	Circularité	15
1.6	Conclusion	16
1.7	Publications annexées en lien avec ce chapitre	17
1.7.1	"Fundamental representation and algebraic properties of biquaternions or complexified quaternions" AACA 2010	18
1.7.2	"Fast Complexified Quaternion Fourier Transform" IEEE TSP 2008	48
1.7.3	"On Properness Of Quaternion Random Variables" IMA 2004	58

Ce premier chapitre présente quelques définitions sur les quaternions ainsi que quelques contributions qui seront utilisées dans les algorithmes présentés dans le chapitre 2.

Depuis longtemps en traitement du signal, les signaux à valeurs complexes ont occupé une place importante, soit pour décrire l'évolution temporelle du contenu spectral d'un signal monovarié (via le signal analytique [Gabor 1946, Ville 1948] et ses extensions [Hahn 1991, Bulow 2001, Felsberg 2001], soit récemment pour étudier des signaux bivariés [Lilly 2010]. Comme souligné régulièrement par les auteurs (en l'occurrence par B. Picinbono [Picinbono 1994]), un signal complexe n'est qu'un signal bidimensionnel, mais l'intérêt de l'étudier via sa représentation complexe réside dans les outils d'analyse complexe. Un autre avantage, qui a largement motivé mes travaux utilisant les extensions des complexes, est le fait que les nombres complexes font le lien entre algèbre et géométrie. Ce lien est résumé dans la formule d'Euler :

$$e^{i\theta} = \cos \theta + i \sin \theta$$

que Feynmann considérait comme "notre joyau" [Feynman 1963]. Les nombres complexes encodent les transformations géométriques du plan et permettent de les considérer comme des scalaires, plutôt que des matrices. Malgré cela, les signaux complexes sont rarement interprétés géométriquement, et leur étude est quasi systématiquement conduite via l'étude d'un couple de signaux [Schreier 2010, Lilly 2010].

L'idée d'utiliser les quaternions en traitement du signal est venue avec l'introduction de signaux ou images dont les échantillons sont à trois ou quatre dimensions. Dans les premières tentatives, la géométrie a été mise de côté et l'on a cherché à se soustraire au cas calcul vectoriel en passant par le formalisme des quaternions [Sangwine 1996]. De plus en plus maintenant, l'aspect géométrique est exploité en signal, ce qui rend l'utilisation des quaternions plus pertinente. Cet aspect géométrique (lien avec le groupe des rotations dans $\mathbb{R}^3$) a par ailleurs été exploité depuis longtemps en *computer graphics* [Shoemake 1985].

Un fait particulièrement remarquable est que la formule d'Euler est valable pour les quaternions. Cela place les quaternions au même rang que les complexes pour faire le lien entre algèbre et géométrie, mais cette fois-ci dans l'espace 3D ou 4D. Avant de présenter quelques aspects des quaternions, notions de base et contributions, rappelons quelques faits historiques.

1.1 Historique

Les quaternions ont été découverts en 1843, par Sir William Ronan Hamilton, Physicien, astronome et mathématicien irlandais. C'est après une dizaine d'années de tentatives d'extension des nombres complexes à l'espace 3D que Sir W.R. Hamilton réalisa qu'il fallait quatre dimensions pour construire une algèbre géométrique. Il inventa les quaternions et la fameuse relation entre les trois imaginaires purs. La légende dit qu'il grava cette équation de relation entre les trois imaginaires purs quaternioniques sur le pont de Brougham (auss appelé pont de Broom) à Dublin. Hamilton développa le calcul pour les quaternions et les biquaternions (quaternions à coefficients complexes), mais le calcul vectoriel pris l'ascendant sur les quaternions rapidement. Les quaternions ne sont pas tombés dans l'oubli, et leur utilité est reconnue dans plusieurs domaines, entre autre récemment en animation graphique [Shoemake 1985, Hanson 2006]. En traitement du signal, si l'on omet leur utilisation pour coder les rotations, ils sont apparus *via* les transformations de Fourier des images couleurs au milieu des années 90 [Sangwine 1996], et depuis ont trouvé quelques applications. Certaines d'entre elles vont être exposées dans ce manuscrit.

Le passage des réels aux complexes fait perdre l'ordonancement (impossible de dire si un complexe est plus grand ou plus petit qu'un autre). Le passage des complexes aux quaternions fait, lui, perdre la commutativité du produit. D'un point de vue purement théorique, les quaternions sont donc la structure algébrique non-commutative quasiment la plus "simple" sur laquelle on peut commencer à appréhender le traitement du signal non-commutatif. Ce manque de commutativité est une constante dans les travaux présentés dans ce manuscrit et on la retrouvera sur d'autres structures algébriques (qui sont en fait intimement liées aux quaternions) dans la seconde partie du manuscrit.

Il n'est pas vraiment possible de faire une bibliographie complète sur les quaternions. Il est préférable de se fier à quelques ouvrages de référence. De nombreux ouvrages présentent les propriétés des quaternions. Parmi les références les plus complètes, on peut citer le livre de Ward [Ward 1997], celui de Kantor [Kantor 1989] qui présente les quaternions via leur nature hypercomplexe, ou le livre de Girard [Girard 2004] plus orienté vers les formulations quaternioniques en Physique et le lien avec les algèbres de Clifford. Il existe également nombre d'ouvrage présentant le lien entre les quaternions et les groupes de rotations. Une référence assez complète sur ce sujet est le livre d'Altman [Altman 1986].

1.2 Définitions et propriétés

Les quaternions sont des nombres hypercomplexes de dimension 4. L'ensemble des quaternions est noté $\mathbb{H}$ en l'honneur de Sir W.R. Hamilton. Un quaternion $q \in \mathbb{H}$ s'écrit, dans sa forme Cartésienne, comme :

$$q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k} \quad (1.1)$$

avec $a, b, c, d \in \mathbb{R}$, et $\mathbf{i}, \mathbf{j}, \mathbf{k}$ des nombres imaginaires purs obéissant aux fameuses¹ relations :

$$\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\mathbf{j}\mathbf{k} = -1 \quad (1.2)$$

On en déduit les relations suivantes :

$$\mathbf{i}\mathbf{j} = -\mathbf{j}\mathbf{i} = \mathbf{k} \quad (1.3)$$

$$\mathbf{k}\mathbf{i} = -\mathbf{i}\mathbf{k} = \mathbf{j} \quad (1.4)$$

$$\mathbf{j}\mathbf{k} = -\mathbf{k}\mathbf{j} = \mathbf{i} \quad (1.5)$$

Le conjugué de q est $\bar{q} = a - b\mathbf{i} - c\mathbf{j} - d\mathbf{k}$ et son module est $|q| = \sqrt{q\bar{q}} = \sqrt{\bar{q}q} = (a^2 + b^2 + c^2 + d^2)^{1/2}$. Un quaternion de module égal à 1 est dit *unitaire*. Pour tout $q \neq 0$, son inverse est donné par : $q^{-1} = \bar{q}/|q|^2$. On dénomme a la partie réelle (ou scalaire) de q , notée $\mathcal{S}_q$ et $\mathcal{V}_q = q - a$ sa partie vectorielle. De manière évidente, on retrouve la partie réelle par : $\mathcal{S}_q = (q + \bar{q})/2$ et la partie vectorielle $\mathcal{V}_q = (q - \bar{q})/2$. La conjugaison n'est pas une involution sur $\mathbb{H}$, du fait que $\forall q, p$ on a $\overline{pq} = \bar{q}\bar{p}$. Il existe tout de même des involutions sur $\mathbb{H}$, elles sont de la forme $-\mu q \mu$, avec μ un quaternion *pure unitaire*, i.e. une racine de -1. Parmi toutes ces involutions, trois sont particulières, et permettent d'obtenir les composantes réelles (a, b, c, d) de q par combinaison linéaire comme suit :

$$\begin{aligned} a &= \frac{q - \mathbf{i}q\mathbf{i} - \mathbf{j}q\mathbf{j} - \mathbf{k}q\mathbf{k}}{4} & b &= \frac{q - \mathbf{i}q\mathbf{i} + \mathbf{j}q\mathbf{j} + \mathbf{k}q\mathbf{k}}{4\mathbf{i}} \\ c &= \frac{q + \mathbf{i}q\mathbf{i} - \mathbf{j}q\mathbf{j} + \mathbf{k}q\mathbf{k}}{4\mathbf{j}} & d &= \frac{q + \mathbf{i}q\mathbf{i} + \mathbf{j}q\mathbf{j} - \mathbf{k}q\mathbf{k}}{4\mathbf{k}} \end{aligned}$$

La règle d'addition est triviale pour les quaternions, et pour la multiplication, la règle est la suivante, pour $p, q \in \mathbb{H}$:

$$pq = \mathcal{S}_p\mathcal{S}_q - \mathcal{V}_p \cdot \mathcal{V}_q + \mathcal{S}_p\mathcal{V}_q + \mathcal{S}_q\mathcal{V}_p + \mathcal{V}_p \wedge \mathcal{V}_q$$

ou $\wedge$ et le produit vectoriel classique, en considérant les parties vectorielles de p et q comme des vecteurs de $\mathbb{R}^3$. La non-commutativité du produit dans $\mathbb{H}$ peut se voir ici dans la non-commutativité du produit vectoriel. En complément de ces quelques propriétés, le lecteur pourra consulter par exemple [Ward 1997] pour une liste plus complète des propriétés algébriques des quaternions.

1.2.1 Théorèmes de Frobenius et Hurwitz

Sans entrer dans les détails, nous mentionnons ici les théorèmes de Frobenius et Hurwitz². Pour une explication particulièrement claire, on consultera les chapitres 17 et 19 de [Kantor 1989]. Ces théorèmes démontrent la place particulière des quaternions (ainsi que celle des complexes et des octonions). En substance, ces deux théorèmes disent que toute

1. Ces relations auraient été gravées par Sir W.R. Hamilton au matin du 16 octobre 1843, sur le pont de Brougham à Dublin.

2. Le théorème d'Hurwitz est aussi appelé théorème de Frobenius généralisé.

algèbre normée³ de division⁴ est isomorphe à $\mathbb{R}$ (réels), $\mathbb{C}$ (complexes), $\mathbb{H}$ (quaternions) ou $\mathbb{O}$ (octonions).

Ces théorèmes motivent l'utilisation des quaternions en traitement des signaux à échantillons 3D ou 4D. En effet, il peut s'avérer intéressant de conserver des propriétés liées à la norme des échantillons qui soient proches de celles connues sur $\mathbb{R}$ et $\mathbb{C}$. De plus, l'utilisation d'autres modélisations (algèbres de Clifford, nombres hypercomplexes, bicomplexes) ne garantit pas que tout élément non-nul possède un inverse. Ceci pourrait avoir des conséquences éventuelles sur la bonne marche de certains algorithmes.

Cette argumentation est éventuellement discutable. En pratique, les conséquences sont infimes. Malgré tout, l'utilisation des quaternions pour modéliser les échantillons 3D et 4D paraît plus naturelle du fait de la position exceptionnelle de cette algèbre.

1.2.2 Représentations

En complément de la représentation Cartésienne donnée en (1.1) et de l'expression scalaire/vectorielle, il est possible de représenter un quaternion de plusieurs manières. Ces représentations autorisent pour certaines des interprétations géométriques et d'autres permettent d'appréhender le passage entre les complexes (2D) et les quaternions (4D)⁵.

Forme polaire : Une propriété remarquable des quaternions est que la formule d'Euler est toujours valide sur $\mathbb{H}$. Ainsi, tout quaternion q s'écrit :

$$q = |q| (\cos \theta + \mu_q \sin \theta) = |q| e^{\mu_q \theta}$$

avec $|q| \in \mathbb{R}^+$, μ_q un quaternion unitaire pure ($\mu_q^2 = -1$) appelé *axe* de q et $\theta \in [0, 2\pi[$ l'*angle* de q . L'existence de l'exponentielle quaternionique est due à la convergence de sa série, qui est possible sur $\mathbb{H}$ grâce à l'existence de la norme quaternionique. Il est à noter que le comportement de l'exponentielle sur $\mathbb{H}$ est différent de celui connu sur $\mathbb{R}$ ou $\mathbb{C}$. En effet, d'une manière générale, on a, pour $p, q \in \mathbb{H}$ non nuls :

$$e^{\mu_p \theta_p} e^{\mu_q \theta_q} \neq e^{\mu_p \theta_p + \mu_q \theta_q}$$

C'est une manifestation de la formule de Baker-Campbell-Hausdorff, bien connue en théorie des groupes de Lie (voir par exemple [Belinfante 1966]).

Forme de Cayley-Dickson : Une façon assez naturelle de voir les quaternions est de les considérer comme des nombres complexes dont les coefficients sont eux-même complexes, mais avec un axe imaginaire orthogonal au premier. Ainsi, on réécrit q comme :

$$q = z_1 + z_2 j = a + b i + c j + d k$$

avec $z_1 = a + i b$ et $z_2 = c + i d$. On comprend avec cette notation que z_1 et z_2 sont deux nombres complexes qui vivent dans deux plans orthogonaux de $\mathbb{R}^4$ ne s'intersectant qu'à l'origine. Cette notation est obtenue par un processus de doublement connu dans l'étude des nombres hypercomplexes [Kantor 1989]. Cette notation est largement utilisée dans les algorithmes de traitement du signal et la modélisation des signaux quaternioniques décrits dans le chapitre 2.

3. Une algèbre normée est un corps muni d'une norme $\|\cdot\|$ sur lequel on a $\|x\|\|y\| = \|xy\| \forall x, y$ appartenant à ce corps.

4. Une algèbre de division est en fait un corps. En anglais, on parle de "division algebra".

5. Cela s'avère utile lors de la généralisation de techniques/outils de traitement du signal complexe au cas quaternionique.

Forme de Cayley-Dickson polaire : Nous avons récemment introduit une nouvelle notation pour les quaternions [Sangwine 2010]. Cette notation a originalement été proposée dans un but d'interprétation du signal hyperanalytique décrit dans le paragraphe 2.3 et dans [Le Bihan 2008]. Tout quaternion q peut d'écrire :

$$q = Ae^{Bj} \quad (1.6)$$

avec $A, B \in \mathbb{C}$. A est un *module complexe* et B une *phase complexe*. Plus de détails sur cette notation ainsi que son lien avec celles précédemment citées se trouvent dans [Sangwine 2010].

Représentations matricielles : Il est possible (voir par exemple [Ward 1997]) de représenter les quaternions par certaines matrices 2×2 complexes :

$$\mathbb{H} \ni a + ib + jc + kd \sim M = \begin{bmatrix} a + ib & c + id \\ -c + id & a - ib \end{bmatrix} = \begin{bmatrix} z_1 & z_2 \\ -z_2^* & z_1^* \end{bmatrix} \in \mathcal{M}_{2 \times 2}(\mathbb{C})$$

avec z_1 et z_2 issus de la notation de Cayley-Dickson. Cette notation matricielle permet d'intuiter le lien particulier entre les quaternions unitaires (*i.e.* $q \in \mathbb{H}$ t.q. $|q| = 1$) et le groupe spécial unitaire $SU(2)$ (voir [Altman 1986]).

Représentations vectorielles : Les quaternions étant des nombres 4D, il est également possible de les considérer comme des vecteurs de dimension 4. Cette notation a été utilisée dans l'étude de la circularité des variables aléatoires quaternioniques [Amblard 2004]. Il existe principalement trois représentations pour un quaternion $q \in \mathbb{H}$:

- Réelle : $q_R = [a \ b \ c \ d]^t$
- Complexe : $q_C = [z_1 \ z_1^* \ z_2 \ z_2^*]^t$
- Quaternionique $q_H = [q \ q_i \ q_j \ q_k]^t$

On peut donc voir un quaternion alternativement comme un vecteur de $\mathbb{R}^4$, de $\mathbb{C}^4$ ou de $\mathbb{H}^4$. Cette façon d'appréhender les quaternions est similaire à la représentation dite de "vecteur augmenté" utilisée en traitement des signaux complexes [Schreier 2010].

1.2.3 Géométrie

La faculté des quaternions à représenter les transformations géométriques en 3D est un fait connu depuis longtemps. Ils ont été utilisés dans de nombreux domaines (animation [Shoemake 1985], aéronautique [Kuipers 1999]) et la description de leur lien avec le groupe des rotations est décrite dans de nombreux ouvrages (par exemple [Altman 1986]). Nous présentons quelques résultats bien connus sur les transformations géométriques à l'aide des quaternions. Pour un inventaire plus exhaustif, nous renvoyons à [Coxeter 1946] et [Ward 1997].

Rotations en 3D Étant donné un point P dans l'espace tridimensionnel, représenté par le quaternion $p = xi + yj + zk$, alors $p_{\mathcal{R}_q}$ donné par :

$$p_{\mathcal{R}_q} = q^{-1}pq$$

avec $q = \cos\theta + \mu \sin\theta$ un quaternion unitaire (*i.e.* $|q| = 1$), est le point obtenu par rotation du vecteur p autour de μ d'un angle 2θ . Un quaternion pur peut être vu alternativement comme un vecteur ou un point de $\mathbb{R}^3$. Attention tout de même car rigoureusement c'est un bivecteur.

Rotations en 4D Étant donné un point P de l'espace 4D, représenté par un quaternion $p = s + xi + yj + zk$, alors la transformation suivante :

$$p\mathcal{R}_{ab} = apb$$

avec $a \in \mathbb{H}$ et $b \in \mathbb{H}$ deux quaternions unitaires, est une rotation de p autour du plan défini par les parties vectorielles de a et b .

Translation de Clifford Étant donné un point de $\mathbb{R}^4$, noté P et représenté par le quaternion $p = s + xi + yj + zk$, la transformation suivante :

$$p\mathcal{T}_a^g = ap$$

avec $a \in \mathbb{H}$ un quaternion unitaire, est une translation de Clifford à *gauche*. On définit de même une translation à droite comme $p\mathcal{T}_a^g = pa$. Géométriquement, la translation de Clifford consiste en une rotation simultanée d'un même angle (si $a = e^{\eta\varphi}$ cette rotation est d'angle φ), et dans deux plans distincts de $\mathbb{R}^4$, définis par les deux nombres complexes qui forment le quaternion q ⁶.

Les quaternions, tout comme les complexes en 2D, représentent des transformations géométriques par des opérations algébriques sur des scalaires, évitant ainsi de recourir à l'usage des matrices. Ceci a des conséquences, en particulier en stabilité numérique dans l'accumulation des rotations.

1.2.4 $\mathbb{H}$ et les groupes $SU(2)$ et $SO(3)$

Comme nous venons de le voir, les quaternions permettent d'écrire de manière concise les transformations géométriques dans $\mathbb{R}^3$ (ainsi que dans $\mathbb{R}^4$). En fait, le lien entre les quaternions et le groupe des rotations est bien connu [Altman 1986]. Les quaternions unitaires sont isomorphes à $SU(2)$, le groupe des matrices complexes de déterminant égal à 1 et unitaires. On peut se convaincre de cet isomorphisme en remarquant l'isomorphisme entre les éléments de base des quaternions, *i.e.* 1, i , j et k , et les matrices de Pauli [Altman 1986].

Pour $SO(3)$, chaque élément de ce groupe est en correspondance avec deux quaternions unitaires, $+q$ et $-q$. Cet homomorphisme entre $SO(3)$ et $SU(2)$ est bien connu, avec quelques implications amusantes comme le “belt trick” [Hanson 2006] qui traduit le fait qu'il faut “deux copies de $SU(2)$ pour recouvrir $SO(3)$ ”.

Nous n'utiliserons pas dans ce chapitre cette correspondance entre quaternions unitaires et le groupe des rotations $SO(3)$, mais ce lien permet de voir les connexions possibles entre les deux parties de ce manuscrit.

1.2.5 Quaternions complexes, octonions et algèbres de Clifford

Parmi les généralisations à la dimension 8 des nombres hypercomplexes, nous en mentionnons ici deux.

Les biquaternions $\mathbb{H}_{\mathbb{C}}$ sont construits par la procédure de doublement. Cela revient à “complexifier” les coefficients d'un quaternion $q \in \mathbb{H}$ comme défini dans l'équation 1.1. Ainsi les coefficients a , b , c et d d'un biquaternion sont des complexes, avec pour nombre imaginaire I (par exemple $a = \Re(a) + I\Im(b)$), et évidemment $I^2 = -1$. Le nombre imaginaire I commute avec tous éléments de la base canonique de $\mathbb{H}$, *i.e.* $Ii = iI$, $Ij = jI$ et $Ik = kI$. Comme le prédit le théorème de Frobenius, les biquaternions ne forment pas une algèbre

6. Dans la translation de Clifford du quaternion q , les deux parties de la décomposition de Cayley-Dickson de q vont être “tournées” est fonction de l'axe de la translation de Clifford.

normée de division. En fait, il existe des biquaternions non-nuls dont la norme est 0 (diviseurs de zéros) et la norme d'un produit de biquaternions n'est pas le produit des normes⁷ des biquaternions. L'article [Sangwine 2011] qui est inclus dans la section 1.7 fait un état de l'art des propriétés des biquaternions.

Les octonions $\mathbb{O}$ sont les nombres hypercomplexes de dimension 8 qui possèdent les propriétés intéressantes vis à vis de la norme (définie de manière standard). Malheureusement, ils souffrent d'un problème assez important, ils ne sont pas associatifs par rapport au produit, *i.e.* $(ab)c \neq a(bc)$ pour $a, b, c \in \mathbb{O}$ en général. Quelques utilisations des octonions existent, principalement en Physique théorique (formulation octonionique en électrodynamique par exemple [Lounesto 2001]). On peut dire que si les octonions n'ont pas trouvé d'applications en traitement du signal, c'est sans doute parce qu'ils ne possèdent pas de représentation matricielle (une représentation matricielle est forcément associative). Ceci explique que la TF octonionique, entre autre, ne peut être obtenue *via* une représentation matricielle réelle, complexe ou quaternionique.

Les limitations des nombres hypercomplexes dans les dimensions "élevées" sont peut-être le signe que la solution se trouve du côté des algèbres de Clifford (parfois appelées également "algèbres géométriques"). Plusieurs ouvrages présentent des applications de ces algèbres en traitement des images et en robotique [Sommer 2001], et il est sans doute possible d'utiliser ces algèbres en traitement du signal (1D, à échantillons multivariés). Ce point sera abordé de nouveau dans les perspectives de mes travaux liés au signal hyperanalytique et présenté dans la section 2.3.

1.3 Matrices de quaternions

La non-commutativité des quaternions induit quelques difficultés dans l'étude des matrices de quaternions. Les matrices de quaternions ont été étudiées depuis le milieu du XX^{eme} siècle [Wolf 1936, Lee 1949, Brenner 1951] et il existe toujours une activité de recherche sur le sujet [Wu 2008]. L'article de référence est celui de F. Zhang [Zhang 1997] qui fait une synthèse des résultats théoriques connus à l'époque. À titre d'exemple, on mentionne ici deux concepts très connus dans l'étude des matrices qui se compliquent lors du passage à $\mathbb{H}$: 1) les valeurs propres d'une matrice quaternioniques peuvent être *droites* ou *gauches* [Zhang 1997]; 2) le calcul du déterminant d'une matrice dépend du sens de parcours des éléments. Le problème 1) est lâchement évité⁸ dans ce manuscrit, car nous ne parlons que de valeurs propres *droites* (les seules pour lesquelles la théorie est bien établie à ce jour). Dans le cas du déterminant, plusieurs définitions existent, par exemple, nous citons celle de J. Dieudonné [Dieudonné 1943].

Singular Value Decomposition (SVD) L'existence de la SVD pour une matrice dont les éléments sont des quaternions est connue depuis un certain temps [Zhang 1997]. Ainsi, tout matrice $\mathbf{A} \in \mathcal{M}_{N \times M}(\mathbb{H})$ peut s'écrire :

$$\mathbf{A} = \mathbf{U} \mathbf{\Delta} \mathbf{V}^t$$

avec $\mathbf{U} \in \mathcal{M}_{N \times N}(\mathbb{H})$ et $\mathbf{V} \in \mathcal{M}_{M \times M}(\mathbb{H})$ deux matrices unitaires et $\mathbf{\Delta} \in \mathcal{M}_{N \times M}(\mathbb{R})$ une matrice diagonale réelle. En 2001, nous avons proposé un algorithme de calcul de cette SVD

7. Il est à noter que la définition de la norme pose problème également pour les biquaternions. Voir [Ward 1997, Sangwine 2011]

8. ou habilement contourné, selon le point de vue ...

[Le Bihan 2004], se basant sur la *matrice adjointe*, notée $\chi_{\mathbf{A}}$ ⁹, de la matrice quaternionique $\mathbf{A}$.

En 2006 et 2007, nous avons proposé avec S.J. Sangwine, deux nouveaux algorithmes en arithmétique quaternionique pour le calcul de cette décomposition [Sangwine 2006, Le Bihan 2007b]. L'approche algorithmique purement quaternionique permet d'atteindre une précision plus importante dans le calcul des valeurs singulières et nécessite moins d'opérations arithmétiques, réduisant ainsi le coup de calcul. Ces deux algorithmes sont basés sur des extensions quaternioniques des transformations de Householder [Bunse-Gerstner 1989] et de Givens [Janovská 2003].

Eigenvalue decomposition (EVD) Comme précisé auparavant, nous parlons ici des valeurs propres droites, c'est à dire les λ t.q. $\mathbf{A}\lambda = \mathbf{v}\lambda$ avec $\mathbf{A} \in \mathcal{M}_{N \times N}(\mathbb{H})$ et $\mathbf{v} \in \mathbb{H}^N$. Plus de détails sur ces valeurs propres se trouvent dans [Zhang 1997]. Ici, nous mentionnons la contribution que nous avons apporté à l'étude des valeurs propres des matrices de biquaternions (quaternions à coefficients complexes). Nous avons proposé un algorithme de calcul des valeurs propres d'une matrice de biquaternions *Hermitienne* à l'aide d'une technique de matrice adjointe quaternionique [Le Bihan 2007a].

Matrices polynomiales Quelques résultats sont connus pour les polynômes de quaternions, et en particulier pour le calcul des zéros de ces polynômes [Serôdio 2001]. Dans la continuité des résultats de diagonalisation des matrices quaternioniques et biquaternioniques, nous avons proposé un algorithme de diagonalisation pour les matrices polynomiales quaternioniques [Le Bihan 2005]. Il est ainsi possible de montrer qu'une matrice polynomiale quaternionique $\mathbf{A}[z] \in \mathcal{M}_{N \times N}(\mathbb{H}[z])$ de degré p , donnée par :

$$\mathbf{A}[z] = \sum_{l=0}^{p-1} \mathbf{A}_l z^l = \mathbf{A}_0 + \mathbf{A}_1 z + \mathbf{A}_2 z^2 + \dots + \mathbf{A}_{p-1} z^{p-1} \quad (1.7)$$

avec $\mathbf{A}_\alpha \in \mathcal{M}_{N \times N}(\mathbb{H}) \forall \alpha$ et $z = \Re(z) + I\Im(z) \in \mathbb{C}$, peut être diagonalisée dans le cas parahermitien (voir les détails dans [Le Bihan 2005]). L'algorithme que nous avons proposé pour cette diagonalisation est basé sur les transformations de Givens quaternioniques. Cet algorithme a été utilisé à des fins de séparation de mélanges convolutifs d'ondes polarisées [Le Bihan 2005, Menanno 2010].

1.4 Transformation de Fourier quaternionique

L'idée de définir une Transformée de Fourier Quaternionique (TFQ), ou hypercomplexe, est apparue à plusieurs auteurs, et vraisemblablement de manière indépendante, dans des domaines de recherche distincts. En 1987-88, une TF hypercomplexe (bicomplexe en fait) est introduite pour le traitement des données RMN 2D [Ernst 1987, Delsuc 1988]. Cette transformée a été introduite afin de séparer des fréquences de rotation de spins en RMN.

La transformation de Fourier quaternionique a ensuite été redécouverte en 1992 par Todd A. Ell dans un autre contexte. Il avait pour objectif de procurer un équivalent à la transformée de Fourier dans le cas de systèmes bilinéaires invariants dans le temps. Il a ainsi pu proposer une description en gain-phase de ces systèmes [Ell 1992, Ell 1993].

En 1996, S.J. Sangwine a proposé une transformation de Fourier 2D pour les images couleurs, *i.e.* pour les signaux bidimensionnels à échantillons vectoriels (chaque pixel RGB

9. La matrice adjointe d'une matrice quaternionique $\mathbf{A} \in \mathcal{M}_{N \times M}(\mathbb{H})$ est une matrice $\chi_{\mathbf{A}} \in \mathcal{M}_{2N \times 2M}(\mathbb{C})$, à valeurs complexes, dont les propriétés peuvent être reliées à celle de $\mathbf{A}$, voir [Zhang 1997, Le Bihan 2004] pour plus de détails

est représenté par un quaternion pur) [Sangwine 1996]. S.J. Sangwine a ensuite beaucoup travaillé sur les différentes définitions des TF 2D quaternioniques, établissant le lien entre les différentes définitions possibles (gauche, droite, bilatérale)¹⁰. Bien que sa définition originale de la TFQ 2D soit bilatérale [Ell 1993, Sangwine 1996], S.J. Sangwine a principalement travaillé depuis sur les TFQ unilatérales et leur interprétation dans l'étude des images couleurs. La TFQ unilatérale *droite* d'un signal bivarié $s(x, y) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{H}$ est donnée par :

$$S_\mu(\nu, \omega) = TFQ_\mu[s(x, y)] = \int_{-\infty}^{+\infty} s(x, y) e^{-2\pi\mu(\nu x + \omega y)} dx dy$$

avec μ un quaternion pure unitaire. On remarque tout de suite qu'il existe une infinité de définitions possibles du fait du choix possible de μ . En pratique, pour les images couleurs, le choix de μ n'a pas beaucoup d'importance. On verra plus tard qu'il est plus contraint dans le cas de TFQ 1D.

Au début des années 2000, l'équipe de G. Sommer à Kiel, sous l'impulsion de T. Bulow, s'est intéressée à l'extension du signal analytique au cas des images (signaux 2D), en utilisant la TFQ. Une bonne synthèse de leurs travaux se trouve dans le livre édité par Sommer [Sommer 2001]. Les travaux de Bulow sont basés sur une TFQ bilatérale très similaire à celle de T.A. Ell [Ell 1992] et s'intéressent à proposer une TFQ pour des images dont les pixels sont à valeurs réelles. Ainsi, la TFQ bilatérale de $s(x, y) : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ est donnée par :

$$S(\nu, \omega) = \int_{-\infty}^{+\infty} e^{-2\pi j \nu x} s(x, y) e^{-2\pi k \omega y} dx dy$$

L'utilisation de cette transformée pour la définition d'un signal analytique 2D conduit à la définition de plusieurs phases exploitables dans le cas de signaux 2D séparables [Bulow 2001]. On notera que Bulow utilise dans ses travaux un triplet de phases¹¹ qui sont liées aux angles d'Euler, ce qui revient à utiliser encore une autre représentation des quaternions (voir [Bulow 2001]).

Depuis la publication de tous ces résultats sur les TFQs 2D, quelques applications ont été proposées pour ces algorithmes. Encore une fois, sans vouloir être exhaustif, on peut citer par exemple : l'analyse d'images [Ell 2007], l'estimation de mouvement [Alexiadis 2009], le watermarking [Bas 2003] ou la détection en sonar [Redfield 2002], etc.

L'idée que nous avons avancée avec S.J. Sangwine en 2006 est différente. Nous avons entrepris de définir une transformation de Fourier Quaternionique pour des signaux à valeurs complexes 1D. L'idée de base est en fait la même que celle développée en RMN ou dans les travaux de T. Bulow, elle consiste à séparer les symétries du signal dans des parties imaginaires de la transformée. Dans le cas complexe 1D, on sait que la transformée de Fourier (complexe) d'un tel signal ne satisfait plus la symétrie hermitienne ($S^*(\nu) = S(-\nu)$) quand le signal complexe est "quelconque" (parties réelles et imaginaires possédant une partie paire et une partie impaire). Ainsi, un signal complexe $z(t)$ peut s'écrire d'une manière générale :

$$z(t) = [\Re(z(t))_p + \Re(z(t))_i] + \mathbf{i} [\Im(z(t))_p + \Im(z(t))_i]$$

10. La non-commutativité des quaternions rend possible la définition de plusieurs TF, selon que le noyau de la transformée est à gauche, droite, ou que deux noyaux (à gauche et droite) soient utilisés. On consultera [Ell 2007] pour un état de l'art récent et l'équivalence entre les différentes transformations 2D possibles.

11. En fait, il existe encore une autre notation de type "polaire" pour les quaternions. On peut écrire $q \in \mathbb{H}$ comme : $q = \|q\| \exp(\mathbf{i}\varphi) \exp(\mathbf{i}\theta) \exp(\mathbf{i}\psi)$. Nous n'utiliserons pas cette notation dans ce manuscrit, mais il pourrait être intéressant d'essayer de l'appliquer dans l'interprétation du signal hyperanalytique décrit dans la section 2.3.

$z(t)$	$Z_j(\nu)$
Réel / pair	Réel / pair
Reél / impair	Imaginaire pur $\Im_j$ / impair
Réel / quelconque	Complexe $\Re + j\Im_j \begin{cases} \Re : \text{even} \\ \Im_j : \text{odd} \end{cases}$
Imaginaire pur $\Im_i$ / pair	Imaginaire pur $\Im_i$ / pair
Imaginaire pur $\Im_i$ / impair	Imaginaire pur $\Im_k$ / impair
Imaginaire pure $\Im_i$ / quelconque	Imaginaire $i\Im_i + k\Im_k \begin{cases} \Im_i : \text{pair} \\ \Im_k : \text{impair} \end{cases}$
Complexe / pair	Complexe $\Re + \Im_j$ / pair
Complexe / impaire	Complexe $j\Im_j + k\Im_k$ / impaire
Complexe / $\begin{cases} \Re : \text{pair} \\ \Im_i : \text{impair} \end{cases}$	Complexe $\Re + k\Im_k \begin{cases} \Re : \text{pair} \\ \Im_k : \text{impair} \end{cases}$
Complexe / $\begin{cases} \Re : \text{impair} \\ \Im_i : \text{pair} \end{cases}$	Imaginaire $i\Im_i + j\Im_j \begin{cases} \Im_i : \text{pair} \\ \Im_j : \text{impair} \end{cases}$
Complexe / quelconque	Quaternion 'complet' $\mathbb{H} \begin{cases} \Re : \text{pair} \\ \Im_i : \text{pair} \\ \Im_j : \text{impair} \\ \Im_k : \text{impair} \end{cases}$

FIGURE 1.1 – Symmétries de la TF quaternionique 1D.

les indices p et i désignant les parties paires et impaires. Maintenant, si l'on définit une TFQ pour ce signal $z(t)$ de la façon suivante :

$$Z_\mu(\nu) = \int_{-\infty}^{\infty} z(t) e^{-2\mu\pi\nu t} dt$$

avec μ quaternion unitaire pur, les quatre parties du signal se retrouvent dans les quatre parties de la TFQ (réelle et les trois imaginaires). Ici encore, la définition de la TFQ possède un degré de liberté : l'axe μ . En toute généralité, on pourrait définir toutes les propriétés de la TFQ calculée avec cet axe, mais afin de rendre les choses plus explicites nous allons prendre le cas particulier où $\mu = j$ dans la suite.

On peut trouver les symétries de $Z(\nu)$ assez facilement en fonction des propriétés de parité de $z(t)$. Ces propriétés sont résumées dans le tableau 1.1 (on utilise les notations $\Im_i$, $\Im_j$ et $\Im_k$ pour les parties imaginaires i , j et k). Il existe d'autres symétries que l'on peut obtenir en raisonnant dans le domaine fréquentiel. Elles ne sont pas mentionnées ici mais permettent d'établir des propriétés du signal hyperanalytique qui est présenté dans la partie 2.3. À titre d'exemple, et afin d'illustrer le comportement de la TFQ, nous donnons également d'autres propriétés de cette transformée. Tout d'abord, si la TFQ du signal complexe $z(t)$ est $Z_j(\nu)$ (axe de transformation j), alors la TFQ de $z^*(t)$ est liée à celle de $z(t)$ comme suit :

$$QFT_j[z^*(t)] = -jZ(\nu)j$$

On peut également se demander ce qu'il advient de la symétrie Hermitienne, connue pour la TF des signaux réels, dans le cas de la TFQ des signaux complexes. La propriété équivalente est la suivante : Soit un signal $z(t)$ à valeurs complexes (*i.e.* $z(t) = \Re(z(t)) + i\Im(z(t))$) dont la TFQ_j est notée $Z(\nu)$ (à valeurs quaternioniques donc). On a alors la

relation suivante par calcul direct :

$$X(\nu) = -\mathbf{i}X(-\nu)\mathbf{i}$$

On voit ainsi que dans le cas des TFQ, c'est l'involution dont l'axe est le même que celui de la transformation qui joue le rôle de la conjugaison dans le cas "classique"¹².

Finalement, on peut tenter de réécrire l'identité de **Parseval** dans le cadre des TFQ. Étant donnés $x(t)$ et $y(t)$ deux signaux complexes, et leurs TFQ_j notées $X(\nu)$ et $Y(\nu)$, alors l'égalité suivante est vérifiée :

$$\int_{-\infty}^{+\infty} x(t)y^*(t)dt = \int_{-\infty}^{+\infty} X(\nu)(-\mathbf{j}Y(\nu)\mathbf{j})d\nu$$

On démontre aisément cette égalité avec la propriété de la TFQ de $y^*(t)$ donnée précédemment.

Ces propriétés illustrent les spécificités du calcul impliquant la TFQ et montrent l'importance des involutions dans l'étude de cette transformée. On gardera à l'esprit également que la non-commutativité du produit des TFQ rend leur manipulation délicate et qu'un soin particulier doit être apporté à l'ordre des produits. Toutefois, une étude systématique de la TFQ 1D reste encore à être menée. Une étude complète n'existe pas à l'heure actuelle. Cette étude fait partie de mes perspectives. Dans nos travaux, nous avons utilisé cette transformée afin de définir des grandeurs instantanées pour les signaux complexes non-stationnaires non-circulaires. Ce travail est présenté dans la section 2.3 et illustre l'intérêt de la TFQ dans l'étude des signaux complexes non-circulaires.

Enfin, on peut se demander comment étendre ces résultats obtenus pour la transformée de Fourier Quaternionique à des dimensions supérieures. Plusieurs solutions existent, et nous avons travaillé sur une extension à la dimension 8 : la TF Biquaternionique [Said 2008b]. Nous avons étudié cette transformation et montré son intérêt pour l'étude des signaux à échantillons vectoriels 4D. Les détails sont dans l'article inclus dans la section 1.7. On peut évidemment se poser la question de la généralisation de ces transformées hypercomplexes aux dimensions supérieures à 8. Les travaux que nous avons mené depuis quelques années montrent que l'aspect géométrique qu'apportent les transformées quaternioniques est central. Si cet aspect est accessible en dimension 4 ou 8, au delà, c'est moins évident. Il me semble qu'une théorie plus générale doit pouvoir être établie en se basant sur les algèbres de Clifford [Sommer 2001]. Les transformations de Fourier de Clifford existent et doivent permettre de mener le même type d'analyse que celles quaternioniques que nous avons menées, mais pour n'importe quelle dimension N . L'étude de ces transformées en traitement des signaux est une des perspectives de ces travaux.

1.5 Circularité

La notion de circularité pour les variables aléatoires et les signaux complexes est bien connue en traitement du signal. Elle a donné lieu à un grand nombre de publications depuis les premiers travaux dans le domaine du signal par B. Picinbono [Picinbono 1994]. Le cas circulaire est finalement assez simple et beaucoup des travaux récents s'intéressent en particulier au cas non-circulaire¹³. Ne pouvant être exhaustif, nous renvoyons à l'ouvrage

12. En fait, la conjugaison est une involution sur $\mathbb{C}$, alors qu'elle ne l'est pas (!) sur $\mathbb{H}$. La symétrie Hermitienne est donc liée aux involutions sur l'espace considéré et s'étend donc au cas quaternionique *via* celles-ci.

13. On rappelle qu'une variable aléatoire complexe z est dite circulaire au sens large si $z \stackrel{d}{=} ze^{i\theta}$ pour tout $\theta \in \mathbb{R}$. La conséquence est que la partie réelle et la partie imaginaire de z sont décorréllées et de même variance.

récent [Schreier 2010] et les références s’y trouvant. Ce livre présente une synthèse des travaux dans le domaine.

Dans le cas des quaternions, la notion de circularité a tout d’abord été introduite par N.N. Vakhania [Vakhania 1998]. En 2004, nous avons généralisé la définition de Vakhania [Amblard 2004] pour la circularité des variables aléatoires quaternioniques. Ainsi, pour une variable aléatoire quaternionique q , il existe deux niveaux de circularité : la $\mathbb{C}_\eta$ -circularité et la $\mathbb{H}_\eta$ -circularité. Ainsi, une variable aléatoire quaternionique q est dite $\mathbb{C}_\eta$ -circulaire si :

$$q \stackrel{d}{=} e^{\eta\varphi} q, \quad \forall \varphi$$

pour un et un seul quaternion unitaire pur $\eta = i, j$ ou k (quand q est exprimé dans la base quaternionique classique $\{1, i, j, k\}$). Une variable aléatoire quaternionique q sera dite $\mathbb{H}_\eta$ -circulaire si la même égalité est vérifiée quelque soit le quaternion pur η . On remarquera que dans le cas quaternionique, il ne s’agit pas d’une invariance par rotation de la distribution de q pour la $\mathbb{C}_\eta$ -circularité, mais une invariance par translation de Clifford à gauche.

La différence majeure entre le cas complexe et le cas quaternionique est donc qu’il existe deux niveaux de circularité sur $\mathbb{H}$. Ces niveaux de circularité se répercutent sur la matrice de covariance de la variable q dont la structure peut être étudiée grâce aux représentations vectorielles données en 1.2.2. Ces matrices de covariance ainsi que des détails techniques et un exemple de distribution Gaussienne $\mathbb{C}_j$ -circulaire sont visibles dans l’article [Amblard 2004] inclus dans la section 1.7. Nos travaux ont récemment été repris et les implications de la circularité dans le traitement des signaux quaternioniques sont utilisées dans différents algorithmes : détection/estimation [Le Bihan 2006a, Via 2010a], filtrage [Mandic 2011, Took 2009, Ujang 2009, Took 2010] ou analyse en composantes indépendantes [Le Bihan 2006b, Via 2011, Via 2010c].

1.6 Conclusion

Ce chapitre a présenté quelques résultats connus sur les quaternions ainsi que quelques contributions récentes sur les représentations, les transformations de Fourier ou les matrices quaternioniques. Ces résultats ont été obtenus dans le but de résoudre des problèmes apparaissant lors du traitement de signaux à valeurs complexes ou quaternioniques. Certains de ces traitements seront présentés au chapitre 2.

Depuis 1993, beaucoup de nouveaux résultats ont permis de comprendre l’intérêt des quaternions pour le traitement du signal et depuis le début des années 2000, une petite communauté s’est créée en traitement du signal, qui fait avancer les connaissances sur le sujet et en parallèle fait apparaître de nouveaux problèmes intéressants autour des quaternions. En particulier, l’algèbre linéaire (matriciel) sur $\mathbb{H}$ n’a pas encore complètement été formalisé, et par exemple, l’intérêt des valeurs propres gauches n’a pas été identifié. On peut également se demander comment les décompositions multilinéaires comme PARAFAC ou la HOSVD, qui sont devenues populaires en signal, se comportent quand les tableaux/tenseurs sont à valeurs quaternioniques. C’est peut-être une piste de recherche intéressante.

Finalement, je mentionne que plusieurs des résultats de simulations obtenus dans nos travaux quaternioniques ont été obtenus grâce à la Toolbox Matlab QTFM (Quaternion Toolbox For Matlab) développée en collaboration avec S.J. Sangwine [QTFM 2005].

1.7 Publications annexées en lien avec ce chapitre

Les articles suivants sont inclus ici :

1. “Fundamental representation and algebraic properties of biquaternions or complexified quaternions”, S.J. Sangwine, T.A. Ell and N. Le Bihan, *Advances in Applied Clifford Algebras*, Online First, 2011.
2. “Fast Complexified Quaternion Fourier Transform”, S. Said, N. Le Bihan and S.J. Sangwine, *IEEE Transactions on Signal Processing*, Vol. 56, No. 4, 2008.
3. “On Properness of quaternion random variables”, P.O. Amblard and N. Le Bihan, *IMA Conf. on Math. in Signal Processing*, Cirencester, 2004.

Fundamental Representations and Algebraic Properties of Biquaternions or Complexified Quaternions

Stephen J. Sangwine, Todd A. Ell and Nicolas Le Bihan

Abstract. The fundamental properties of biquaternions (complexified quaternions) are presented including several different representations, some of them new, and definitions of fundamental operations such as the scalar and vector parts, conjugates, semi-norms, polar forms, and inner products. The notation is consistent throughout, even between representations, providing a clear account of the many ways in which the component parts of a biquaternion may be manipulated algebraically.

Mathematics Subject Classification (2010). Primary 11R52; Secondary 15A66.

Keywords. Quaternion, biquaternion.

It is typical of quaternion formulae that, though they be difficult to find, once found they are immediately verifiable.

J. L. Synge (1972) [40, p 34]

1. Introduction

Fundamental properties of the quaternions are relatively accessible in the literature, both in terms of abstract algebraic properties and applied formulae. This is less true for quaternions with complex components (complexified quaternions, or biquaternions¹), even though the algebra, being isomorphic to the Clifford algebra $\mathcal{Cl}_{3,0}$, has been well studied. This paper aims to set out the fundamental definitions of biquaternions and some elementary results, which, although elementary, are often not trivial and thereby render more accessible the fundamental properties of the biquaternions. The emphasis in this paper is on the biquaternions as an applied (and numerical) algebra – that is, a tool for the manipulation of algebraic expressions and formulae to

¹Biquaternions was a word coined by Hamilton himself [19] and [17, § 669, p664]. (The word was used 18 years later by Clifford [6] for a different concept, which is unfortunate.)

allow deep insights into scientific and engineering problems. It is not a study of the abstract properties of the biquaternion algebra, nor its relations with other algebras. Throughout the paper ‘quaternion’ means a quaternion with real components (a quaternion over the reals, $\mathbb{R}$), and ‘biquaternion’ means a quaternion with complex components (a quaternion over the field of complex numbers, $\mathbb{C}$). We denote the set of quaternions by $\mathbb{H}$, and the set of biquaternions by $\mathbb{B}$. $\mathbb{H}$ is, of course, a subset of $\mathbb{B}$.

Some of the material in the paper is based on the book by Ward [42, Chapter 3] which is one of the few readily accessible sources of detail on the biquaternions. We have not followed Ward’s notation in this paper, preferring instead a scheme based on bold or plain symbols without hats and underlines.

We have also drawn upon the paper by Sangwine and Alfsmann [34] which sets out comprehensive results on the divisors of zero, and their subsets the idempotents and nilpotents. Sangwine and Alfsmann’s paper uses the same notations as this paper (having benefited from access to this paper in draft).

The quaternions themselves (with real elements) are well-covered in various books, for example [2, 3, 26, 27, 28]. Hamilton’s works on quaternions were published in book form in [17, 18, 14], and many are also now available freely on the Internet in various digital repositories. The paper by Coxeter [7] is also a useful source.

We begin by setting out various ways in which quaternions and biquaternions may be represented. We start with representations for quaternions, for reference, and because these representations are generalized for biquaternions. In what follows, we use notation as consistently as we can. In particular, the complex operator usually denoted by i (or in electrical engineering j) is represented in this paper as I in every case. This is to keep the complex operator distinct from the first of the three quaternion operators $\mathbf{i}$, since it is independent. The independence of $\mathbf{i}$ and I is perhaps the most fundamental axiomatic aspect of the biquaternions that must be understood. Bold symbols denote vectors and bivectors², whereas normal weight symbols denote scalar or complex quantities or quaternions.

Throughout the paper³ we use the term *norm* or the more specialized term *semi-norm*, both denoted $\|q\|$, to mean the sum of the squares of the components of a quaternion or biquaternion, and *modulus*, denoted $|q|$, to mean the square root of the norm and thus the Euclidean magnitude. This is not universally accepted terminology, many sources using *norm* where we use *modulus*. However, our usage is consistent with several authors who have written on quaternions, including Synge [40] and Ward [42], but it does require care when using statements made about norms in other sources.

Many of the concepts given in this paper are implemented in numerical form in a MATLAB® toolbox [36] which two of the authors first developed in 2005 and are built upon in a toolbox for handling linear quaternion systems,

²Bivectors represent directed areas, and are explained in Table 3 and § 3.7.

³Except in (32), where we use the notation of the cited reference.

first developed in 2007 [11]. The toolbox [36] was essential to the development of this paper and the results presented within it, since otherwise, errors in algebra would go unnoticed. In many cases we have established results first by using the toolbox, and then derived the algebraic proofs or statements which appear here.

2. Quaternions

Classically, quaternions are represented in the form of hypercomplex numbers with three imaginary components. In Cartesian form this is:

$$q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \quad (1)$$

where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are mutually perpendicular unit bivectors⁴ obeying the famous multiplication rules: $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\mathbf{j}\mathbf{k} = -1$, discovered by Hamilton in 1843 [20], and w, x, y, z , are real. Quaternions are generalized to biquaternions by permitting w, x, y and z to be complex, as discussed in § 3, but in this paper we reserve the symbols w, x, y, z for the real case. A quaternion with $w = 0$ is known as a *pure* quaternion (Hamilton's terminology, but still used and widely understood).

The conjugate of a quaternion is given by negating the three imaginary components: $\bar{q} = w - x\mathbf{i} - y\mathbf{j} - z\mathbf{k}$. It is easily shown that $\bar{q}p = p\bar{q}$ for general quaternions p and q . Indeed the formula may be generalized to more than two quaternions (the generalized formula was first noted by Hamilton [22, § 20, p238], and was also included in [27, p60, § 31]): $\overline{pqrst} = \bar{t}\bar{s}\bar{r}\bar{q}\bar{p}$. The quaternion conjugate may be expressed in terms of multiplications and additions [10, Theorem 11] using any system of three mutually orthogonal unit pure quaternions (here $\mathbf{i}, \mathbf{j}, \mathbf{k}$):

$$\bar{q} = -\frac{1}{2}(q + \mathbf{i}q\mathbf{i} + \mathbf{j}q\mathbf{j} + \mathbf{k}q\mathbf{k}) \quad (2)$$

Similar formulae, based on involutions [9, 10], exist for extracting the four Cartesian components of a quaternion⁵ [38]:

$$\begin{aligned} w &= \frac{1}{4}(q - \mathbf{i}q\mathbf{i} - \mathbf{j}q\mathbf{j} - \mathbf{k}q\mathbf{k}), & x &= \frac{1}{4\mathbf{i}}(q - \mathbf{i}q\mathbf{i} + \mathbf{j}q\mathbf{j} + \mathbf{k}q\mathbf{k}) \\ y &= \frac{1}{4\mathbf{j}}(q + \mathbf{i}q\mathbf{i} - \mathbf{j}q\mathbf{j} + \mathbf{k}q\mathbf{k}), & z &= \frac{1}{4\mathbf{k}}(q + \mathbf{i}q\mathbf{i} + \mathbf{j}q\mathbf{j} - \mathbf{k}q\mathbf{k}) \end{aligned} \quad (3)$$

The norm of a quaternion is given by the sum of the squares of its components: $\|q\| = w^2 + x^2 + y^2 + z^2$, $\|q\| \in \mathbb{R}$. It can also be obtained by

⁴Classic texts often refer to the operators $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ as *vectors*, a misconception that has caused considerable confusion over many years, but is understandable, since it could not be cleared up without the concept of geometric algebra and bivectors. We discuss this in § 3.7.

⁵These formulae, and that for the quaternion conjugate, generalize to biquaternions. If an arbitrary set of mutually orthogonal unit pure quaternions or biquaternions $\mu, \xi, \mu\xi$, is substituted for $\mathbf{i}, \mathbf{j}$ and $\mathbf{k}$, the formulae give the four components of the biquaternion q expressed in a new basis defined by $(1, \mu, \xi, \mu\xi)$.

multiplying the quaternion by its conjugate, in either order since a quaternion and its conjugate commute: $\|q\| = q\bar{q} = \bar{q}q$. The modulus of a quaternion is the square root of its norm: $|q| = \sqrt{\|q\|}$.

Every non-zero quaternion has a multiplicative inverse⁶ given by its conjugate divided by its norm: $q^{-1} = \bar{q}/\|q\|$.

The quaternion algebra $\mathbb{H}$ is a normed division algebra, meaning that for any two quaternions p and q , $\|pq\| = \|p\| \|q\|$, and the norm of every non-zero quaternion is non-zero (and positive) and therefore the multiplicative inverse exists for any non-zero quaternion.

Of course, as is well known, multiplication of quaternions is not commutative, so that in general for any two quaternions p and q , $pq \neq qp$. This can have subtle ramifications, for example: $(pq)^2 = pqpq \neq p^2q^2$.

Alternative representations for quaternions are given in Table 1, expressed in terms of the Cartesian form given above in (1), and in selected cases, in terms of other representations given here.

TABLE 1. Representations for quaternions.

Designation	Representation	Details ($w, x, y, z, a, b, r, \theta \in \mathbb{R}$, $q, \mu \in \mathbb{H}$, $S(\mu) = 0$, $\mu^2 = -1$)
Cartesian	$q = w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$	
Scalar + vector	$q = S(q) + \mathbf{V}(q)$	$S(q) = w$ $\mathbf{V}(q) = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$
‘Complex’ form	$q = a + \mu b$	$a = w$ $b = \sqrt{x^2 + y^2 + z^2} = \mathbf{V}(q) $ $\mu = \mathbf{V}(q) / \mathbf{V}(q) $
Cayley-Dickson	$q = (w + x\mathbf{i}) + (y + z\mathbf{i})\mathbf{j}$	This multiplies out to the Cartesian representation.
Polar form	$q = r \exp(\mu\theta)$ $= r(\cos\theta + \mu \sin\theta)$	$r = q $, $r \cos\theta = a$ $r \sin\theta = b$
Cayley-Dickson polar form	$q = \mathcal{A} \exp(\mathcal{B}\mathbf{j})$	$\mathcal{A}, \mathcal{B} \in \mathbb{H}$ with $y = z = 0$ (isomorphic to $\mathbb{C}$) See [35] for formulae defining $\mathcal{A}$ and $\mathcal{B}$.

μ is a unit pure quaternion and is known as the *axis* of the quaternion. It expresses the direction in 3-space of the vector part⁷, $\mathbf{V}(q)$. Hamilton

⁶This does not apply to biquaternions.

⁷We use the term ‘vector part’ throughout this paper to mean that part of the quaternion consisting of the three components containing $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$. It does not necessarily correspond to the concept of a 3-space vector, since it could be a vector, bivector or a combination of both, using the language of geometric algebra described later. The term ‘vector part’ is well-established in the literature and, lacking a good alternative which would be readily understood, we retain it.

himself showed that any unit pure quaternion is a square root of -1 [21, pp 203, 209][17, § 167, p179] and a proof is also given in [9, Lemma 1]. This is why we call the form $a + \boldsymbol{\mu}b$ the ‘complex’ form, since it is isomorphic to a complex number $a + Ib$. This means that the modulus and argument of the quaternion are identical to those of $a + Ib$ (one could think of them as having similar Argand diagrams, where in the case of quaternions, the Argand diagram represents a plane section of 4-space defined by the ‘axis’ $\boldsymbol{\mu}$ and the scalar quaternion axis – along which w is measured).

The polar form of a quaternion is analogous to the polar form of a complex number, with one exception. The argument, θ , is confined to the interval $[0, \pi)$ because the modulus of the vector part is conventionally taken to be positive (there is no universally applicable⁸ coordinate-invariant way to define an orientation in 3-space which would permit the sign of the vector part to be determined). If we negate the argument of the exponential in the polar form, therefore, the negation is conventionally applied to the axis, $\boldsymbol{\mu}$ and not to the argument θ . When θ is computed numerically, the result is always in $[0, \pi)$ because we have to use the (non-negative) modulus of the vector part to compute it (using an `atan2` function, typically).

The Cayley-Dickson polar form [35] has a complex modulus $\mathcal{A}$, and a complex argument $\mathcal{B}$, (both in fact are degenerate quaternions of the form $w + \boldsymbol{i}x$, isomorphic to complex numbers).

3. Biquaternions

To generalize the quaternions to biquaternions we simply permit the four elements to be complex rather than real, thus giving us the Cartesian representation:

$$q = W + X\boldsymbol{i} + Y\boldsymbol{j} + Z\boldsymbol{k} \quad (4)$$

where $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$ are exactly as in § 2 and W , X , Y , Z , are complex. This generalization was first studied by Hamilton himself [19] and [17, § 669, p664], and was also discussed by Cayley [41].

In (4) each of the four elements is of the form $W = \Re(W) + I\Im(W)$, where $I^2 = -1$ is the usual complex operator, distinct from $\boldsymbol{i}$. Axiomatically, I commutes with the three quaternion operators $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$, that is $\boldsymbol{i}I = I\boldsymbol{i}$, $\boldsymbol{j}I = I\boldsymbol{j}$ and $\boldsymbol{k}I = I\boldsymbol{k}$. Since reals commute with the three quaternion operators, so do all complex numbers of the form $a + Ib$, where $a, b \in \mathbb{R}$. It is important to maintain a clear conceptual separation between complex numbers $a + Ib$ and quaternions of the form $a + b\boldsymbol{i}$ or $a + b\boldsymbol{j}$, which, although isomorphic to a complex number, remain quaternions (or biquaternions if the coefficients are complex). Similarly the ‘complex form’ of a quaternion $a + \boldsymbol{\mu}b$ in Table 1 is a quaternion, and not a complex number in the sense we are talking about here since $\boldsymbol{\mu}$ is a unit pure quaternion.

⁸Of course, in *specific applications* it may be possible to define a reference direction.

Some familiar rules of algebra apply to biquaternions. Since the real and imaginary parts are quite separate from the concept of the four quaternion components, real and imaginary parts may be equated just as when working with complex equations. However, some of the elementary properties of quaternions can become non-elementary when the quaternions are complexified. For example, some generalizations of the norm and modulus of a quaternion are complex in general for biquaternions, and so is the argument in one of the polar forms. We discuss each of these non-trivial properties in a later section.

The existence of complex generalizations of the norm, modulus and inner product, yields a problem of terminology, since conventionally these quantities are real, and satisfy properties that their complex generalizations cannot (the *triangle inequality* $\|p + q\| \leq \|p\| + \|q\|$, for example, requires ordering, but complex numbers lack ordering, hence the triangle inequality cannot be applicable to a ‘norm’ with a complex value). Rather than invent new terms, we use the existing accepted terms (with the exception of *norm*, where we substitute *semi-norm*, for reasons discussed in §3.3), but caution the reader that because these quantities are complex, they cannot be assumed to satisfy all the usual properties of their conventional real equivalents. In taking this approach, we are following Synge [40, p 9], and to some extent Ward [42], both of whom use the term *norm* to refer to a complex generalization. Synge also refers to a scalar product with a complex value, and Ward uses the term *inner product* for a complex generalization of the concept in his section on the Minkowski metric [42, § 3.3, p 115].

Although a biquaternion commutes with its quaternion conjugate, and complex numbers commute (including with their complex conjugates), somewhat surprisingly, a biquaternion does not necessarily commute with its complex conjugate (Proposition 1 in §3.1).

Alternative representations for biquaternions are shown in Table 2, expressed in terms of the Cartesian form given above in (4), and in selected cases, in terms of other representations given here. In the scalar/vector form, the scalar part is complex, and the vector part is a pure biquaternion.

Complex form I corresponds to the ‘complex’ form in Table 1. The differences are that A and B are now complex, whereas a and b were real, and the imaginary unit ξ is a pure biquaternion root of -1 of the form $\xi = b\mu + dI\nu$, where $b^2 - d^2 = 1$ and μ and ν are mutually perpendicular unit pure quaternions (themselves also roots of -1) [33]. Note that B can vanish – as discussed in §3.6.2 and in detail in [34, § 2] – although its components (X , Y and Z), and therefore q , do not vanish. ξ can be thought of as a complex axis, in the sense that it has real and imaginary parts which each define directions in 3-space. The geometric interpretation of biquaternions is discussed further in §3.7. As in the quaternion case, B (the result of the square root operation) may be arbitrarily negated, provided ξ is negated to compensate. Conventionally, the computation of a complex square root yields

TABLE 2. Representations for biquaternions.

Designation	Representation	Details: $w_a, x_a, y_a, z_a \in \mathbb{R}, a \in \{r, i\}$ $W, X, Y, Z, A, B, R, \Theta \in \mathbb{C}$ $q_r, q_i, Q, \Psi \in \mathbb{H}$ $q, \boldsymbol{\xi} \in \mathbb{B}, \boldsymbol{\xi}^2 = -1$ [33]
Cartesian	$q = W + X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}$	
Scalar + vector	$q = S(q) + \mathbf{V}(q)$	$S(q) = W$ $\mathbf{V}(q) = X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k}$
‘Complex’ form I	$q = A + \boldsymbol{\xi}B$	$A = W = S(q)$ $\boldsymbol{\xi}B = \mathbf{V}(q)$ $B = \sqrt{X^2 + Y^2 + Z^2}$ $= \mathbf{V}(q) $ $\boldsymbol{\xi} = (X\mathbf{i} + Y\mathbf{j} + Z\mathbf{k})/B$
‘Complex’ form II	$q = q_r + Iq_i$ $= \Re(q) + I\Im(q)$	$q_r = w_r + x_r\mathbf{i} + y_r\mathbf{j} + z_r\mathbf{k}$ $q_i = w_i + x_i\mathbf{i} + y_i\mathbf{j} + z_i\mathbf{k}$
Expanded form	$q = w_r + x_r\mathbf{i} + y_r\mathbf{j} + z_r\mathbf{k}$ $+ I(w_i + x_i\mathbf{i} + y_i\mathbf{j} + z_i\mathbf{k})$	$w_r = \Re(W), w_i = \Im(W)$ $x_r = \Re(X), x_i = \Im(X)$ $y_r = \Re(Y), y_i = \Im(Y)$ $z_r = \Re(Z), z_i = \Im(Z)$
Hamilton polar form	$q = R \exp(\boldsymbol{\xi}\Theta)$ $= R(\cos \Theta + \boldsymbol{\xi} \sin \Theta)$	$R = q $ $A = R \cos \Theta$ $B = R \sin \Theta$
Complex polar form	$q = Q \exp(I\Psi)$ $= Q(\cos \Psi + I \sin \Psi)$	$\Psi = \tan^{-1}(q_r^{-1}q_i)$ $Q = q / \exp(I\Psi)$ $= q \exp(-I\Psi)$ $q_r = Q \cos \Psi$ $q_i = Q \sin \Psi$

a result in the right half of the complex plane with positive real part. If this convention is followed, the ‘sign’ of $\boldsymbol{\xi}$ is uniquely determined.

Complex form II has quaternions in the real and imaginary parts, and is perhaps the most obvious representation for a biquaternion other than the Cartesian form. It is related to the complex polar form described below.

In the expanded form, the biquaternion is represented as a complex number with quaternion real and imaginary parts expressed in Cartesian form.

The polar form of a quaternion depends on Euler’s formula $\exp(I\theta) = \cos \theta + I \sin \theta$ which generalizes by replacing the complex operator I with any

root of -1 . In the case of quaternions, the set of unit pure quaternions provides an infinite number of roots of -1 and the general polar form $r \exp(\mu\theta)$, as given in Table 1 is therefore a straightforward extension of Euler's formula. In the case of biquaternions there are two possibilities for the root of -1 : the complex root I , or any one of the biquaternion roots of -1 defined in §3.5. Thus we have two possible fundamental polar forms.

The first ('Hamilton') polar form generalizes the polar form in the real case. R , the 'modulus' of the biquaternion, is complex; ξ is a biquaternion root of -1 ; and Θ , the 'angle' in the exponential, is also complex⁹. The interpretation of complex angles is discussed further in §3.2.

In the second ('complex') polar form, the standard complex operator I serves as the root of -1 in the exponential, but in this form, the 'argument' of the exponential, Ψ , is a quaternion, and the exponential is scaled by a quaternion 'modulus' Q . We thus have a polar form with a 'modulus' and 'argument' in $\mathbb{H}$. Note that it is not possible to find Q by the obvious direct route of $Q = \sqrt{q_r^2 + q_i^2}$ because $q_r = Q \cos \Psi$ and squaring this gives $Q \cos \Psi Q \cos \Psi$ and not $Q^2 \cos^2 \Psi$. However, it is the case that $\cos^2 \Psi + \sin^2 \Psi = 1$, as would be expected. Further, note that care is needed in computing the inverse tangent in order to find Ψ : it is important that the real part is divided on the left. This is because

$$\tan \Psi = (Q \cos \Psi)^{-1}(Q \sin \Psi) = \cos^{-1} \Psi Q^{-1} Q \sin \Psi = \cos^{-1} \Psi \sin \Psi \quad (5)$$

whereas with a division on the right we have:

$$\tan \Psi \neq (Q \sin \Psi)(Q \cos \Psi)^{-1} = Q \sin \Psi \cos^{-1} \Psi Q^{-1} \quad (6)$$

and Q and its inverse are at opposite ends of the product and do not cancel, in general¹⁰. (Note that $\cos \Psi$ and $\sin \Psi$ commute, so their quotient is the same whether divided on the left or right.) The exponential is a biquaternion because of the presence of I and it has a complex modulus. The modulus of this complex modulus is 1. We discuss both polar forms further in §3.1 in the context of conjugation. Finally, note that in the complex polar form, Q and the exponential do not commute. It is therefore possible to define a variant by placing Q on the right of the exponential. The variant is related to the form in Table 2 by the conjugate rule.

De Leo and Rodrigues [8] discussed polar forms of biquaternions but apparently did not see that there were two simple polar forms as here, with different imaginary units. Instead they described a single polar form containing the product of two exponentials. We can do this with either of our polar

⁹The cosine and sine of a complex angle are simply defined in terms of Euler's formula as $\cos z = \frac{1}{2}(e^{Iz} + e^{-Iz})$ and $\sin z = -\frac{1}{2}I(e^{Iz} - e^{-Iz})$; or, if we write $z = x + Iy$: $\cos z = \cos x \cosh y - I \sin x \sinh y$ and $\sin z = \sin x \cosh y + I \cos x \sinh y$ [5, See **complexes**].

¹⁰It is of course possible for Q and Ψ to commute in special cases. When this is the case it does not matter whether the inverse tangent is computed by dividing the real part on the left or right, since the result will be the same. Since the left division will give the correct result in all cases, this is the one that should be implemented numerically.

forms, representing the ‘modulus’ in each case in its own polar form. Thus the ‘Hamilton’ polar form can be written:

$$q = R \exp(\xi \Theta) = r \exp(I\phi) \exp(\xi \Theta) \quad (7)$$

where $r, \phi \in \mathbb{R}$ are the modulus and argument of R , the complex ‘modulus’ of q . Note that, because I and ξ commute, the two exponentials commute, and it is possible to write this polar form as:

$$q = r \exp(I\phi + \xi \Theta) \quad (8)$$

where the argument of the exponential is now a biquaternion with $I\phi$ as scalar part and $\xi \Theta$ as vector part. (In general, with quaternions as well as biquaternions, $e^p e^q \neq e^{pq}$ because of non-commutativity.)

Similarly, the ‘complex’ polar form can also be written in this way, expanding the quaternion ‘modulus’, Q , into the standard polar form of a quaternion as given in Table 1:

$$q = Q \exp(I\Psi) = r \exp(\mu\theta) \exp(I\Psi) \quad (9)$$

where $\mu \in \mathbb{H}$ and $r, \theta \in \mathbb{R}$. Notice that the single real modulus r is the same in each case, but the various ‘angles’ are different in value and type ($\phi, \theta \in \mathbb{R}, \Theta \in \mathbb{C}, \Psi \in \mathbb{H}$). The real modulus, r , is the absolute value of the square root of the semi-norm as discussed in § 3.3. In this case, the arguments of the two exponentials (specifically μ and Ψ) do not commute *in general*, hence the two exponentials cannot be combined by adding $\mu\theta$ to $I\Psi$, and neither can the order of the exponentials be changed. In the special case when $\mu\Psi = \Psi\mu$, the vector part of Ψ must be a (real or complex) scalar multiple of μ (commuting quaternions must have the same axis). Combining the exponentials and separating Ψ into scalar and vector parts, we may therefore write the quaternion as:

$$q = r \exp(\mu\theta + I(S(\Psi) + \mathbf{V}(\Psi))) \quad (10)$$

$$= r \exp(I S(\Psi)) \exp(\mu\theta + I \mathbf{V}(\Psi)) \quad (11)$$

and writing $\mathbf{V}(\Psi)$ as $\alpha\mu$, $\alpha \in \mathbb{C}$:

$$= r \exp(I S(\Psi)) \exp(\mu(\theta + I\alpha)) \quad (12)$$

we find that the axis of the quaternion q is the real unit quaternion μ which is the axis of Q (the quaternion ‘modulus’ in the polar form). The complex semi-norm of the quaternion is clearly $r \exp(I S(\Psi))$, and the complex angle is $\theta + I\alpha$.

We can usefully combine the ‘complex’ representation of a quaternion from Table 1 with ‘complex’ form II from Table 2 to give the following representation, which was used in [33] to derive the biquaternion roots of -1 :

$$q = (\alpha + \mu\beta) + I(\gamma + \nu\delta) \quad (13)$$

in which μ and ν are real pure unit quaternions, and α, β, γ and δ are real. The four real coefficients in this representation may be related to the

coefficients in the other representations given above. For example: $\alpha + I\gamma = A = W$. The correspondence between ξ , and μ and ν is not so simple. Equating the vector part of (13) with the vector part of ‘complex’ form I, we get:

$$\mu\beta + I\nu\delta = \xi B = \xi\sqrt{X^2 + Y^2 + Z^2} \quad (14)$$

and we can see that dividing $\mu\beta + I\nu\delta$ by its (complex) modulus B , will yield ξ *provided that B does not vanish, as discussed in § 3.6.2.*

We may relate each of the terms in (13) to a concept from geometric algebra [23, 37], as shown in Table 3 with equivalent definitions based on various representations from Table 2.

TABLE 3. Correspondence between geometric algebra concepts in $Cl_{3,0}$ and biquaternion components.

Geometric algebra concept	Elements of equation 13	Corresponding elements of representations in Table 2
Scalar – undirected quantity	α	$\Re(W) = \Re(A) = S(q_r) = w_r$
Bivector – directed area	$\mu\beta$	$\Re(\xi B) = \Re(\mathbf{V}(q)) = \mathbf{V}(q_r)$
Vector – directed magnitude	$I\nu\delta$	$I\Im(\xi B) = I\Im(\mathbf{V}(q)) = I\mathbf{V}(q_i)$
Pseudoscalar – undirected volume	$I\gamma$	$I\Im(W) = I\Im(A) = IS(q_i) = Iw_i$

These equivalences are given by Ward [42, § 3.2, p 112] and they are discussed in more detail in § 3.7. Note carefully that geometric *vectors* are represented by imaginary pure quaternions, and that real pure quaternions are *bivectors*. This is because the product of two perpendicular vectors must yield a bivector. The product of two bivectors gives a bivector (and a scalar, unless the bivectors are perpendicular).

The representation in (13) was used in [33] to derive the solutions of the equation $q^2 = -1$ (that is the biquaternion roots of -1) when q is a biquaternion, and it was shown that the solutions required $\alpha = \gamma = 0$, $\mu \perp \nu$, and $\beta^2 - \delta^2 = 1$. We return to this result in Theorem 2.

The biquaternion algebra $\mathbb{B}$ is not a division algebra because non-zero biquaternions exist that lack a multiplicative inverse. The set of such biquaternions is known as the *divisors of zero* [34] and is defined in § 3.6.

3.1. Conjugates

The conjugate of a biquaternion may be defined exactly as for a quaternion by negating the vector part. Thus we have

$$\bar{q} = W - Xi - Yj - Zk = S(q) - \mathbf{V}(q) = \bar{q}_r + I\bar{q}_i \quad (15)$$

We call this the ‘Hamiltonian’ or *quaternion conjugate*, in agreement with Synge [40, Equation 3.4, p 8]. The conjugate rule $\bar{q}\bar{p} = \overline{pq}$ and its generalization to more than two quaternions applies equally to biquaternions. Similarly, a biquaternion commutes with its quaternion conjugate, and the product of the two is the semi-norm, as discussed in § 3.3.

However, there is another possible conjugate – that obtained by taking the complex conjugate of the complex components of the quaternion. We denote this complex conjugate by a superscript star, which is a common convention with complex numbers. Thus the *complex conjugate* is given by:

$$q^* = W^* + X^*i + Y^*j + Z^*k = S(q)^* + \mathbf{V}(q)^* = q_r - Iq_i \quad (16)$$

Our definition in equation 16 agrees with that of Synge [40, p8, equation 3.4] and appears to be the obvious way to define the complex conjugate, but it differs from that of Ward [42] who defines a complex conjugate which is a quaternion conjugate with complex conjugate components. There is no way to define the complex conjugate in terms of additions and multiplications as can be done with the quaternion conjugate — if a means existed, it would also apply to complex numbers because a degenerate biquaternion of the form $\alpha + I\gamma$ with $\alpha, \gamma \in \mathbb{R}$, is isomorphic to a complex number. Instead the complex conjugate must be seen as a fundamental operation. Note very carefully that a biquaternion *does not commute* with its complex conjugate, a surprising result that we examine in Proposition 1 below.

Of course, we can apply both conjugates [42] (we call this a *total conjugate*, but the term *biconjugate* has also been suggested in [12], and we consider it apposite). It is not difficult to show the following results:

$$\overline{q^*} = \bar{q}^* \quad (17)$$

$$(pq)^* = p^*q^* \quad (18)$$

$$\overline{(pq)^*} = \overline{q^*p^*} \quad (19)$$

In terms of the various representations above, the total or biconjugate is:

$$\begin{aligned} \overline{q^*} &= w_r - x_r i - y_r j - z_r k - I(w_i - x_i i - y_i j - z_i k) \\ &= \overline{q_r} - I \overline{q_i} \\ &= S(q)^* - \mathbf{V}(q)^* \end{aligned} \quad (20)$$

However, the ramifications of non-commutativity are deep, as shown by the following Proposition.

Proposition 1. *A biquaternion does not, in general, commute with its complex conjugate: reversing the order of the product yields a complex conjugate result.*

Proof. Represent an arbitrary biquaternion in complex form I, as given in Table 2: $q = q_r + Iq_i$. Then the two products of q with its complex conjugate q^* are:

$$q q^* = (q_r + Iq_i)(q_r - Iq_i) = q_r^2 + q_i^2 + I(q_i q_r - q_r q_i) \quad (21)$$

$$q^* q = (q_r - Iq_i)(q_r + Iq_i) = q_r^2 + q_i^2 - I(q_i q_r - q_r q_i) \quad (22)$$

The results are a complex conjugate pair, and therefore differ. $\square$

Note that important exceptions to Proposition 1 are:

- Quaternions, that is biquaternions with $q_i = 0$ (trivial).
- Imaginary biquaternions, with $q_r = 0$. In this case the product is q_i^2 , regardless of order.
- Biquaternions with real and imaginary parts that commute (co-planar, with a common axis). In this case the imaginary part of the product vanishes and the result is $q_r^2 + q_i^2$, regardless of order.

The Hamilton and complex conjugates correspond to negation of the argument of the exponential in the two polar forms in Table 2. In the ‘Hamilton’ polar form, negating the argument of the exponential negates the sine term, which corresponds to the vector part of q , and thus yields the classical Hamilton conjugate. In the ‘complex’ polar form, negating the argument of the exponential again negates the sine term, but in this case the sine term corresponds to the imaginary part of the quaternion, and thus it is the complex conjugate that is obtained. A third possibility would be a polar form that would correspond (in the sense just outlined) to the total conjugate. However, this would require the sine term to represent all the components in (20) except w_r , which would be represented in the cosine term. This appears improbable.

The three types of conjugate allow us to construct formulae for extracting the geometric components of a biquaternion, as defined in Table 3. The formulae in Table 4 are obtained by substitution from four formulae for the scalar/vector and real/imaginary parts of a quaternion based on sums and differences of the Hamilton or complex conjugates. Unlike the formulae in equation 3 however, which are based on involutions, and therefore require multiplication and addition only, these formulae contain complex conjugates, which are primitive operations that cannot be expressed in terms of multiplications and additions. Similarly, and fairly trivially, by combining the formulae in equation 3 with formulae based on sums and differences of complex conjugates, one can obtain formulae for extracting any of the eight real components of a biquaternion¹¹. For example, that for the real part of the scalar part is $w_r = \frac{1}{8}(q + q^* - iq i - iq^* i - jq j - jq^* j - kq k - kq^* k)$. Finally we note that the operations represented by the formulae in Table 4 are known in geometric algebra as *ungrading* operations which return the different ‘grades’ within a multivector [30, § 6.1.3].

3.2. Inner Product

The definition of the inner product in the classic quaternion case was given by Porteous [32, Prop. 10.8, p 177] and we take this as the basis for defining the inner product of two biquaternions, since it is consistent with the quaternion case, and also yields the semi-norm (see the next section) in the case of the

¹¹The utility of such formulae is not in numeric computation, where explicit access to the components is a much better approach, but in algebraic manipulation, where the formulae may allow an algebraic solution to an otherwise seemingly difficult derivation.

TABLE 4. Formulae for the geometric components of a biquaternion. (See also Table 3.)

Scalar	$\Re(S(q))$	$\frac{1}{4} (q + \bar{q} + q^\star + \bar{q}^\star)$
Bivector	$\Re(\mathbf{V}(q))$	$\frac{1}{4} (q - \bar{q} + q^\star - \bar{q}^\star)$
Vector	$\Im(\mathbf{V}(q))I$	$\frac{1}{4} (q - \bar{q} - q^\star + \bar{q}^\star)$
Pseudoscalar	$\Im(S(q))I$	$\frac{1}{4} (q + \bar{q} - q^\star - \bar{q}^\star)$

inner product of a biquaternion with itself. The same formula is given by Synge [40, Equation 3.8, p 9] in the context of a discussion of basic properties of biquaternions. Ward, in his discussion of biquaternions and the Minkowski metric [42, § 3.3, p114], utilises the same definition, but he also discusses other definitions of the inner product (see [42, § 3.2, p109]). The definitions of Porteous and Synge are:

$$\langle p, q \rangle = \frac{1}{2} (\bar{p}q + \bar{q}p) = \frac{1}{2} (p\bar{q} + q\bar{p}) \quad (23)$$

where the overbar represents a (Hamilton or quaternion) conjugate. The result of this expression will have zero vector part (since the vector parts of the two terms inside the parentheses cancel). In general, the scalar part of the result, and therefore the inner product, will be complex. The inner product may also be defined in terms of a simple elementwise product of the two quaternions, as can be seen by expanding out (23). If we let $p = W_p + X_p\mathbf{i} + Y_p\mathbf{j} + Z_p\mathbf{k}$ and $q = W_q + X_q\mathbf{i} + Y_q\mathbf{j} + Z_q\mathbf{k}$, then

$$\langle p, q \rangle = W_p W_q + X_p X_q + Y_p Y_q + Z_p Z_q \quad (24)$$

which is, of course, complex¹².

As already discussed in § 3, the use of the term *inner product* here does not imply that all the usual properties of an inner product will be satisfied. Ward [42, § 3.3, p 115] states the following properties that are satisfied by the inner product defined in (23), and these are easily verified:

$$\langle p, q \rangle = \langle q, p \rangle, \quad p, q \in \mathbb{B} \quad (25)$$

$$\langle p, q + r \rangle = \langle p, q \rangle + \langle p, r \rangle, \quad r \in \mathbb{B} \quad (26)$$

$$\alpha \langle p, q \rangle = \langle \alpha p, q \rangle = \langle p, \alpha q \rangle, \quad \alpha \in \mathbb{C} \quad (27)$$

Conventionally, but not always [31, See: **Scalar product**], the inner product is *positive definite*, that is greater than or equal to zero, or non-negative, which cannot apply to a complex-valued inner product.

Classically, the inner product of two vectors is given by $|v_1||v_2|\cos\theta$ where θ is the angle between the two vectors. If extended to quaternions, it is not difficult to see that the angle is that between the two quaternions in 4-space, in the common plane defined by the two quaternions. When we

¹²The result will be equal to the scalar part of the quaternion with zero vector part resulting from (23).

extend this concept to biquaternions, the angle between them, and their moduli, must be complex, in general, so the geometric interpretation of the inner product is slightly more difficult. In the quaternion case, orthogonality arises from the angle between the two quaternions (having a vanishing cosine), but in the biquaternion case there is the additional possibility that one or both moduli are zero, resulting in a vanishing inner product. Expressing the inner product of two biquaternions in terms of inner products of the real and imaginary parts makes it possible to understand what the inner product represents geometrically. Representing p as follows: $p = p_r + Ip_i$ and similarly for q , the inner product may be expanded as:

$$\langle p, q \rangle = \langle p_r, q_r \rangle - \langle p_i, q_i \rangle + I(\langle p_r, q_i \rangle + \langle p_i, q_r \rangle) \quad (28)$$

Certain special cases are apparent after careful inspection:

- $\langle p, q \rangle = 0$ indicates that the two biquaternions are orthogonal or ‘perpendicular’, and it requires the real and imaginary parts of the inner product to vanish separately. The detailed conditions required for this are many, since the four inner product components (two real, two imaginary) may have positive or negative real values and can cancel in several ways. We list here four different ways in which the scalar product can vanish:

Strongest constraint. This is when all four real and imaginary parts of the two biquaternions are mutually orthogonal. A simple example is $p = 1 + I\mathbf{i}$ and $q = \mathbf{j} + I\mathbf{k}$, but it is easy to construct a more general example starting from a random quaternion. Let p_1 be a randomly chosen (pure or full) quaternion. Then let $p_2 = p_1\mathbf{i}$, $p_3 = p_1\mathbf{j}$ and $p_4 = p_1\mathbf{k}$ ¹³. The four quaternions thus constructed have the same four numeric components permuted with sign changes in such a way that any two will have a vanishing inner product¹⁴. Two biquaternions $p = p_1 + Ip_2$ and $q = p_3 + Ip_4$ (or any other permutation) will have $\langle p, q \rangle = 0$. This pair of biquaternions will be found to be divisors of zero (see § 3.6), but scaling the real and imaginary parts with different scale factors will result in biquaternions that are not divisors of zero, but are still orthogonal (orthogonality not being dependent on norm). Choosing the initial random value to be pure does not result in the pair of biquaternions being pure, because of permutation of the components.

Weaker constraint I.

1. The real parts of the two biquaternions are orthogonal, and the imaginary parts of the two biquaternions are orthogonal, hence the real part of the scalar product vanishes because it is the difference of two vanishing scalar products.

¹³Any set of three mutually orthogonal unit pure quaternions may be used here in place of $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$, and similarly in the next two cases.

¹⁴Three other permutations can be obtained by multiplication on the right, or by division on either side by $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$.

2. The real part of each biquaternion is not orthogonal to the imaginary part of the other. However, the scalar products of the real part of one biquaternion with the imaginary part of the other have the same values, but *opposite signs*. This means the imaginary part of the scalar product of the two biquaternions vanishes *by cancellation*.

A pair of biquaternions satisfying this constraint can be obtained from an arbitrary biquaternion p , by choosing $q = p\mathbf{i}$ (and two others also orthogonal to the first may be constructed by choosing $q = p\mathbf{j}$ or $q = p\mathbf{k}$)¹⁵.

Weaker constraint II.

1. The real parts of the two biquaternions are not orthogonal, and neither are the imaginary parts, but the scalar product of the two real parts has the same value *and sign* as the scalar product of the two imaginary parts. This means the real part of the scalar product of the two biquaternions vanishes *by cancellation*.
2. The real part of each biquaternion is orthogonal to the imaginary part of the other, hence the imaginary part of the scalar product of the two biquaternions vanishes because it is the sum of two vanishing scalar products.

A pair of biquaternions satisfying this constraint can be obtained from an arbitrary biquaternion p , by choosing $q = pI\mathbf{i}$ (and ditto, *mutatis mutandis*).

Weakest constraint. This is when none of the real and imaginary parts of the two biquaternions are orthogonal, but the real and imaginary parts of the scalar product of the two biquaternions vanish by cancellation. A pair of biquaternions satisfying this constraint can be constructed from an arbitrary biquaternion p , by choosing $q = p(\mathbf{i} + I\mathbf{j})$ or similar. Note that this results in q being a divisor of zero (see § 3.6) because $(\mathbf{i} + I\mathbf{j})$ is a divisor of zero.

- $\langle p, q \rangle$ is real, that is with zero imaginary part. There is a trivial case: the two biquaternions may be imaginary with zero real parts. Otherwise, the imaginary part of the inner product can vanish in two ways, as in the weaker cases above of $\langle p, q \rangle = 0$: either the imaginary part vanishes by cancellation, or it vanishes because the real part of each biquaternion is orthogonal to the imaginary part of the other. It is easy to construct a biquaternion to satisfy the second condition using the technique outlined above for the strongest constraint: construct two pairs of orthogonal quaternions, and then construct two biquaternions using the components of these pairs.
- $\langle p, q \rangle$ is imaginary, that is with zero real part. The conditions required for this are very similar to the previous case, with the appropriate

¹⁵As in the stronger case, it is also possible to multiply by $\mathbf{i}$ on the left, *etc*.

changes. A trivial case is where one biquaternion is real and the other is imaginary.

Clearly, from the above analysis, *orthogonality* of biquaternions is not as simple as orthogonality of real quaternions where a geometrical interpretation in terms of a plane and a real angle is straightforward. In the biquaternion case, interpreting the inner product $\langle p, q \rangle$ as $|p| |q| \cos \Theta$, where the two moduli and the angle are complex, is not at all obvious. This remains a topic for further work.

3.3. Semi-norm

In this section we consider the generalization of the quaternion norm $\|q\| = w^2 + x^2 + y^2 + z^2$, $\|q\| \in \mathbb{R}$ to the biquaternion case by allowing the four Cartesian components to become complex. Conventionally a *norm* is real and positive definite, that is non-negative, but in the case of biquaternions, this convention must be relaxed because the norm has a complex value. Additionally, the existence of divisors of zero (see § 3.6) means that a non-zero biquaternion can have a vanishing norm. For this reason, we use the term *semi-norm*. A semi-norm is a generalization of the concept of a norm, with no requirement that the norm be zero only at the origin (of a vector space) [4, See: **semi-norm**]. As with our use of the term *inner product* in § 3.2 for a complex-valued generalization of the inner product, we are here extending the term *semi-norm* to a complex-valued quantity analogous to a norm, with the additional property of vanishing in the case of divisors of zero. We also use the term *modulus* for the square root of the semi-norm – and since the semi-norm is complex, we define the modulus to be complex, in general (in some cases it may be real).

The semi-norm can be real and negative in special cases, as well as purely imaginary. The semi-norm can be defined in terms of the inner product of a biquaternion with itself, that is $\|q\| = \langle q, q \rangle$ or directly in terms of the four complex components: $\|q\| = |q|^2 = W^2 + X^2 + Y^2 + Z^2$. Although the semi-norm is complex-valued the result given by Coxeter for quaternions [7, § 2] still holds: $\|q\| = |q|^2 = q \bar{q} = \bar{q} q$. This is, of course, a special case of the formula for the inner product, in this case of q with itself (23).

Lemma 1. [34, Lemma 1] *Let $q = q_r + Iq_i$ be a non-zero biquaternion with real part $q_r \in \mathbb{H}$ and imaginary part $q_i \in \mathbb{H}$. The real part of $\|q\|$ is equal to the difference between the norms of q_r and q_i , and the imaginary part of $\|q\|$ is equal to twice the inner product of q_r and q_i :*

$$\|q\| = \|q_r\| - \|q_i\| + 2I \langle q_r, q_i \rangle \quad (29)$$

Proof. We express the semi-norm of q as the product of q with its quaternion conjugate: $\|q\| = q \bar{q}$. Writing this explicitly:

$$\|q\| = (q_r + Iq_i) \overline{(q_r + Iq_i)} = (q_r + Iq_i)(\bar{q}_r + I\bar{q}_i) \quad (30)$$

and multiplying out we get:

$$\|q\| = q_r \overline{q_r} - q_i \overline{q_i} + I(q_r \overline{q_i} + q_i \overline{q_r}) \quad (31)$$

The real part of $\|q\|$ can be recognised as $\|q_r\| - \|q_i\|$ from Coxeter's result [7, § 2]. The imaginary part of $\|q\|$ is twice the inner product of q_r and q_i as given by Porteous [32, Prop. 10.8, p 177] and (23). $\square$

There are some special cases of the result given in Lemma 1:

- biquaternions with perpendicular real and imaginary parts have a real semi-norm (because the imaginary part vanishes), but the semi-norm may be negative if the norm of the imaginary part exceeds the norm of the real part;
- biquaternions with real and imaginary parts with equal norms have an imaginary semi-norm (because the real part vanishes).
- imaginary biquaternions (with $q_r = 0$) have a negative real semi-norm, and therefore an imaginary modulus.
- biquaternions with perpendicular real and imaginary parts with equal norms have a vanishing semi-norm. This important special case is covered in more detail in § 3.6.

The semi-norm is invariant under quaternion conjugation, that is $\|\overline{q}\| = \|q\|$, but under complex conjugation the semi-norm of the conjugate equals the complex conjugate of the semi-norm: $\|p^*\| = \|p\|^*$. This can be seen easily from Lemma 1, since quaternion conjugation does not affect the norms of the real and imaginary parts, nor their inner product. However, although complex conjugation does not alter the norms of the real and imaginary parts, it does negate their inner product and therefore negates the imaginary part of the semi-norm.

The semi-norm as defined here (and by Ward [42, § 3.3, p 115]) obeys the ‘rule of the norms’¹⁶ so that for any two biquaternions, the semi-norm of their product equals the product of their semi-norms: $\|pq\| = \|p\| \|q\|$. This is true even in the case where one or both of p and q is a divisor of zero, with vanishing semi-norm, as discussed in § 3.6 (but the result will be zero, computed as the semi-norm of the product, or the product of the semi-norms). Hence $\mathbb{B}$ is a *normed algebra* using the definition of the semi-norm used in this section. The non-zero biquaternions that have zero semi-norm (the *divisors of zero*) lack a multiplicative inverse.

We now consider the properties of the modulus, or square root of the semi-norm as defined above. What we here call the modulus is usually required to satisfy the following three properties [4, See: **norm**]¹⁷. In fact only

¹⁶The ‘rule of the norms’ holds also for the modulus (the square root of the semi-norm), which is what the term ‘norm’ refers to in many sources in the literature.

¹⁷The cited reference states these properties using the notation $\|x\|$ where we use $|x|$ and it uses the term ‘norm’ to mean what is here called ‘modulus’.

the second of these is valid for biquaternions.

$$|\lambda x| = |\lambda| |x| \quad (\text{Not valid for } x \in \mathbb{B}) \quad (32)$$

$$|-x| = |x| \quad (33)$$

$$|x + y| \leq |x| + |y| \quad (\text{Triangle inequality – not applicable for } x \in \mathbb{B}) \quad (34)$$

The first of these properties does not hold for biquaternions because λ is usually taken to be a scalar, that is in the biquaternion context, a complex number. The formula assumes that $|\lambda|$ equals λ apart from sign, which is not true for complex numbers. However, a similar formula for the semi-norm does apply to biquaternions: $\|\lambda x\| = \lambda^2 \|x\|$. This can be shown as follows: noting that $\|q\| = \langle q, q \rangle$ for all $q \in \mathbb{B}$ and using the third property of the inner product given in (27), we have that $\lambda^2 \langle q, q \rangle = \langle \lambda q, \lambda q \rangle$ and hence the result stated. However, taking the square root of both sides gives $|\lambda x| = \sqrt{\lambda^2} |x|$. The square root of the square of a complex number is neither the modulus of the complex number, nor necessarily the original complex number: the result of the square root is always in the right half-plane (because the argument is halved by the square root operation). The second of these properties is easily seen to be satisfied for all $q \in \mathbb{B}$ because negating a complex number does not change its square, and therefore does not alter the sum of the squares of the four components of a biquaternion, nor the square root of the sum of the squares. The triangle inequality, as has already been noted in § 3 is not applicable to the biquaternion semi-norm, because ordering is not defined for complex numbers.

3.4. Real norms

It is possible to define more than one norm for the biquaternions including some which are real-valued (see for example [42, §§ 3.2 and 3.3]).

There is only one possible real-valued norm which satisfies the ‘rule of the norms’. Gürlebeck and Sprössig [13, Lemma 1.30] state that there is a unique real norm satisfying $\|pq\| = \|p\| \|q\|$ and $\|\mu\sigma_0\| = |\mu|$, $\mu \in \mathbb{R}$ where σ_0 is the 2×2 identity (Pauli) matrix. This norm is simply the square root of the absolute value of the semi-norm, in other words, r in equations 7 and 9, or $|Q|$ in Table 2.

There is another way to define a real norm [42, § 3.2, p 109], which does not satisfy the ‘rule of the norms’, as we would expect, since this norm differs from that given by Gürlebeck and Sprössig:

$$\|q\| = S(q\bar{q}^\star) = \|q_r\| + \|q_i\| \quad (35)$$

$$= w_r^2 + x_r^2 + y_r^2 + z_r^2 + w_i^2 + x_i^2 + y_i^2 + z_i^2 \quad (36)$$

If we define a modulus (as the square root of this norm), the modulus satisfies the properties in (32), including the triangle inequality. Finally, note that this norm can be seen as a special case of the following real-valued inner product: $\langle p, q \rangle = \langle p_r, q_r \rangle + \langle p_i, q_i \rangle$. This is easily seen from (36).

3.5. Roots of -1

This section draws on and builds on results first presented by Sangwine in [33]. These results have been generalized to Clifford algebras by Hitzer and Ablamovicz [25] where the concept is aptly referred to as *geometric roots* of -1 .

Biquaternion roots of -1 appear as ξ in Table 2, in the ‘complex’ and ‘Hamilton polar’ forms.

Theorem 1. *Any non-zero quaternion or biquaternion with a non-vanishing modulus, divided by its modulus has a unit real norm.*

Proof. Let q be an arbitrary quaternion or biquaternion and let $p = q/|q|$. For any quaternion or biquaternion x we have $\|x\| = x\bar{x}$ and $|x| = \sqrt{\|x\|}$. Therefore $p = q/\sqrt{q\bar{q}}$ and we can write the norm of p as:

$$\|p\| = p\bar{p} = \frac{q}{\sqrt{q\bar{q}}} \overline{\left(\frac{q}{\sqrt{q\bar{q}}}\right)} \quad (37)$$

Since $\sqrt{q\bar{q}}$ is complex, it is unaffected by the quaternion conjugate and we can simplify this to:

$$= \frac{q}{\sqrt{q\bar{q}}} \frac{\bar{q}}{\sqrt{q\bar{q}}} = \frac{q\bar{q}}{q\bar{q}} = 1 \quad (38)$$

□

Theorem 2. *Any pure quaternion or biquaternion with a non-vanishing modulus, divided by its modulus, is a root of -1 .*

Proof. Let q be an arbitrary quaternion or biquaternion with $|q| \neq 0$. Then $(q/|q|)^2 = q^2/q\bar{q} = q/\bar{q}$. If we restrict q to have zero scalar part, the conjugate reduces to negation and we have $(q/|q|)^2 = q/-q = -1$. □

Theorem 2 is simply a restatement of the well-known fact that unit pure quaternions are roots of -1 [17, § 167, p179], but in the case of biquaternions, the ramifications of the result are not so obvious. Theorem 2 is also easily demonstrated using the Cartesian form:

$$\begin{aligned} \left(\frac{Xi + Yj + Zk}{\sqrt{X^2 + Y^2 + Z^2}} \right)^2 &= \begin{pmatrix} -X^2 + XY(ij + ji) \\ -Y^2 + XZ(ik + ki) \\ -Z^2 + YZ(jk + kj) \end{pmatrix} / (X^2 + Y^2 + Z^2) \\ &= \frac{-(X^2 + Y^2 + Z^2)}{X^2 + Y^2 + Z^2} = -1 \end{aligned} \quad (39)$$

It was shown in [33, Theorem 2.1] that any pure biquaternion ξ satisfying the constraints $\Re(\xi) \perp \Im(\xi)$ and $\|\Re(\xi)\| - \|\Im(\xi)\| = 1$ is a root of -1 and that no other biquaternions are roots of -1 . It follows therefore that dividing an arbitrary pure biquaternion by its (complex) modulus produces a result that satisfies the constraints stated. Since these constraints are by no means obvious from (39), it is interesting to verify them directly. To verify the

constraints stated, we take an arbitrary pure biquaternion p , divide it by its (complex) modulus and show that the result satisfies the constraints stated. Dividing p by its modulus gives a root of -1 : $\xi = p/|p|$. We have to show that the real and imaginary parts of ξ are orthogonal, that is $\langle \Re(\xi), \Im(\xi) \rangle = 0$, and also that the difference between the norms of the real and imaginary parts of ξ is 1. We can express the real and imaginary parts of ξ as follows, by taking the sum and difference of ξ with its complex conjugate:

$$\Re(\xi) = \frac{1}{2}(\xi + \xi^*), \quad \Im(\xi) = \frac{1}{2I}(\xi - \xi^*) \quad (40)$$

The inner product of two quaternions is given by (23), but since we are dealing with pure quaternions, the conjugates reduce to negation, and the inner product simplifies to:

$$\langle u, v \rangle = -\frac{1}{2}(uv + vu) \quad (41)$$

Substituting the results above for the real and imaginary parts of ξ , we obtain the following:

$$\begin{aligned} \langle \Re(\xi), \Im(\xi) \rangle &= -\frac{1}{2} \left[\frac{1}{2}(\xi + \xi^*) \frac{1}{2I}(\xi - \xi^*) + \frac{1}{2I}(\xi - \xi^*) \frac{1}{2}(\xi + \xi^*) \right] \\ &= -\frac{1}{8I} [(\xi + \xi^*)(\xi - \xi^*) + (\xi - \xi^*)(\xi + \xi^*)] \\ &= -\frac{1}{8I} [\xi^2 + \xi^* \xi - \xi^{*2} - \xi \xi^* + \xi^2 - \xi^* \xi - \xi^{*2} + \xi \xi^*] \\ &= -\frac{1}{4I} [\xi^2 - \xi^{*2}] \end{aligned} \quad (42)$$

Since ξ is a root of -1 by Theorem 2, its complex conjugate must also be a root of minus one¹⁸, thus $\xi^2 = \xi^{*2} = -1$ and therefore $\langle \Re(\xi), \Im(\xi) \rangle = 0$ as expected.

To show that the real and imaginary parts of ξ have norms differing by 1, we can use a similar approach. Starting with the fact that the norm is the inner product of a quaternion with itself, as stated in §3.3, the formula in (41) simplifies to $\|u\| = -u^2$ when u is pure and $v = u$. Thus the difference between the norms of the real and imaginary parts of ξ is given by:

$$\begin{aligned} \|\Re(\xi)\| - \|\Im(\xi)\| &= -\Re(\xi)^2 + \Im(\xi)^2 \\ &= -\left[\frac{1}{2}(\xi + \xi^*) \right]^2 + \left[\frac{1}{2I}(\xi - \xi^*) \right]^2 \\ &= -\frac{1}{4} [\xi^2 + \xi^{*2} + \xi \xi^* + \xi^* \xi] - \frac{1}{4} [\xi^2 + \xi^{*2} - \xi \xi^* - \xi^* \xi] \\ &= -\frac{1}{2} [\xi^2 + \xi^{*2}] \end{aligned} \quad (43)$$

which is 1 because ξ and its conjugate are roots of -1 .

¹⁸This is easily shown using the ‘complex’ form $\xi = \xi_r + I\xi_i$, since the imaginary part must be zero in both ξ^2 and ξ^{*2} if they are each equal to -1 .

3.6. Divisors of zero

It is possible for the semi-norm, $\|q\|$, of a biquaternion to be zero, even though $q \neq 0$. In modern terminology, the biquaternions with vanishing semi-norm are better known as *divisors of zero*. The ‘rule of the norms’ given in § 3.3 shows that multiplying an arbitrary biquaternion by a divisor of zero (with vanishing semi-norm) yields a result which is also a divisor of zero (with vanishing semi-norm).

The conditions for the semi-norm to vanish were discovered by Hamilton [17, Lecture VII, § 672, p669]¹⁹ — if we represent a non-zero biquaternion in the form: $q = q_r + Iq_i$ then its norm is zero iff $\|q_r\| = \|q_i\|$ and $\langle q_r, q_i \rangle = 0$, that is, the real and imaginary parts must have equal norms and be perpendicular in 4-space. This result follows easily from (29) in Lemma 1 (§ 3.3). The conditions stated are equivalent to the condition $W^2 + X^2 + Y^2 + Z^2 = 0$, which can be expanded into a pair of simultaneous equations each equivalent to one of the two conditions above, by equating the real and imaginary parts to zero separately [34, Proposition 1]. Notice that the conditions required for a vanishing semi-norm may be satisfied in the case of a pure biquaternion (the real and imaginary components are then perpendicular in 3-space). It follows from these conditions that any non-zero quaternion or any non-zero imaginary biquaternion must have a non-zero norm, since the constraint of real and imaginary parts having equal norm is violated.

Sangwine and Alfsmann have shown in [34] that every biquaternion divisor of zero is one of:

- an idempotent as defined in the next section in Theorem 3;
- a complex multiple of an idempotent (dividing the divisor of zero by twice its scalar part yields an idempotent);
- a nilpotent as defined in § 3.6.2.

3.6.1. Idempotents. An idempotent is a value that squares to give itself. An example of a biquaternion idempotent is given in [12] (quoting a paper of Lanczos from 1929) but no general case is given in the paper. More recently biquaternion idempotents were analysed in [1, § 7.1] using a matrix representation of the isomorphic Clifford algebra $\mathcal{Cl}_{3,0}$.

We present here a derivation of the set of idempotents in $\mathbb{B}$, drawing on results presented in [34] by Sangwine and Alfsmann.

Theorem 3. [34, Theorem 2] *Any biquaternion of the form $q = \frac{1}{2} \pm \frac{1}{2} \xi I$, where $\xi \in \mathbb{B}$ is a root of -1 , is an idempotent. There are no other idempotents in $\mathbb{B}$.*

Proof. The proof is by construction from an arbitrary biquaternion represented in the form $q = A + \xi B$ ($A, B \in \mathbb{C}$). Squaring and equating to q gives:

$$q^2 = A^2 - B^2 + 2AB\xi = A + \xi B = q \quad (44)$$

¹⁹In [17, Lecture VII, § 673, p671], Hamilton rather quaintly refers to an *evanescent tensor* — a tensor being what we here call a semi-norm.

Equating coefficients of 1 and ξ (that is, equating the real and imaginary parts with respect to ξ) gives:

$$A = A^2 - B^2 \quad (45)$$

$$B = 2AB \quad (46)$$

The second of these equations requires that $A = \frac{1}{2}$. Making this substitution into the first equation gives: $\frac{1}{2} = \frac{1}{4} - B^2$, from which $B^2 = -\frac{1}{4}$. Since B is complex, the only solutions are $B = \pm I/2$. Thus $q = \frac{1}{2} \pm \frac{\xi I}{2}$. $\square$

The roots of -1 in $\mathbb{B}$ include two trivial cases, as noted in [33, Theorem 2.1]. These are $\pm I$ which yields the trivial idempotents $q = 0$ and $q = 1$; and $\pm \mu$ where μ is a real pure quaternion root of -1 .

Since every biquaternion idempotent, q , is a solution of $q^2 = q$, it must also be a solution of $q(q - 1) = 0$. Therefore every biquaternion idempotent is also a divisor of zero and must satisfy the conditions stated in § 3.6. This is demonstrated in detail in [34, § 4, Proposition 2].

All non-pure biquaternion divisors of zero have been shown to be of the form $p = \alpha q$ where α is a non-zero complex number, and q is a biquaternion idempotent [34, § 4, Theorem 3]. Further, dividing a non-pure biquaternion divisor of zero by twice its scalar part yields an idempotent [34, § 4, Corollary 1].

3.6.2. Nilpotents. A nilpotent is a quantity whose square vanishes. Hamilton discovered nilpotent biquaternions and was aware that all such biquaternions were divisors of zero [17, Lecture VII, § 674, pp671-3].

We first present the case of a pure biquaternion with vanishing semi-norm and show a result due to Hamilton [17, Lecture VII, § 672, p669] that the square of such a biquaternion vanishes.

Theorem 4. *Let $p = p_r + Ip_i$ be a non-zero pure biquaternion with vanishing semi-norm. Then $p^2 = 0$.*

Proof. From the fact that p has a vanishing semi-norm, $\|p_r\| = \|p_i\|$, therefore we may divide by their common norm, and obtain $p/\|p_r\| = \mu + I\nu$ where μ and ν are perpendicular unit pure quaternions. Then:

$$\left(\frac{p}{\|p_r\|}\right)^2 = (\mu + I\nu)^2 = \mu^2 - \nu^2 + I(\mu\nu + \nu\mu) \quad (47)$$

which vanishes because the squares of the two unit pure quaternions are -1 and the imaginary part vanishes because the products of two perpendicular unit pure quaternions changes sign when the order of the product is reversed. $\square$

Sangwine and Alfsmann, in [34, § 5, Lemma 1] have shown that all nilpotent biquaternions are pure, and in [34, § 5, Proposition 3] that all nilpotent biquaternions are divisors of zero. It follows that all nilpotent biquaternions have real and imaginary parts with equal norm and therefore that any nilpotent biquaternion can be normalised to the form $\mu + I\nu$ as in Theorem 4. We

see therefore that all nilpotent biquaternions are constructed from a pair of mutually orthogonal unit pure quaternions.

All nilpotent biquaternions are pure. *Non-nilpotent* pure biquaternions may be expressed in the form $q = \xi B$ where ξ is a root of -1 and B is complex (the square root of the semi-norm: $B = \sqrt{X^2 + Y^2 + Z^2}$). Since a nilpotent biquaternion q is a divisor of zero, its semi-norm, and therefore B , is zero, even though $q \neq 0$. A consequence is that it is not possible to compute the ‘axis’ ξ of a nilpotent biquaternion q by dividing q by B .

3.7. The biquaternions as a geometric algebra

In this section we discuss equivalences between the biquaternions and geometric algebra [23, 24, 29, 37, 39]. In particular, since there are many geometric algebras of different dimension, we consider the particular case of the Clifford algebra $Cl_{3,0}$ to which the biquaternion algebra is isomorphic.

Hamilton himself, in his *Lectures on Quaternions* [17, Lectures I, II, III] was interested in an algebra of points, lines and planes, and clearly was motivated by a desire to have an algebra applicable to 3-dimensional geometry. The modern notion of a geometric algebra uses different language, and some of the concepts are more sophisticated, but essentially the aim is the same: to represent and manipulate geometric objects algebraically.

The geometric algebra $Cl_{3,0}$ (and therefore the biquaternion algebra, through isomorphism) contains four types of element, as already introduced in § 3 and Tables 3 and 4: scalars, vectors, bivectors and pseudoscalars. Scalars are quantities without geometric form, vectors represent directed magnitudes, bivectors represent directed areas, and pseudoscalars represent signed volumes. The product of two perpendicular vectors is a bivector, representing the area swept out by one vector when moved along the other [39, § 2.1]. This is why vectors are identified with pure imaginary biquaternions. Given two perpendicular pure quaternions, μ and ν , such that $\langle \mu, \nu \rangle = 0$, the two vectors $I\mu$ and $I\nu$ have a product $-\mu\nu$ which is perpendicular to both μ and ν . The negative sign is a consequence of squaring I . Changing the order of the two vectors changes the sign of the resulting bivector (it is oppositely directed), so the minus sign is of little consequence. It is usual in geometric algebra to have to choose the sense of the product of two vectors by convention and this applies also to the quaternions/biquaternions, where the choice is between $\mathbf{i}\mathbf{j}\mathbf{k} = -1$ (the usual choice) or $\mathbf{k}\mathbf{j}\mathbf{i} = -1$, giving oppositely handed coordinate systems. The product of three orthogonal vectors defines a volume, represented by a pseudoscalar. For example, $(I\mathbf{i})(I\mathbf{j})(I\mathbf{k}) = I$, as given by Ward [42, p 113].

There has been much confusion caused by the use of the term *vector* when in fact the objects under consideration were *bivectors*. In physics the concepts of *axial* and *polar* vectors have been used, causing further confusion, because the two terms suggest different types of vector, rather than fundamentally different types of quantity. Even the present authors, in their earlier works, used the term *vector* to refer to quantities which are now clearly

TABLE 5. Biquaternion basis multiplication table.

	1	iI	jI	kI	i	j	k	I
1	1	iI	jI	kI	i	j	k	I
iI	iI	1	$-k$	j	$-I$	kI	$-jI$	$-i$
jI	jI	k	1	$-i$	$-kI$	$-I$	iI	$-j$
kI	kI	$-j$	i	1	jI	$-iI$	$-I$	$-k$
i	i	$-I$	kI	$-jI$	-1	k	$-j$	iI
j	j	$-kI$	$-I$	iI	$-k$	-1	i	jI
k	k	jI	$-iI$	$-I$	j	$-i$	-1	kI
I	I	$-i$	$-j$	$-k$	iI	jI	kI	-1

seen to be bivectors. Unfortunately, the scalar/vector part terminology commonly applied to quaternions has this wrong, and we are stuck with it: the vector part of a (bi)quaternion is simply that part which is not the scalar part – it may be a (geometric algebra) vector, a bivector or a combination of both²⁰. Similarly, we must be careful not to confuse the scalar part of a biquaternion with the geometric algebra concept of a scalar: the scalar part of a biquaternion is a complex number consisting of the scalar and pseudoscalar parts of the biquaternion.

In the geometric algebra $\mathcal{Cl}_{3,0}$ a composite quantity consisting of two or more of the four types of quantity (scalar, vector, bivector, pseudoscalar) is called a *multivector*. A biquaternion may contain all four types of geometric quantity, and except in degenerate cases (such as a pure imaginary biquaternion), corresponds to the concept of a multivector.

Although multivectors in $\mathcal{Cl}_{3,0}$ are regarded as composites of scalar, vector, bivector and pseudoscalar, a biquaternion is not constructed in this way. Instead it has four complex components, but as we have seen in § 3, it is possible to interpret these in multiple ways, for example as a complex number with quaternion components. It is also possible to consider the biquaternions as 8-dimensional quantities with basis $(1, iI, jI, kI, i, j, k, I)$ where we place the vector basis first. The multiplication table is then easily derived from the rules of quaternion multiplication (Table 5). The table is in 1:1 correspondence with the multiplication table for $\mathcal{Cl}_{3,0}$ — this is the basis of the isomorphism between the two algebras.

From the rules of quaternion multiplication, it is fairly easy to draw up a table showing how the geometric components of a biquaternion multiply together. Table 6 shows the result. The four centre entries apply in the general case, but in the specific cases where the bivector/vector are perpendicular, the scalar or pseudoscalar part of the product will be zero. Multiplication by

²⁰The same word causes another terminology problem in the field of matrices of quaternions, where a matrix with one row or column is known as a vector, using the usual matrix terminology. Of course it is nothing to do with the vector part of a quaternion, or indeed the vector part of a quaternion matrix.

TABLE 6. Multiplication table for components of a biquaternion multivector

	S	B	V	P
S – scalar	S	B	V	P
B – bivector	B	$S + B$	$P + V$	V
V – vector	V	$P + V$	$S + B$	B
P – pseudoscalar	P	V	B	S

 TABLE 7. Correspondence between components of the $\mathcal{Cl}_{3,0}$ algebra and the biquaternions.

Geometric algebra concept	Biquaternion algebra	$\mathcal{Cl}_{3,0}$ algebra
Scalar	1	1
Bivectors	$\mathbf{i}$ $\mathbf{j}$ $\mathbf{k}$	$-e_2e_3$ $-e_3e_1$ $-e_1e_2$
Vectors	$I\mathbf{i}$ $I\mathbf{j}$ $I\mathbf{k}$	e_1 e_2 e_3
Pseudoscalar	I	$e_1e_2e_3$

the inverse of the pseudoscalar²¹ is known as the *dual* operation: it maps a bivector into a perpendicular vector, and a scalar into a pseudoscalar or *vice versa*. The notion of *perpendicular* depends on an inner product [24, p 10]:

Two multivectors are said to be *orthogonal* or *perpendicular* to one another, if their inner product is zero.

However, the inner product defined in (23) is not zero for $\mathbf{j}$ and $-\mathbf{j}I$ which are a perpendicular pair of bivector and vector respectively. It is not straightforward to consider other possible inner products by simply replacing the quaternion conjugation in (23) with a complex or total conjugate, or with no conjugate. We note the following in support of this assertion, and then leave it for future work: a biquaternion does not commute with its complex conjugate. This means there are several possible variants of (23) if we replace the quaternion conjugates with complex conjugates, by varying the order of the multiplicands in each term.

The biquaternions are isomorphic to the Clifford algebra $\mathcal{Cl}_{3,0}$ [13, § 1.1.3 and p 22]. The basis vectors of $\mathcal{Cl}_{3,0}$ are conventionally written e_1 , e_2 and e_3 although sometimes other notations are used²². Table 7 shows the correspondences between the basis elements of a biquaternion and a $\mathcal{Cl}_{3,0}$ multivector. Notice that the Clifford basis vectors are multiplied cyclically to

²¹The inverse of the pseudoscalar in the biquaternion algebra is $1/I$, that is $-I$.

²²The basis elements in [13] are written as σ_i , $i \in 0, 1, 2, 3$ otherwise known as the Pauli matrices.

give the basis bivectors. Thus the biquaternion $\mathbf{j}$ is isomorphic to the Clifford bivector $-e_3e_1 = +e_1e_3$. The quaternions are a subset of the biquaternions, and the corresponding subset in the Clifford algebra $\mathcal{Cl}_{3,0}$ is known as the *even grade* part $\mathcal{Cl}_{3,0}^+$.

3.7.1. Outer products. In geometric algebra, the algebraic or geometric product of two vectors is the sum of the scalar product and the wedge ($\wedge$) or outer product:

$$pq = p \cdot q + p \wedge q \quad (48)$$

It is possible to define an outer product for quaternions and biquaternions using a formula widely used in geometric algebra [39, Equation 3.10]:

$$\frac{1}{2}(pq - qp) \quad (49)$$

This formula gives the cross product of the vector parts (because the components of the products involving the scalar parts, and the dot product of the vector parts, commute and therefore cancel). If applied to a pair of vectors (i.e. pure biquaternions with zero real part), the result is a bivector (i.e. a pure biquaternion with zero imaginary part). Applied to a pair of bivectors (pure quaternions) the result is another bivector. Applied to a vector and a bivector, the result is a vector. This indicates that the formula as stated mixes several geometric algebra concepts. As with the various possible formulations of the inner product discussed in the previous section, we could consider variants of the formula using one or more of the three types of biquaternion conjugation, but as with the inner products, we leave this for further work.

4. Further Work

There is further interpretation needed of the biquaternions as a geometric algebra, making use of the isomorphism between biquaternions and $\mathcal{Cl}_{3,0}$. As discussed in the previous section, the inner and outer product formulae can be studied, and matched to formulae already known for $\mathcal{Cl}_{3,0}$ to give an explicit set of geometric formulae expressed in terms of $\mathbb{B}$.

The interpretation of complex angles in the polar form and in the context of the inner product also requires further work.

Acknowledgements

This paper was started in 2005 at the Laboratoire des Images et des Signaux (now part of the GIPSA-Lab), Grenoble, France with financial support from the Royal Academy of Engineering of the United Kingdom and the Centre National de la Recherche Scientifique (CNRS).

Thanks are due to Dr Sebastian Miron, who gave a short series of seminars on biquaternions in Grenoble in 2005, and thus started us on the road

to this paper. We also acknowledge the contribution made by Daniel Alfsmann who, working with one of us, and using his knowledge of hypercomplex algebras in general, greatly clarified the topic of divisors of zero [34].

Dr T. A. Ell is a Visiting Fellow at the University of Essex, funded by grant numbers GR/S58621 and EP/E010334/1 from the United Kingdom Engineering and Physical Sciences Research Council from September 2003 – September 2009.

We also thank the anonymous referee who made over 50 suggestions on the paper, many of them non-trivial, which have improved the paper considerably.

References

- [1] R. Ablamowicz, B. Fauser, K. Podlaski, and J. Rembielinski, *Idempotents of Clifford algebras*. Czechoslovak Journal of Physics, **53** (11) (2003), 949–954.
- [2] Simon L. Altmann, *Rotations, Quaternions, and Double Groups*. Oxford University Press, Oxford, 1986.
- [3] Benno Artmann, *The concept of number : from quaternions to monads and topological fields*. Ellis Horwood series in mathematics and its applications. Ellis Horwood, Halsted, Chichester, 1988. Translation of: Der Zahlbegriff, Göttingen: Vandenhoeck & Rupprecht, 1983. Translated with additional exercises and material by H.B. Griffiths.
- [4] E. J. Borowski and J. M. Borwein, editors, *Collins Dictionary of Mathematics*. HarperCollins, Glasgow, 2nd edition, 2002.
- [5] Alain Bouvier, Michel George, and Francois Le Lionnais, editors. *Dictionnaire des Mathématiques*. Quadrige/Puf, Paris, 2e edition, 2005.
- [6] William K. Clifford, *Preliminary sketch of biquaternions*. Proc. London Math. Soc., **s1-4** (1) (1871), 381–395.
- [7] H. S. M. Coxeter, *Quaternions and reflections*. American Mathematical Monthly, **53** (3) (1946), 136–146.
- [8] Stefano de Leo and Waldyr A. Rodrigues, Jr. *Quaternionic electron theory: Geometry, algebra, and Dirac's spinors*. International Journal of Theoretical Physics, **37** (6) (1998), 1707–1720. ISSN 0020-7748 (Print) 1572-9575 (Online).
- [9] T. A. Ell and S. J. Sangwine. *Quaternion involutions*. Preprint: <http://www.arxiv.org/abs/math.RA/0506034>, June 2005.
- [10] T. A. Ell and S. J. Sangwine, *Quaternion involutions and anti-involutions*. Computers and Mathematics with Applications, **53** (1) (2007), 137–143.
- [11] Todd A. Ell and Stephen J. Sangwine, *Linear quaternion systems Toolbox for Matlab®*. <http://lqstfm.sourceforge.net/>, 2007Software library, licensed under the GNU General Public License.
- [12] Andre Gsponer and Jean-Pierre Hurni, *Lanczos – Einstein – Petiau: From Dirac's equation to nonlinear wave mechanics*. Preprint, August 2005. URL <http://www.arxiv.org/abs/physics/0508036>.
- [13] Klaus Gürlebeck and Wolfgang Sprössig, *Quaternionic and Clifford Calculus for Physicists and Engineers*. John Wiley, Chichester, 1997.

- [14] H. Halberstam and R. E. Ingram, editors, *The Mathematical Papers of Sir William Rowan Hamilton*, volume III Algebra. Cambridge University Press, Cambridge, 1967.
- [15] W. R. Hamilton, *On a new species of imaginary quantities connected with the theory of quaternions*. Proceedings of the Royal Irish Academy, **2** (1844), 424–434.
- [16] W. R. Hamilton, *Researches respecting quaternions*. Transactions of the Royal Irish Academy, **21** (1848), 199–296.
- [17] W. R. Hamilton, *Lectures on Quaternions*. Hodges and Smith, Dublin, 1853. Available online at Cornell University Library: <http://historical.library.cornell.edu/math/>.
- [18] W. R. Hamilton, *Elements of Quaternions*. Longmans, Green and Co., London, 1866.
- [19] W. R. Hamilton, *On the geometrical interpretation of some results obtained by calculation with biquaternions*. In Halberstam and Ingram [14], chapter 35, pages 424–5. First published in *Proceedings of the Royal Irish Academy*, 1853.
- [20] W. R. Hamilton, *On a new species of imaginary quantities connected with the theory of quaternions*. In Halberstam and Ingram [14], chapter 5, pages 111–116. First published as [15].
- [21] W. R. Hamilton, *Researches respecting quaternions. First series (1843)*. In Halberstam and Ingram [14], chapter 7, pages 159–226. First published as [16].
- [22] W. R. Hamilton, *On quaternions*. In Halberstam and Ingram [14], chapter 8, pages 227–297. First published in various articles in *Philosophical Magazine*, 1844–1850.
- [23] David Hestenes and Garret Sobczyk, *Clifford Algebra to Geometric Calculus*. D. Reidel Publishing Company, Dordrecht, 1984.
- [24] David Hestenes, Hongbo Li, and Alyn Rockwood, *New algebraic tools for classical geometry*. In Sommer [37], chapter 1, pages 3–26.
- [25] Eckhard Hitzer and Rafał Abłamowicz. *Geometric roots of -1 in Clifford algebras $Cl_{p,q}$ with $p + q \leq 4$* . Preprint: <http://arxiv.org/abs/arxiv:0905.3019>, May 2009.
- [26] I. L. Kantor and A. S. Solodnikov, *Hypercomplex numbers, an elementary introduction to algebras*. Springer-Verlag, New York, 1989.
- [27] Philip Kelland and Peter Guthrie Tait, *Introduction to quaternions*. Macmillan, London, 3rd edition, 1904.
- [28] J. B. Kuipers, *Quaternions and Rotation Sequences*. Princeton University Press, Princeton, New Jersey, 1999.
- [29] Joan Lasenby, Anthony N. Lasenby, and Chris J. L. Doran, *A unified mathematical language for physics and engineering in the 21st century*. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, **358** (2000) 21–39.
- [30] H. Li, *Invariant Algebras and Geometric Reasoning*. World Scientific, Singapore, 2008.
- [31] David Nelson, editor, *The Penguin Dictionary of Mathematics*. Penguin Books, London, third edition, 2003.

- [32] I. R. Porteous, *Topological Geometry*. Cambridge University Press, Cambridge, second edition, 1981.
- [33] S. J. Sangwine, *Biquaternion (complexified quaternion) roots of -1*. Adv. appl. Clifford alg., **16** (1) (2006), 63–68.
- [34] S. J. Sangwine and Daniel Alfsmann, *Determination of the biquaternion divisors of zero, including the idempotents and nilpotents*. Adv. appl. Clifford alg., **20** (2) (2010), 401–410.
- [35] S. J. Sangwine and N. Le Bihan, *Quaternion polar representation with a complex modulus and complex argument inspired by the Cayley-Dickson form*. Adv. appl. Clifford alg., **20** (1) (2010), 111–120.
- [36] Stephen J. Sangwine and Nicolas Le Bihan, *Quaternion Toolbox for Matlab®*. <http://qtfm.sourceforge.net/>, 2005. Software library, licensed under the GNU General Public License.
- [37] G. Sommer, editor, *Geometric computing with Clifford algebras: theoretical foundations and applications in computer vision and robotics*. Springer-Verlag, London, UK, 2001.
- [38] A. Sudbery, *Quaternionic analysis*. Mathematical Proceedings of the Cambridge Philosophical Society, **85** (2) (1979), 199–225.
- [39] Jaap Suter, *Geometric algebra primer*. Self-published on personal website: <http://www.jaapsuter.com/data/2003-3-12-geometric-algebra/geometric-algebra.pdf>
- [40] J. L. Synge, *Quaternions, Lorentz transformations, and the Conway-Dirac-Eddington matrices*. Communications of the Dublin Institute for Advanced Studies, Series A 21, Dublin Institute for Advanced Studies, Dublin, 1972.
- [41] P. G. Tait, *Sketch of the analytical theory of quaternions*. In An elementary treatise on Quaternions, chapter VI, pages 146–159. Cambridge University Press, third edition, 1890. Chapter by ‘Prof Cayley’ (Arthur Cayley).
- [42] J. P. Ward, *Quaternions and Cayley Numbers: Algebra and Applications*. volume 403 of Mathematics and Its Applications. Kluwer, Dordrecht, 1997.

Stephen J. Sangwine
 School of Computer Science and Electronic Engineering
 University of Essex
 Wivenhoe Park
 Colchester, CO4 3SQ
 United Kingdom
 e-mail: s.sangwine@ieee.org

Todd A. Ell
 5620 Oak View Court, Savage, MN 55378-4695
 USA
 e-mail: t.ell@ieee.org

Nicolas Le Bihan
GIPSA-Lab, Département Images et Signal
961 Rue de la Houille Blanche
Domaine Universitaire, BP 46
38402 Saint Martin d'Hères cedex
France
e-mail: `nicolas.le-bihan@gipsa-lab.inpg.fr`

Received: January 4, 2010.

Revised: August 19, 2010.

Accepted: August 30, 2010.

Fast Complexified Quaternion Fourier Transform

Salem Said, Nicolas Le Bihan, and Stephen J. Sangwine, *Senior Member, IEEE*

Abstract—In this paper, we consider the extension of the Fourier transform to biquaternion-valued signals. We introduce a transform that we call the biquaternion Fourier transform (BiQFT). After giving some general properties of this transform, we show how it can be used to generalize the notion of analytic signal to complex-valued signals. We introduce the notion of hyperanalytic signal. We also study the Hermitian symmetries of the BiQFT and their relation to the geometric nature of a biquaternion-valued signal. Finally, we present a fast algorithm for the computation of the BiQFT. This algorithm is based on a (complex) change of basis and four standard complex FFTs.

Index Terms—Biquaternion Fourier transform (BiQFT), biquaternion-valued signals, fast algorithm (BiQFFT), Hermitian symmetries, hyperanalytic signal.

I. INTRODUCTION

THE aim of this paper is to introduce a tool of harmonic analysis for biquaternion-valued signals, the Biquaternion Fourier transform, or BiQFT. Our interest in biquaternion-valued signals stems from the geometric properties of biquaternions. These properties allow a biquaternion-valued signal to encode several signals of different geometric nature that might occur together. This is an interesting goal, since signals of different geometric nature that occur/propagate together arise in many domains. In electromagnetics one has to deal with an electric field (vector signal) propagating along with a magnetic field (bivector signal). In acoustics one has to deal with pressure (a pseudoscalar signal) and particle velocity (a vector signal) propagating together, and there is a similar situation in seismology for pressure and particle displacement.

A biquaternion, which is a quaternion with complex components, is a composite quantity. It has a scalar, pseudoscalar, vector and a bivector part.¹ This makes it possible to encode several coexisting/copropagating signals of different geometric nature in a single biquaternion-valued signal. We thus have, for these different signals, a single-signal representation that preserves the difference between their geometries. It is not possible to realize such a representation by simply incorporating

the signals into a kind of long-vector—such as for instance a 6-component electric and magnetic field vector—nor by using quaternions. From this point of view, the use of biquaternions seems necessary.

Biquaternions are isomorphic to the Clifford algebra $\mathcal{C}\ell(3)$. We could have used the notations and formalism of Clifford algebra in this paper. However, we made the choice of using the biquaternion formalism. This choice is made for consistency with the previous work of one of the authors [1], and also because two authors of the present paper have developed a Matlab toolbox for quaternions which has many biquaternion capabilities [2].

This paper contributes to the study of generalizations of the Fourier transform to hypercomplex number systems, which has gained interest in the last decade. The first major step was performed by Ell [3] who considered quaternion Fourier transforms in his Ph.D. thesis in 1992. Chernov [4]–[6] introduced the quaternion and Clifford-valued Fourier transform for use in fast algorithms of conventional complex-valued 2-D Fourier transforms in 1995. These ideas were then generalized by Sommer *et al.* (see [7] for a thorough review of their work) to Clifford algebras. Shortly after the work of Chernov, the problem of quaternion Fourier transforms was studied by Sangwine [1] (later in collaboration with Ell) and in parallel by Pei [8] for the definition of a Fourier transform for color images.

The transform we propose in the present work, the BiQFT, will be applied to signals with biquaternion-valued samples. This is a significant difference from previous work where the components of the signals considered took values over the reals. The BiQFT generalizes many of the interesting properties of the usual Fourier transform to biquaternion-valued signals. Furthermore, we will see that it can be used to define a generalization of the notion of analytic signal to complex-valued signals. Indeed, we will use it to introduce the notion of *hyperanalytic signal*. We will also study the Hermitian symmetries of the BiQFT. We will arrive at the interesting result that signals with different geometric natures transform, via the BiQFT, to spectra with different (and complementary) Hermitian symmetries. This means that the BiQFT preserves, in the spectral domain, the separation that exists in the time domain between signals with different geometries. In this sense, the BiQFT is adapted to the motivation for its introduction mentioned at the beginning of this section.

In the following, it will be possible to note that the BiQFT is different from the ones developed by Bülöw [9], who proposed a quaternion Fourier transform for 2-D scalar-valued images, and by Felsberg [10] who defined a Clifford Fourier transform for N -dimensional scalar-valued signals, and by Ebling and Scheuermann [11] who proposed a Clifford Fourier transform (with bivector or pseudo-scalar exponential) to analyze vector-field images. In particular, our approach is different from this last example in the definition of the exponential kernel. In fact, the transform of Ebling and Scheuermann is a special case

Manuscript received October 5, 2006; revised September 6, 2007. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Markus Püschel.

S. Said and N. Le Bihan are with the GIPSA-Lab, Département Images et Signal, 38 402 Saint Martin d'Hères Cedex, France (e-mail: salem.said@gipsa-lab.inpg.fr; Nicolas.Le-Bihan@gipsa-lab.inpg.fr).

S. J. Sangwine is with the Department of Computing and Electronic Systems, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: s.sangwine@ieee.org).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TSP.2007.910477

¹Biquaternions are quaternions with complex components, sometimes known as *complexified quaternions*. The biquaternions form a *geometric algebra*—More precisely, biquaternions are isomorphic to $\mathcal{C}\ell(3)$ the geometric algebra of the 3-D Euclidian space.

of the transform proposed in this paper. In our approach, the axis of the exponential is biquaternion-valued, introducing a degree of freedom in the transform orientation. This point will be detailed in Section III. Thus, the proposed transform is a generalization of the quaternion transform proposed by Ell and Sangwine for color images [1], [12].

The paper is organized as follows. First, in Section II we introduce the biquaternion algebra and some of the properties of biquaternions including the exponential of a biquaternion and the biquaternion roots of -1 . Also, biquaternion-valued signals are presented and it is seen how they can be used to encode coexisting signals of different geometric natures. Then, in Section III we introduce the definition and properties of the biquaternion Fourier transform. In Section III-C, we give an application of the BiQFT in the form of the definition—and some examples of—the notion of hyperanalytic signal. In Section IV we study the Hermitian symmetries of the BiQFT and their relation to the geometric properties of biquaternion-valued signals. Finally in Section V we propose a fast algorithm for the calculation of the transform using four complex fast Fourier transforms (FFTs) and allowing calculation of the transform with complexity of order $\mathcal{O}(N \log(N))$. Conclusions are given in Section VI.

II. BIQUATERNIONS

Biquaternions form an 8-dimensional algebra first discovered by Hamilton in 1853 [13]. In the following, we present their definitions and useful properties. For a more complete discussion, a good reference is Ward's book [14].

A. Basics

Biquaternions, also known as *complexified quaternions*, are quaternions with complex coefficients. The set of biquaternions is generally denoted $\mathbb{H}_{\mathbb{C}}$.

Definition 1 (Biquaternions): A biquaternion $q \in \mathbb{H}_{\mathbb{C}}$, is written in its Cartesian form as

$$q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} \quad (1)$$

where $q_0, q_1, q_2, q_3 \in \mathbb{C}$ are complex numbers. We write complex numbers using the notation

$$\mathbf{I}^2 = -1 \quad (2)$$

and so any complex number $z \in \mathbb{C}$ is noted: $z = \Re(z) + \mathbf{I}\Im(z)$.

It is also possible to write a biquaternion in the following form:

$$\begin{aligned} S(q) &= q_0 \\ V(q) &= q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} \end{aligned} \quad (3)$$

where $S(q)$ is the (complex-valued) scalar part of q while $V(q)$ is its (complex quaternion-valued) vector part. Also, the real and imaginary parts (with respect to $\mathbf{I}$) of a biquaternion are respectively defined as

$$\begin{aligned} \Re(q) &= \Re(q_0) + \Re(q_1)\mathbf{i} + \Re(q_2)\mathbf{j} + \Re(q_3)\mathbf{k} \\ \Im(q) &= \Im(q_0) + \Im(q_1)\mathbf{i} + \Im(q_2)\mathbf{j} + \Im(q_3)\mathbf{k} \end{aligned} \quad (4)$$

where $\Re(q)$ and $\Im(q)$ are (real) quaternions, allowing us to write any biquaternion $q \in \mathbb{H}_{\mathbb{C}}$ as $q = \Re(q) + \mathbf{I}\Im(q)$. The multiplication of biquaternions is carried out using the usual rules of quaternion multiplication [14], along with the additional rules:

$$\mathbf{i}\mathbf{I} = \mathbf{I}\mathbf{i}; \quad \mathbf{j}\mathbf{I} = \mathbf{I}\mathbf{j}; \quad \mathbf{k}\mathbf{I} = \mathbf{I}\mathbf{k} \quad (5)$$

which are equivalent to stating that the complex imaginary unit $\mathbf{I}$ commutes with the quaternion imaginary units $\mathbf{i}, \mathbf{j}, \mathbf{k}$.

From these two definitions we can see that biquaternions form an 8-dimensional associative algebra. However, they do not form, as we will see, a *normed* division algebra because there exist nonzero biquaternions with vanishing norm.

B. Conjugation and Norms

1) *Conjugation:* There are two basic ways of conjugating a biquaternion. *Quaternion conjugation* is related to the imaginary units $\mathbf{i}, \mathbf{j}, \mathbf{k}$ and *complex conjugation* to $\mathbf{I}$.

Definition 2: The *quaternion conjugate* of a biquaternion $q \in \mathbb{H}_{\mathbb{C}}$ is

$$q^* = q_0 - q_1\mathbf{i} - q_2\mathbf{j} - q_3\mathbf{k} = S(q) - V(q). \quad (6)$$

Definition 3: The *complex conjugate* of a biquaternion $q \in \mathbb{H}_{\mathbb{C}}$ is defined as

$$q^* = q_0^* + q_1^*\mathbf{i} + q_2^*\mathbf{j} + q_3^*\mathbf{k} \quad (7)$$

where $q_0^*, q_1^*, q_2^*, q_3^*$ are the complex conjugates of the complex coefficients of q .

Definition 4: Biquaternion conjugation is the combination of the two conjugations that have just been defined. The *biquaternion conjugate* of q is thus

$$\bar{q} = q_0^* - q_1^*\mathbf{i} - q_2^*\mathbf{j} - q_3^*\mathbf{k} = (q^*)^* = (q^*)^*. \quad (8)$$

Note that biquaternion conjugation is known as the *reversion* operation in the Clifford algebra formalism. Similarly, quaternion conjugation corresponds to Clifford conjugation and complex conjugation to grade involution. It is interesting to note that, while *complex conjugation* is multiplicative, i.e., $(pq)^* = p^*q^*$, *quaternion conjugation* and *biquaternion conjugation* are involutive. That is, for instance, $\overline{\overline{q}} = q$.

2) *Norm and Semi-norm of a Biquaternion:* The definitions of complex conjugation and quaternion conjugation can be used to define a norm and a semi-norm on biquaternions.

Definition 5: The following real positive-definite expression:

$$|q|^2 = S(q\bar{q}) = |q_0|^2 + |q_1|^2 + |q_2|^2 + |q_3|^2 \quad (9)$$

defines a norm, $|\bullet|$, on the vector space of biquaternions.

As mentioned before, biquaternions are not a normed algebra. So the norm is not multiplicative, $|pq| \neq |p||q|$. It is however possible to construct a complex-valued semi-norm which is multiplicative.

Definition 6: The semi-norm of a biquaternion $q \in \mathbb{H}_{\mathbb{C}}$ is defined as the complex number

$$\langle q \rangle^2 = S(qq^*) = q_0^2 + q_1^2 + q_2^2 + q_3^2. \quad (10)$$

Note that $\langle pq \rangle = \langle p \rangle \langle q \rangle$. The semi-norm is multiplicative.

However, the semi-norm is complex-valued and it can vanish for a nonzero biquaternion.² For example, the biquaternion $q = 1 + \mathbf{I}i$ satisfies $\langle q \rangle = 0$ and $|q| = \sqrt{2}$. If $\Im(q) = 0$ then $\langle q \rangle = |q|$, and if $\Re(q) = 0$ then $\langle q \rangle = \mathbf{I}|q|$.

C. Biquaternion Roots of -1

An important issue in defining a biquaternion Fourier transform is to define a biquaternion-valued exponential for the kernel of the transform. In the following, we will mainly consider exponential kernels containing a biquaternion root of -1 . The definition of such roots was recently introduced by Sangwine [15]. Their presence in the exponential kernel is essential to many properties of the BiQFT as well as for the fast Fourier transform (FFT) decomposition presented in this paper, Section V. Note, however, that in Section IV, we will consider an exponential kernel which does not contain a biquaternion root of -1 .

A biquaternion $\mu \in \mathbb{H}_{\mathbb{C}}$ is a biquaternion root of -1 iff

$$\mu^2 = \mu\mu = -1. \quad (11)$$

Assuming that μ is given by $\mu = \Re(\mu) + \mathbf{I}\Im(\mu)$, it has been shown in [15] that, except for the trivial cases $\mu = \pm \mathbf{I}, \pm \mathbf{i}, \pm \mathbf{j}, \pm \mathbf{k}$, μ is a root of -1 if it is a *pure* biquaternion (i.e., $\Re(\mu)$ and $\Im(\mu)$ are pure (real) quaternions³) and the following two conditions are satisfied

$$\begin{aligned} |\Im(\mu)|^2 - |\Re(\mu)|^2 &= -1 \\ (\Re(\mu)|\Im(\mu)) &= 0 \end{aligned} \quad (12)$$

where $(\cdot|\cdot)$ is a standard (real) quaternion-valued scalar product. Given two (real) *pure* quaternions $q = b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ and $p = w\mathbf{i} + x\mathbf{j} + y\mathbf{k}$, their scalar product is given by $(q|p) = bw + cx + dy$. Any three mutually orthogonal roots of -1 can be used as a basis to decompose a biquaternion. Indeed, given any biquaternion root of -1 , μ , and any biquaternion $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, it is possible to rewrite q as

$$q = (q'_0 + q'_1\mu) + (q'_2 + q'_3\mu)\nu \quad (13)$$

or equivalently

$$q = q'_0 + q'_1\mu + q'_2\nu + q'_3\mu\nu \quad (14)$$

where ν is a biquaternion root of -1 orthogonal to μ and q'_0, q'_1, q'_2, q'_3 are complex numbers. Equation (13) allows the definition of a decomposition for any biquaternion q , with respect to μ and ν , into its *simplex* $\text{Simp}(q)$ and *perplex* $\text{Perp}(q)$ parts in the following way:

- $\text{Simp}(q) = (q'_0 + q'_1\mu)$;
- $\text{Perp}(q) = (q'_2 + q'_3\mu)\nu$;

²The conditions for the semi-norm to vanish are quite specific, and were discovered by Hamilton. In short the real and imaginary parts of the biquaternion must have equal norms and a vanishing inner product.

³A real quaternion $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ where $a, b, c, d \in \mathbb{R}$ is called *pure* if $a = 0$.

and with $q = \text{Simp}(q) + \text{Perp}(q)$. This decomposition can be carried out in the same way as described in [16]. It will be of use for the definition of a fast algorithm in Section V.

D. Exponential of a Biquaternion

The kernel of the biquaternion Fourier Transform is, like that of the usual Fourier transform, an exponential. As stated before—see last paragraph Section II-C—we do not restrict our attention to exponential kernels containing biquaternion roots of -1 .

Here, the precise meaning of the exponential of a biquaternion is explained.

Definition 7: The exponential of a biquaternion $q \in \mathbb{H}_{\mathbb{C}}$ is a biquaternion denoted e^q , and given by the sum of the series

$$\sum_{n \in \mathbb{N}} \frac{q^n}{n!}. \quad (15)$$

This series is convergent in the norm $|\bullet|$, for all $q \in \mathbb{H}_{\mathbb{C}}$. Indeed, the norm of the n th term of the series is less than $|q|^n/n!$ so that the norms clearly converge.

Thus, any biquaternion can be written in its exponential form as

$$q = \langle q \rangle e^{\arctan\left(\frac{\langle V(q) \rangle}{S(q)}\right)} \widehat{V(q)} \quad (16)$$

where $\langle q \rangle$ is its semi-norm and $\widehat{V(q)}$ its normalized vector part, i.e., $\widehat{V(q)} = V(q)/\langle V(q) \rangle$. If q is a *pure biquaternion*⁴ and provided that $\langle q \rangle \neq 0$, then $\hat{q} = q/\langle q \rangle$ is a biquaternion root of -1 . Using (15), it is possible to express the exponential of a pure biquaternion as

$$e^q = \cos \langle q \rangle + \hat{q} \sin \langle q \rangle. \quad (17)$$

Note that, in the general case, $\cos \langle q \rangle$ and $\sin \langle q \rangle$ are complex. Furthermore, for any biquaternion q one can write

$$e^q = e^{S(q)} e^{V(q)} \quad (18)$$

where $e^{S(q)}$ is a complex number.⁵

Finally, we mention the following properties which will be used in the study of the Hermitian symmetries of the BiQFT, Section IV. Given a biquaternion $q \in \mathbb{H}_{\mathbb{C}}$, then

$$(e^q)^* = e^{q^*}. \quad (19)$$

In particular, in the case where q is a *pure biquaternion*, then

$$(e^q)^* = e^{-q} \quad (20)$$

and if $\Re(q) = 0$, then

$$(e^q)^* = e^{-q}. \quad (21)$$

⁴A biquaternion $q \in \mathbb{H}_{\mathbb{C}}$ is called *pure* if its scalar part vanishes, i.e., $S(q) = 0$.

⁵This identity follows from the fact that if $p, q \in \mathbb{H}_{\mathbb{C}}$ commute then $e^{p+q} = e^p e^q$. In general, $e^{p+q} \neq e^p e^q$, since biquaternion algebra is not commutative.

E. Biquaternions and Vector Signals

We now show how signals with different physical natures can be embedded in a biquaternion-valued signal. In Section IV, a generalization of the BiQFT is presented that allows to exhibit the physical nature of a signal as a symmetry of its BiQFT spectrum.

A biquaternion-valued signal $f(t)$ is given by

$$\begin{aligned} f(t) &= f_0(t) + f_1(t)\mathbf{i} + f_2(t)\mathbf{j} + f_3(t)\mathbf{k} \\ &= S(f(t)) + V(f(t)) \end{aligned} \quad (22)$$

where f_0, f_1, f_2 , and f_3 are complex-valued signals. Such signals could represent various physical quantities captured on co-located sensors (for example dipoles and magnetic loops for electromagnetic signal recording).

It is a well known fact [14] that the components of a biquaternion $q \in \mathbb{H}_{\mathbb{C}}$ are scalar, pseudo-scalar, vector and bivector. The identification is made in the following way, with possible signal association.

- $\Re(q)$ is a scalar, invariant under all geometric transformations. An example of a scalar signal is an electric potential (voltage).
- $\Im(S(q))$ is a pseudo-scalar, invariant under rotations but changing its sign in parity transformations.
- $\Re(V(q))$ is a bivector. An example of a bivector valued signal is a magnetic field.
- $\Im(V(q))$ is a vector, invariant in parity transformations but otherwise transforming like a vector. An example of a vector signal is an electric field.

It is then obvious that biquaternion-valued signals can describe coexisting signals with different geometrical natures, for example, an electric signal and a magnetic signal propagating together. They can be encoded in $\Re(V(f(t)))$ for the magnetic field and in $\Im(V(f(t)))$ for the electric field. It is then possible to process these signals together but still treat the electric signal as a vector quantity and the magnetic signal as a bivector quantity (preservation of their physical natures). Also, the biquaternion model allows us to take advantage of all the information available in the electromagnetic field.

Finally, we note that the polarization-dependent effects of propagation may be directly modeled on electromagnetic signals via Clifford (biquaternion) transformations [17]. This motivates our development of a Fourier transform for such signals. We stress again the fact that this is not possible with a long vector containing the components of both fields, nor with a quaternion-valued signal.

III. THE BIQUATERNION FOURIER TRANSFORM: BiQFT

In the sequel, we will consider discretized biquaternion-valued signals. This approach is sufficient for the properties we will discuss. It is also closer to numerical implementation. Thus, the discretized signal f will be indexed as $f(n)$ with $n = 0, \dots, N-1$ the sample number. A biquaternion-valued signal will be considered as a vector of $\mathbb{H}_{\mathbb{C}}^N$, i.e., $f = [f(0)f(1)\dots f(N-1)]^T$.

After introducing definitions and properties of the biquaternion Fourier transform, we give, in Section III-C, an example

of an application of the BiQFT that takes advantage of its biquaternion nature. This application involves defining the notion of *hyperanalytic signal* which generalizes the concept of the (complex) analytic signal of a real signal to the (hypercomplex) hyperanalytic signal of a complex signal. In Section IV, we consider a generalized version of the BiQFT which has interesting Hermitian symmetry properties.

In this paper, we limit the transforms considered to the case of a single exponential kernel on the left or right of the signal function (this is usually called a one-sided hypercomplex transform). It is possible (indeed likely) that various two-sided transforms can be formulated with exponentials on both sides of the signal function, but we have not considered this to date, because the one-sided transforms defined by Ell and Sangwine [12] for quaternions have proved to be versatile, and we have found that the one-sided biquaternion transforms defined in this paper have similar properties. It may be worthwhile to study the two-sided case, but there are many possibilities and establishing which ones are invertible is not a trivial task.

A. The Left and Right BiQFT

We now introduce the definitions of the left and right biquaternion Fourier transform.

Definition 8: Given a biquaternion root of $-1, \mu$, then for every biquaternion-valued signal f , its Left BiQFT is defined as

$$\hat{f}_L(u) = \text{LQFT}(f)(u) = \sum_{n=0}^{N-1} e^{-2\pi\mu \frac{un}{N}} f(n) \quad (23)$$

where u is a frequency variable.

In a similar way, we define the Right BiQFT by changing the order of the signal and the exponential.

Definition 9: Given a biquaternion root of $-1, \mu$, then for every biquaternion-valued function f , its Right BiQFT is defined as

$$\hat{f}_R(u) = \text{RQFT}(f)(u) = \sum_{n=0}^{N-1} f(n) e^{-2\pi\mu \frac{un}{N}} \quad (24)$$

where u is a frequency variable.

It is very important to note that to every different chosen μ there corresponds a different transform. We call μ the *axis of the transform*. If, as a special case, $\mu = \mathbf{I}$ is chosen, then one obtains a transform equivalent to that defined by Ebling and Scheuermann [11]. Since the exponential in this case is complex, and not quaternion- or biquaternion-valued, the transform is equivalent to independent complex Fourier transforms applied to the components of the biquaternion signal. Another special case is when μ is chosen to be a *quaternion* (not biquaternion) root of -1 , in which case the transform is equivalent to two independent quaternion transforms as defined in [12] applied to the real and imaginary parts of the biquaternion signal. It is only when μ is chosen to be a biquaternion that we obtain a truly biquaternion Fourier transform in which the samples of the biquaternion signal are treated as a whole, and not as independent components.

The left and right transforms are both invertible, as follows⁶:

$$f(n) = \frac{1}{N} \sum_{u=0}^{N-1} e^{2\pi\mu \frac{nu}{N}} \hat{f}_L(u) \quad (25)$$

and

$$f(n) = \frac{1}{N} \sum_{u=0}^{N-1} \hat{f}_R(u) e^{2\pi\mu \frac{nu}{N}}. \quad (26)$$

The proof of these inversion formulae is the same as for a usual discrete Fourier transform. It is based on the fact that $e^{2\pi\mu(nu/N)}$ is an N^{th} root of unity.

The left and right BiQFTs are related by the two following identities:

$$\begin{aligned} \hat{f}_L(u) &= \hat{f}_R(u) - 2V \left(\mu \sum_{n=0}^{N-1} V(f(n)) \sin(\pi \frac{nu}{N}) \right) \end{aligned} \quad (27)$$

$$\begin{aligned} \text{LQFT}(f)(u) &= \text{RQFT}(\text{Simp}(f))(u) + \text{RQFT}(\text{Perp}(f))(-u) \end{aligned} \quad (28)$$

where $\text{Simp}(f)$ and $\text{Perp}(f)$ denote the simplex and perplex parts of f with respect to μ , i.e., $f = \text{Simp}(f) + \text{Perp}(f)$ where $\text{Simp}(f)$ and $\text{Perp}(f)$ are biquaternion-valued in μ . The proof of these identities is given in the Appendix A.

In the following, when the type of the transform, left or right, is not specified, then the property considered is true for both.

B. The Convolution Theorem

The convolution of two functions f_1 and f_2 is defined as

$$\begin{aligned} (f_1 * f_2)(n) &= \sum_{m=0}^{N-1} f_1(m) f_2(n-m) \\ &= \sum_{m=0}^{N-1} f_1(n-m) f_2(m). \end{aligned} \quad (29)$$

Since biquaternions are not commutative $f_1 * f_2 \neq f_2 * f_1$.

Using a BiQFT, whose axis is any root of -1 , a convolution can be transformed into a product. This simplifies and accelerates its calculation

$$\begin{aligned} f_1 * f_2 &= \text{RQFT}^{-1}(\hat{f}_1(u) \widehat{\text{Simp}(f_2)}(u)) \\ &\quad + \frac{1}{N} \text{RQFT}(\hat{f}_1(u) \widehat{\text{Perp}(f_2)}(-u)) \end{aligned} \quad (30)$$

where $\hat{f}_1 = \text{RQFT}(f_1)$, $\widehat{\text{Simp}(f_2)} = \text{RQFT}(\text{Simp}(f_2))$, and $\widehat{\text{Perp}(f_2)} = \text{RQFT}(\text{Perp}(f_2))$. This identity is a generalization of the convolution theorem for quaternion-valued functions, proved in [18]. It can be proved in the same way.

C. Hyperanalytic Signal

We present here a brief exposition of an application of the BiQFT which makes use of the biquaternion nature of the

⁶We are here adopting the convention of applying the scale factor $(1/N)$ to the inverse transform, which is not, of course, the only possibility.

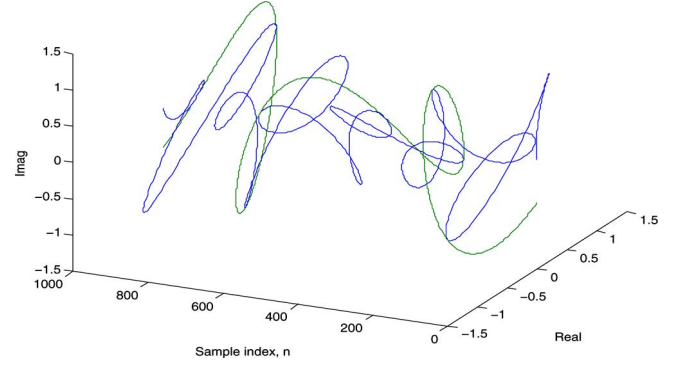


Fig. 1. Complex signal $f[n]$ (blue) and its complex envelope $|a[n]|$ (green).

transform by generalizing the concept of the (complex) *analytic signal* of a real signal [19], [20] to the case of a complex signal with a hypercomplex (in fact biquaternion) analytic signal. Given a real-valued signal $f(t)$, its analytic signal $a(t)$ is a complex-valued signal with the following important properties: $\Re(a(t)) = f(t)$, $\Im(a(t)) \perp f(t)$ (where $\perp$ means orthogonal). The modulus of $a(t)$ is the *envelope* of $f(t)$. The analytic signal is defined in terms of the Hilbert transform of $f(t)$. In the discrete case it can be computed by suppressing the negative frequency components of the Fourier transform of $f(t)$: the inverse Fourier transform of the modified spectrum yields the analytic signal. This is explained thoroughly by Bracewell [21, Chapter 13].

We now generalize this idea to the case where f is complex-valued and demonstrate a result that shows the validity of the idea. There is much work to be done to develop the theoretical explanation of this result. Nevertheless it shows that the BiQFT makes possible a generalization of the analytic signal in a very straightforward way. Using the BiQFT implementation in [2], we have taken a complex-valued discrete signal defined as

$$f[n] = \sin(16\pi n/N)(\sin(4\pi n/N) + \mathbf{I}\sin(6\pi n/N))$$

where $N = 1000$, and computed its BiQFFT by placing $f[n]$ in the scalar part of a biquaternion signal with zero vector part. Suppressing the negative frequency coefficients, and doubling the amplitude of the positive frequency coefficients in exactly the same way as in the standard case with a complex FFT we obtain the Fourier spectrum of a *hyperanalytic signal*. The inverse BiQFFT yields the hyperanalytic signal itself $a[n]$, which is, of course, biquaternion-valued. The semi-norm (defined in (10)) of the hyperanalytic signal is complex, and we find that it is indeed a *complex envelope* as shown in Fig. 1: notice that the complex envelope is tangent to the original signal at some of the points of inflexion. The computation of this complex envelope requires phase unwrapping, otherwise it exhibits discontinuities, but otherwise it is simply the complex modulus of $a[n]$. We have found that the complex envelope is independent of the transform axis μ , even though each choice of transform axis yields a different hyperanalytic signal. We find also that the modulus of the complex envelope is the envelope of the modulus of $f[n]$, as shown

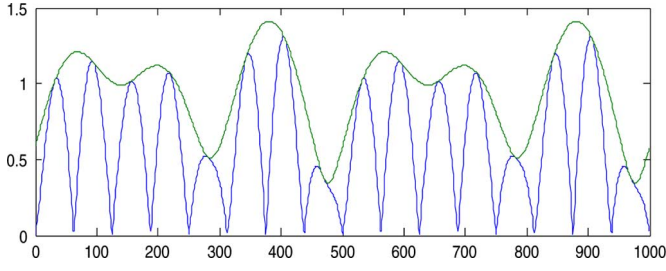


Fig. 2. Modulus of the complex envelope (green) and modulus of $f[n]$ (blue).

in Fig. 2, suggesting that the complex envelope is indeed fundamental.⁷ Furthermore, it is possible to derive a complex signal orthogonal to $f[n]$ by projection of the hyperanalytic signal $a[n]$ onto any biquaternion direction perpendicular to the transform axis μ . (Projection of $a[n]$ onto μ yields $f[n]$.)

A full study of the properties of the hyperanalytic signal must be deferred to a later paper but the results demonstrated here show that the BiQFT provides a simple and natural way to extend the idea of the analytic signal to the case of a complex signal with a hypercomplex hyperanalytic signal. Of course, the BiQFT may not be the only way in which a hyperanalytic signal can be computed, but that also remains a subject for further study.

IV. HERMITIAN SYMMETRIES

In this section, the Hermitian symmetries of the BiQFT are studied. As in the case of the classical Fourier transform, the BiQFT representation translates the symmetries of the signal to Hermitian symmetries of the spectrum. It has been mentioned, Section II-E, that a biquaternion-valued signal can be used to encode several coexisting/copropagating physical signals of different geometric natures. In this section it will be seen that by allowing a more general choice of the axis of the BiQFT—in fact, by allowing an axis of the transform which is not a biquaternion root of -1 —the transform gains additional Hermitian symmetry properties, that allow it to transform signals with different geometric natures into spectra with complementary Hermitian symmetries. In other words, the Hermitian symmetry properties obtained in this way preserve the separation, explained in Section II-E, between signals of various geometries which are the components of a biquaternion-valued signal. This means that the BiQFT of a biquaternion-valued signal will map the different components of the signal having different geometries to parts of the spectrum with different symmetries.

Let us consider the left and right BiQFT, as defined in formulae (23) and (24). We now consider that in these formulae μ is a pure biquaternion, not necessarily a root of -1 . The transforms obtained in this way generalize the BiQFT. These generalizations retain many of the properties given in Section III. We discuss which of these properties continue to hold and which need to be modified in Appendix B. In the following we continue to refer to these generalized transforms, whose axis is not a root of -1 , as the left and right BiQFT.

⁷The value of μ used for the computation of Fig. 2 was $\mathbf{i} + (1 + \mathbf{I})\mathbf{j} + (1 - \mathbf{I})\mathbf{k}$. The same result is obtained for other values of μ .

The specificity of biquaternions is that there are two essential kinds of conjugation, quaternion conjugation and complex conjugation. With an adequate choice of the transform axis, the BiQFT will have Hermitian symmetry properties with respect to each of these two kinds of conjugation. We examine these properties in the two following paragraphs.

A. Symmetries With Respect to Quaternion Conjugation

Let f be a biquaternion-valued function. Then its BiQFT satisfies the following property:

$$(\hat{f}_R)^*(u) = (\hat{f}_L^*)(-u). \quad (31)$$

This property can be proved by direct calculation of $(\hat{f}_R)^*(u)$ using identity (20). Note that in order to apply (20), the axis μ must satisfy $S(\mu) = 0$. Using identity (31), it is possible to prove the two following properties—see Appendix C:

Property 1: Let $a(n) \in \mathbb{C}$ be a complex-valued signal. Its BiQFT, which we denote $A(u)$, has the following property:

- $S(A(u))$ is even;
- $V(A(u))$ is odd.

Property 2: Let $l(n) = l_1(n)\mathbf{i} + l_2(n)\mathbf{j} + l_3(n)\mathbf{k}$ be a pure biquaternion-valued signal, $l_1(n), l_2(n), l_3(n) \in \mathbb{C}$. Its BiQFT, $L(u)$ satisfies the following:

- $S(L(u))$ is odd.

It is, in general, impossible to conclude anything about the symmetry of the vector part of the spectrum.

It is obvious that the quaternion conjugation in the biquaternion Fourier domain allows us to exhibit some symmetries of the considered signal. In order to extract more symmetries, complex conjugation has to be considered.

B. Symmetries With Respect To Complex Conjugation

The Hermitian symmetry of the BiQFT with respect to quaternion conjugation, described by property (31), allows the separation of the spectra of $S(f)$ and $V(f)$ —for a biquaternion-valued function f —based on their symmetries. To be able to separate $\Re(f)$ and $\Im(f)$ in the same way, we need to use the Hermitian symmetry of the BiQFT with respect to complex conjugation. The Hermitian symmetry with respect to complex conjugation is described by the following property:

Let μ be a pure biquaternion such that $\Re(\mu) = 0$. Then the BiQFT of f with respect to axis μ satisfies the following property:

$$\hat{f}^*(u) = \widehat{f^*}(-u). \quad (32)$$

Again, this can be proved by direct calculation, using identity (21). In order for (21) to apply, the axis μ needs to satisfy $\Re(\mu) = 0$. This means that it cannot be a root of -1 . If we limit the choice of μ to roots of -1 we can only obtain the symmetry with respect to quaternion conjugation (31).

The two following properties result from (32), in the same way as Property 1 and Property 2 result from (31):

Property 3: Let $R(n) = R_0(n) + R_1(n)\mathbf{i} + R_2(n)\mathbf{j} + R_3(n)\mathbf{k}$ with $R_0, R_1, R_2, R_3 \in \mathbb{R}$ be a (real) quaternion-valued signal, then:

- $\Re(\hat{R})$ is even;
- $\Im(\hat{R})$ is odd.

TABLE I
SYMMETRIES OF THE BiQFT

Signal: q	$\Re(S(q))$	$\Im(S(q))$	$\Re(V(q))$	$\Im(V(q))$
scalar	even	zero	zero	odd
pseudoscalar	zero	even	odd	zero
bivector	zero	odd	even	odd
vector	odd	zero	odd	even

Property 4: Let $C(n) = C_0(n)\mathbf{1} + C_1(n)\mathbf{i} + C_2(n)\mathbf{j} + C_3(n)\mathbf{k}$ with $C_0, C_1, C_2, C_3 \in \mathbb{R}$ be a pure imaginary biquaternion-valued signal, then:

- $\Re(\hat{C})$ is odd;
- $\Im(\hat{C})$ is even.

These separations between the components of the biquaternion-valued signal are only accessible through complex conjugation.

C. Synthesis: Separating the Spectra of Signals With Different Geometries

Now we will see how the BiQFT maps signals that are scalar, pseudoscalar, vector or bivector, into spectra that have different Hermitian symmetries. This is based on the simultaneous use of the Hermitian symmetries of the BiQFT with respect to quaternion and to complex conjugation (i.e., properties (31) and (32)). Property (31) allows the parts $S(q)$ and $V(q)$ of a signal q to be separated in the Fourier domain. Property (32) separates $\Re(q)$ and $\Im(q)$. If we use the two properties simultaneously, we can separate the spectra of the four parts $\Re(S(q))$, $\Im(S(q))$, $\Re(V(q))$ and $\Im(V(q))$, that is of the—respectively, as it was seen in Section II-E—scalar, pseudoscalar, vector and bivector parts of the signal q .

Remember that property (31) is only verified if the axis of the transform μ verifies $S(\mu) = 0$. Property (32) is only verified if $\Re(\mu) = 0$. Indeed these conditions are necessary for (20) and (21). So in order to have (31) and (32) together, it is necessary that the axis μ verifies $S(\mu) = 0$ and $\Re(\mu) = 0$. This choice of axis allows maximum Hermitian symmetry of the BiQFT and makes possible the separation of all four parts/components of the signal (scalar, pseudoscalar, vector and bivector) in the Fourier domain.

Let μ verify $S(\mu) = 0$ and $\Re(\mu) = 0$. We consider the BiQFT, with axis μ , of a scalar, a pseudoscalar, a vector and a bivector signal. Each of these four signals verifies one of the two properties 1 and 2 of paragraph IV-B, and one of the two properties 3 and 4 of paragraph IV-B. Thus, the BiQFT of these signals are subject to four conditions, two from each property which is verified. For instance, a scalar signal s verifies properties 1 and 3 and its BiQFT is subject to the symmetry conditions stated in each of these two properties.

For each of the four kinds of signals, we can solve the symmetry conditions imposed by the verified properties and obtain Table I, where we denote the signal by q and its BiQFT by Q .⁸

No two lines of this table are identical. This means that the BiQFT of each of the four kinds of signals has completely different symmetry. In other words, scalar, pseudoscalar, vector

⁸For unique separation, we assume general scalar, pseudoscalar, bivector signals q with nonzero even and odd parts.

and bivector signals are mapped by the BiQFT to spectra that have complementary symmetries and that are, in this way, distinguishable from one another.

Just as in the original biquaternion-valued signal, the four kinds of signals are separated on different components. They are, in the BiQFT of the signal, separated on orthogonal subspaces (due to the multivector structure of biquaternion vector space) of functions with a particular symmetry.

V. THE BiQFFT ALGORITHM

In this section, a fast algorithm for calculating the BiQFT is presented: the fast BiQFT, or BiQFFT. The complexity of this algorithm is the same as the usual FFT, i.e., $\mathcal{O}(N \log(N))$, apart from a scale factor. The BiQFFT algorithm is based on the one introduced in [16] for the quaternion Fourier transform (see also [12]).

The BiQFFT algorithm first decomposes the BiQFT into four complex Fourier transforms. This is done using a linear transformation. The four complex Fourier transforms can then be calculated using an existing FFT algorithm or code. The linear transformation is then inverted and the BiQFT is retrieved. This means that the BiQFFT algorithm costs four complex FFTs plus a constant time for the linear transformation, which will be small compared to the time taken to compute the four complex transforms. We will first define the BiQFFT for the calculation of the left BiQFT and then explain how it can be modified in order to calculate the right BiQFT.

A. Factorization of the BiQFT Into Four Complex Transforms

Consider the biquaternion-valued signal $f(n)$ given in (22) and express it in terms of a complex orthonormal basis defined by μ , the transform axis. The complex orthonormal basis is defined by μ and two other unit pure biquaternions ν and ξ such that $\mu \perp \nu \perp \xi$ and $\mu\nu = \xi$ (and $\mu\nu\xi = -1$).⁹ This basis induces a simplex/perplex decomposition which is the basis of the algorithm presented here. As in the real quaternion case presented in [12], it is much more efficient numerically to implement the decomposition by a change of basis. The only difference here is that the basis is complex rather than real. The basis may be represented by a 3×3 complex orthogonal (not Hermitian) matrix

$$\begin{pmatrix} \mu_x & \mu_y & \mu_z \\ \nu_x & \nu_y & \nu_z \\ \xi_x & \xi_y & \xi_z \end{pmatrix}$$

Given a value for μ , the choice of ν and ξ is arbitrary within the constraints given. As an example, if we take the real part of μ to be in the direction $x = y = z$ thus: $\Re(\mu) = \mathbf{i} + \mathbf{j} + \mathbf{k}$, then a suitable value for the imaginary part is $\Im(\mu) = \mathbf{j} - \mathbf{k}$, so that $\mu = \mathbf{i} + (1 + \mathbf{i})\mathbf{j} + (1 - \mathbf{i})\mathbf{k}$, which has a (real) modulus of 1 and a square of -1 as can easily be verified algebraically. This value for μ could be said to be a natural choice since it does not favor any of the x , y or z directions (at least in the real part). In general of course, the choice of (complex) direction for μ must be dependent on the application and the nature of the signals, particularly where some preferred direction exists. This

⁹ $\mu \perp \nu$ means $\sum_{i=1}^3 \mu_i \nu_i = 0$.

remains a topic to be studied as the transforms defined in this paper are applied to real-world applications.

In the new basis, we have

$$\begin{aligned} f(n) &= f_0(n) + f'_1(n)\mu + f'_2(n)\nu + f'_3(n)\xi \\ &= [f_0(n) + f'_1(n)\mu] + [f'_2(n) + f'_3(n)\mu]\nu \end{aligned}$$

Using this change of basis, we are able to separate the transform into the sum of two transforms:

$$F(u) = \sum_{n=0}^{N-1} \begin{pmatrix} e^{-2\pi\mu\frac{nu}{N}} [f_0(n) + f'_1(n)\mu] \\ + e^{-2\pi\mu\frac{nu}{N}} [f'_2(n) + f'_3(n)\mu]\nu \end{pmatrix}$$

We now separate the terms on the right into real and imaginary parts and group the real components together and the imaginary components together to make four complex terms.

$$F(u) = \sum_{n=0}^{N-1} \begin{pmatrix} e^{-2\pi\mu\frac{nu}{N}} [\Re(f_0(n)) + \Re(f'_1(n))\mu] \\ + e^{-2\pi\mu\frac{nu}{N}} \mathbf{I} [\Im(f_0(n)) + \Im(f'_1(n))\mu] \\ + e^{-2\pi\mu\frac{nu}{N}} [\Re(f'_2(n)) + \Re(f'_3(n))\mu]\nu \\ + e^{-2\pi\mu\frac{nu}{N}} \mathbf{I} [\Im(f'_2(n)) + \Im(f'_3(n))\mu]\nu \end{pmatrix}.$$

All four of the transforms within this expression are now isomorphic to a complex Fourier transform. That is, we may replace μ with $\mathbf{I}$ (the complex root of -1) in order to compute the transform, and we will obtain the same numeric results. In computing these four ‘complex’ transforms, we can use any available FFT algorithm. After computing the four complex transforms, all that remains is to re-assemble the parts of $F(u)$ and invert the change of basis. The latter step is equivalent to multiplying out the factors of μ and ν appearing above, but it is more easily performed by a change of basis using the transpose of the original basis matrix used to change from the standard basis to the (μ, ν, ξ) basis.

B. Algorithm for the Right BiQFT

We use formula (28) plus a standard DFT trick. Formula (28) shows that the left BiQFT can be expressed in terms of the right BiQFFT provided the signal is split into simplex and perplex components, as shown in Appendix A. The first term in formula (28) is the right BiQFFT of the simplex part. In fact, because the simplex part commutes with the hypercomplex exponential, the right BiQFFT of the simplex part is identical to the left BiQFFT of the simplex part. The second term is the right inverse BiQFFT of the perplex part, or it may be regarded as the right BiQFFT of the perplex part with negated u , apart from the scale factor of $(1/N)$. Since the computation of the left BiQFFT requires separation into simplex and perplex parts, as shown in the previous section, the only change needed to compute the right BiQFFT is to the computation of the BiQFFT of the perplex part. This requires the hypercomplex exponential to be conjugated, which is a trivial implementation problem. In fact a trick is used: instead of conjugating the exponential (which would require a special modification to the FFT code, or a call to an inverse FFT), the signal is conjugated before computing the FFT, and the result is

conjugated afterwards.¹⁰ The interested reader is referred to our open-source code in [2] for details.¹¹

In this way, a right BiQFFT algorithm, similar to the one given in Section V-A, can be defined. It has the same complexity, $O(N \log(N))$.

Note that in order for the two fast algorithms that have been presented for the left and right BiQFT to work, the transform axis μ must be a biquaternion root of -1 . It was explained in Section IV-C that in order to take full advantage of the Hermitian symmetries of the BiQFT, we must choose an axis μ such that $S(\mu) = 0$ and $\Re(\mu) = 0$. Since such an axis can not be a root of -1 , the BiQFFT algorithm we have presented can not be applied in this case.

VI. CONCLUSION

We have introduced the Biquaternion Fourier transform, or BiQFT. This transform is a tool for the harmonic analysis of biquaternion-valued signals. Our interest for these signals is motivated by their ability to encapsulate several physical signals with different geometric properties into one biquaternion signal representation. We have considered an application of this transform in the form of the definition of the hyperanalytic signal. This notion generalizes that of the analytic signal to complex-valued signals. It is based on the biquaternion structure of the BiQFT. We have also discussed the Hermitian symmetries of the BiQFT. We have seen under what conditions these symmetries retain, in the spectral domain, the separation between signals with different geometric natures that exists in the time domain. Finally, the BiQFFT algorithm was introduced. This fast algorithm allows the calculation of the BiQFT in a time of the order of $N \log(N)$, using a complex change of basis and four usual FFTs. This allows efficient implementation of the BiQFT using commonly existing libraries. The BiQFFT algorithm is implemented in the quaternion toolbox for Matlab [2], developed by two of the authors.

APPENDIX

In this Appendix, we prove some of the properties given in the paper. This is an interesting exercise as it shows instances of how we can manipulate the BiQFT. In part A we prove the identities (27) and (28), relating the left and right transforms. In part B we discuss the properties of the BiQFT with an axis which is not a biquaternion root of -1 , which was proposed in Section IV. In particular, we see how it can be inverted. Finally, in part C we prove the Hermitian symmetry properties of Section IV.

A. Relations Between the Left and Right Transforms

Here we prove the identities (27) and (28) relating the left and right BiQFT.

The proof of (27) is based on the commutation relation, true for any two biquaternions p and q :

$$pq - qp = 2V(V(p)V(q)) \quad (33)$$

¹⁰This is the standard DFT trick referred to above: an inverse DFT can always be computed by conjugating the signal, computing a forwards DFT, and conjugating the result. This follows simply from the identity $AB^* = (A^*B)^*$.

¹¹See the file qfft.m

Now, notice that—from formula (17)— $e^{-2\pi\mu(nu/N)} = \cos[2\pi(nu/N)] - \mu \sin[2\pi(nu/N)]$. It follows that

$$e^{-2\pi\mu \frac{nu}{N}} f(n) = f(n)e^{-2\pi\mu \frac{nu}{N}} - 2V \times \left(\mu V(f(n)) \sin \left[2\pi \frac{nu}{N} \right] \right).$$

Now summing on n we obtain (27).

The proof of (28) is also based on commutation relations. Let $\text{Simp}(f)$ and $\text{Perp}(f)$ be the simplex and perplex parts of f with respect to μ , as defined in Section II-C. Then:

$$e^{-2\pi\mu \frac{nu}{N}} f(n) = e^{-2\pi\mu \frac{nu}{N}} \text{Simp}(f) + e^{-2\pi\mu \frac{nu}{N}} \text{Perp}(f). \quad (34)$$

Noting that $\text{Simp}(f)$ commutes with the exponential kernel, while $\text{Perp}(f)$ commutes with it after conjugation. The last equation becomes

$$e^{-2\pi\mu \frac{nu}{N}} f(n) = \text{Simp}(f)e^{-2\pi\mu \frac{nu}{N}} + \text{Perp}(f)e^{2\pi\mu \frac{nu}{N}}.$$

By summing both sides, we obtain (28).

B. The BiQFT With an Axis Which Is Not a Root of -1

In Section IV, we proposed to use a BiQFT with an axis which is not a root of -1 . By relaxing this restriction on the axis of the transform, some of the properties of the BiQFT are changed. For example, all the properties in Section III that use the decomposition of the signal into simplex and perplex parts do not hold for a BiQFT with μ not a root of -1 . This is for the simple reason that the simplex and perplex parts have *no meaning* in this case—see Section II-C. The most important properties that are lost when we consider $\mu^2 \neq -1$ are those related to the inversion of the BiQFT. It is not possible, for $\mu^2 \neq -1$, to invert the BiQFT using the inversion formulae (25) and (26). Indeed, if μ is not a root of -1 then $e^{2\pi\mu(nu/N)}$ is not an N^{th} root of unity. In this part of the Appendix, we will discuss the inversion of the BiQFT, in this more general case. However, let us first give a more general form of identity (27), which remains valid when $\mu^2 \neq -1$. This will be of interest in the following part C of the Appendix.

Let $f(n)$ be some biquaternion valued signal, then its left and right BiQFT, with respect to any axis μ , are related by the identity

$$\hat{f}_L(u) = \hat{f}_R(u) - 2V \left(\hat{\mu} \sum_{n=0}^{N-1} V(f(n)) \sin \left(2\pi \frac{nu}{N} \langle \mu \rangle \right) \right) \quad (35)$$

where $\langle \mu \rangle$ is the semi-norm of μ , given by (10), and $\hat{\mu} = (\mu / \langle \mu \rangle)$. This identity can be proved in the same way that (27) was proved in part A—with the same use of formula (17). It reduces to (27) when $\mu^2 = -1$.

In Section IV, we propose to use an axis of the transform μ which is pure—i.e., has zero scalar part- and pure imaginary—i.e., has zero real part. This is the choice of the axis of the transform that achieves the highest degree of symmetry. In order to invert a BiQFT with an axis of this kind, it is possible to perform the following trick: Start by complexifying the frequency

index u , write $u = u' + Iu''$. The “BiQFT” obtained in this way, which is a function of the complex index u , is given by

$$\hat{f}(u', u'') = \sum_{n=0}^{N-1} e^{-2\pi\mu \frac{n(u'+Iu'')}{N}} f(n). \quad (36)$$

Of course the BiQFT of f is just the above quantity taken for $u'' = 0$.

Note that

$$\hat{f}(u', u'') = \sum_{n=0}^{N-1} e^{-2\pi I\mu \frac{nu''}{N}} \left(e^{-2\pi\mu \frac{nu'}{N}} f(n) \right).$$

Now since $I\mu$ is a biquaternion root of -1 , we can rewrite the above equation

$$\hat{f}(u', u'') = \text{LQFT} \left(e^{-2\pi\mu \frac{nu'}{N}} f(n) \right)$$

where the axis of the transform is a root of -1 . Now, this transform can be inverted using the standard formula (25). Using this fact we obtain

$$f(n) = \frac{1}{N} \sum_{u''=0}^{N-1} e^{2\pi\mu \frac{n(u'+Iu'')}{N}} \hat{f}(u', u''). \quad (37)$$

We see that a BiQFT with an axis which is not a root of -1 can not be inverted by itself. However it is possible to invert it by complexifying its frequency variable.

C. Hermitian Symmetries of the BiQFT

In Section IV, we use relations (31) and (32) to establish Hermitian symmetries of signals with specific geometry. These symmetries are included in the four properties given in that section, which we have labelled Property1, Property2, Property3, and Property4. We show here how (31) leads to Property1 and Property2.

Property1 concerns the BiQFT of a complex-valued signal $a(n) \in \mathbb{C}$. Its BiQFT $A(u)$ verifies (31). Replacing in this property the two following identities verifies for any complex-valued signal:

$$\begin{aligned} a^*(n) &= a(n) \\ \text{LQFT}(a) &= \text{RQFT}(a). \end{aligned} \quad (38)$$

The second identity is verified since a complex number commutes with any biquaternion. The following equation follows:

$$A^*(u) = A(-u).$$

Property1 follows from decomposing this last equation into its vector and scalar parts.

We turn to Property2. The signal $l(n)$ is pure biquaternion, its BiQFT is noted $L(u)$. The proof can be done by replacing in (31) the identities

$$\begin{aligned} l^*(n) &= -l(n) \\ S(\text{LQFT}(l)) &= S(\text{RQFT}(l)). \end{aligned} \quad (39)$$

The second of these identities can be seen to follow from (35) or (27).

In the same way as these two properties—Property1 and Property2—have been shown to follow from (31), Property3 and Property4 can be shown to follow from (32).

REFERENCES

- [1] S. J. Sangwine, "Fourier transforms of colour images using quaternions, or hypercomplex, numbers," *Electron. Lett.*, vol. 32, no. 21, pp. 1979–1980, 1996.
- [2] S. Sangwine and N. Le Bihan, Quaternion Toolbox for Matlab [Online]. Available: <http://qtfm.sourceforge.net/> (online software library).
- [3] T. A. Ell, "Hypercomplex spectral transformations," Ph.D. dissertation, University of Minnesota, Minneapolis, 1992.
- [4] V. M. Chernov, "Arithmetic method in the theory of discrete orthogonal transforms," *Proc. SPIE*, vol. 2363, pp. 134–141, 1995.
- [5] V. M. Chernov, "Discrete orthogonal transforms with data representation in composition algebras," in *Proc. 9th Scandinavian Conf Image Analysis (SCIA)*, Uppsala, Sweden, 1995, vol. 1, pp. 357–364.
- [6] V. M. Chernov, "Algorithms of two-dimensional discrete orthogonal transforms realized in Hamilton-Eisenstein codes," *Transl.: Problemy Peredachi Informatsii Prob. Inf. Transmission*, vol. 31, no. 3, pp. 228–235, 1995.
- [7] G. Sommer, Ed., *Geometric Computing With Clifford Algebras*. New York: Springer, 2001.
- [8] S.-C. Pei and C.-M. Cheng, "A novel block truncation coding of color images by using quaternion-moment preserving principle," in *Proc. IEEE Int. Symp. Circuits systems*, Atlanta, GA, May 12–15, 1996, vol. 2, pp. 684–687.
- [9] T. Bülöw and G. Sommer, "Hypercomplex signals—A novel extension of the analytic signal to the multidimensional case," *IEEE Trans. Signal Process.*, vol. 49, no. 11, pp. 2844–2852, Nov. 2001.
- [10] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Process.*, vol. 49, no. 12, pp. 3136–3144, Dec. 2001.
- [11] J. Ebling and G. Scheuermann, "Clifford Fourier transform on vector fields," *IEEE Trans. Vis. Comput. Graphics*, vol. 11, no. 4, pp. 469–479, Jul.–Aug. 2005.
- [12] T. A. Ell and S. J. Sangwine, "Hypercomplex Fourier transforms of color images," *IEEE Trans. Image Process.*, vol. 16, no. 1, pp. 22–35, Jan. 2007.
- [13] W. R. Hamilton, "On the geometrical interpretation of some results obtained by calculation with biquaternions," *Proc. Roy. Irish Acad.*, vol. V, pp. 388–390, 1853.
- [14] J. P. Ward, *Quaternions and Cayley Numbers, Algebra and Applications*. Norwell, MA: Kluwer Academic, 1997.
- [15] S. J. Sangwine, "Biquaternion (complexified quaternion) roots of—1," *Adv. Appl. Clifford Alg.*, vol. 16, no. 1, pp. 63–68, 2006.
- [16] T. A. Ell and S. J. Sangwine, "Decomposition of 2-D hypercomplex Fourier transforms into pairs of complex Fourier transforms," in *Proc. Eur. Conf. Signal Image Processing (EUSIPCO)*, Tampere, Finland, 2000, pp. 1061–1064.
- [17] M. Reimer and D. Yevick, "A clifford algebra analysis of polarization-mode dispersion and polarization-dependent loss," *IEEE Photon. Technol. Lett.*, vol. 18, no. 6, pp. 734–736, 2006.
- [18] T. A. Ell and S. J. Sangwine, "Hypercomplex Wiener–Kintchine theorem with application to color image processing," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, Vancouver, Canada, 2000, pp. 792–795.
- [19] D. Gabor, "Theory of communication," *J. Inst. Elec. Eng.*, vol. 93, no. 26, pt. III, pp. 429–457, 1946.
- [20] J. Ville, "Théorie et applications de la notion de signal analytique," *Cables et Transmission*, vol. 2A, pp. 61–74, 1948.
- [21] R. N. Bracewell, *The Fourier Transform and Its Applications*, 3rd ed. Boston, MA: McGraw-Hill, 2000.



Salem Said was born in Paris, France, in 1983. He received the Engineering Diploma degree in telecommunications engineering and the Master's degree in image, signal, and speech processing from the Institut National Polytechnique de Grenoble (INPG), France, in 2005. He is currently working towards the Ph.D. degree at the Department of Images and Signals at the GIPSA-Lab (CNRS UMR 5083), Grenoble, France. His Ph.D. work is focusing on Clifford algebra and manifold-valued signals.



Nicolas Le Bihan was born in Morlaix, France, in 1974. He received the B.Sc. in physics from the Université de Bretagne Occidentale (UBO), Brest, France, in 1997 and the M.Sc. and Ph.D. degrees in signal processing from the Institut National Polytechnique de Grenoble (INPG), Grenoble, France, in 1998 and 2001, respectively.

Since 2002, he has been a Chargé de Recherche with the Centre National de la Recherche Scientifique (CNRS) and is working with the Department of Images and Signals at the GIPSA-Lab. (CNRS UMR 5083), Grenoble, France. His research interests include polarized signal processing using multilinear, geometric algebra and group theory techniques, and applications of signal processing in geophysics.



Stephen J. Sangwine (SM'90) was born in London, U.K., in 1956. He received the B.Sc. degree in electronic engineering from the University of Southampton, Southampton, U.K., in 1979 and the Ph.D. degree from the University of Reading, Reading U.K., in 1991.

He is a Senior Lecturer with the Department of Computing and Electronic Systems, University of Essex, Colchester, U.K. His interests include linear vector filtering and transforms of vector signals and images using hypercomplex algebras, and digital

hardware design.

On Properness Of Quaternion Valued Random Variables

By Pierre-Olivier Amblard and Nicolas Le Bihan

Laboratoire des Images et des Signaux,
CNRS UMR 5083,
961 Rue de la Houille Blanche,
Domaine Universitaire, BP 46,
38402 Saint Martin d'Hères Cedex, France
e-mail: Bidou.Amblard@lis.inpg.fr, Nicolas.Le-Bihan@lis.inpg.fr

Abstract

In this paper, we present the concept of properness for quaternion random variables and emphasize some *second order* geometrical consequences on the four dimensional probability distribution for such variables. Properness is stated in terms of the invariance of the variable distribution under *Clifford translations*.

1. Introduction

Properness of complex valued random variables and vectors is a well-known subject in signal processing and information theory (Neeser, F.D. & Massey, J.L. (1993)). In the last decade, geometric algebras, and among them quaternions, have found applications in signal and image processing. Examples are the modelling and the analysis of color images (Sangwine S.J. (1996)), the definition of quaternion valued Fourier transforms for greyscale images (Bülow & Sommer (2001)), or the quaternion representation of 3D- or 4D-sensor measurements for polarization modelling (Le Bihan, N. & Mars, J. (2004)). In all these works, variables, signals or images were considered as deterministic quantities. However, many applications need a stochastic modelling of the observed phenomena (*e.g.* polarized magnetic disturbances, electromagnetic waves carrying random codes, noise in color image processing...). In this paper, we examine the extension of properness to the case of quaternion random variables, and link it with some geometrical invariance properties of the distributions of such variables.

2. Quaternions and quaternion random variables

After a recall on the definition of quaternions and their properties, we introduce the concept of quaternion random variable and pay attention to its possible *vector* representations.

2.1. Definition and properties

Quaternions algebra is a four dimensional hypercomplex numbers system discovered by Sir R.W. Hamilton in 1843 (Hamilton W.R. (1843)). A quaternion q has a real part and a three dimensional imaginary part such as:

$$q = a + \mathbf{i}b + \mathbf{j}c + \mathbf{k}d \quad (2.1)$$

and with the following relations between the imaginary units: $\mathbf{i}\mathbf{j} = \mathbf{k} = -\mathbf{j}\mathbf{i}$ and $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\mathbf{j}\mathbf{k} = -1$.

Quaternions form a noncommutative division algebra, noted $\mathbb{H}$, so that for $q_1, q_2 \in \mathbb{H}$, $q_1q_2 \neq q_2q_1$ generally. The conjugate of $q \in \mathbb{H}$ is $\bar{q} = a - \mathbf{i}b - \mathbf{j}c - \mathbf{k}d$, its norm is $|q| = \sqrt{a^2 + b^2 + c^2 + d^2}$ and its inverse is $q^{-1} = \bar{q}/|q|^2$. Note that, as $\mathbb{C} \subset \mathbb{H}$, conjugation of a complex number z will be noted $\bar{z}$. A quaternion is called *pure* when its real part is null and *unit* if its norm equals one. If q is a *pure unit* quaternion, then: $|q|^2 = 1$ and $q^2 = -1$.

Euler formula extends to $\mathbb{H}$, so that any quaternion q can be written: $q = |q|e^{\xi\theta}$, where ξ is a *pure unit* quaternion usually called the axis; θ is the angle. Any *pure unit* quaternions μ can thus be written: $\mu = e^{\mu\frac{\pi}{2}}$.

Conjugation is an anti-involution over $\mathbb{H}$ ($\overline{\bar{q}_1\bar{q}_2} = \bar{q}_2\bar{q}_1$), but there exist three important involutions, noted $q_{\mathbf{i}}$, $q_{\mathbf{j}}$ and $q_{\mathbf{k}}$, defined as:

$$q_{\mathbf{i}} = -\mathbf{i}q\mathbf{i} \quad , \quad q_{\mathbf{j}} = -\mathbf{j}q\mathbf{j} \quad , \quad q_{\mathbf{k}} = -\mathbf{k}q\mathbf{k} \quad (2.2)$$

These involutions are isometries and are special cases of 4D rotations (see Coxeter H.S.M. (1946)).

In fact, for any quaternion q the mapping:

$$q \rightarrow e^{\mu_1\theta} q e^{-\mu_1\theta} \quad (2.3)$$

where μ_1 is a pure unit quaternion and $\theta \in [0, 2\pi)$, leaves invariant the plane spanned by $\{1, \mu_1\}$ while it performs a clockwise rotation of angle 2θ in the plane spanned by $\{\mu_2, \mu_3\}$, assuming $\{1, \mu_1, \mu_2, \mu_3\}$ is an orthonormal basis of $\mathbb{H}$. Note that for involutions in (2.2), the mapping axis are respectively $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ with angle $\theta = \pi/2$.

Particular isometries of interest in this work are the so called *Clifford translations*. There are two types of such translations (see Coxeter H.S.M. (1946)). A *left Clifford translation* is the mapping:

$$q \rightarrow e^{\mu_1 \theta} q \quad (2.4)$$

while the mapping $q \rightarrow q e^{\mu_1 \theta}$ is a *right Clifford translation*. A *left Clifford translation* performs a clockwise rotation of angle θ in the plane spanned by $\{1, \mu_1\}$ as well as in the plane spanned by $\{\mu_2, \mu_3\}$. However, a *right Clifford translation* performs a clockwise rotation in $\{1, \mu_1\}$ and an counterclockwise rotation in $\{\mu_2, \mu_3\}$ (both of angle θ).

It possible to consider quaternions as complexified complex numbers, such that: $q = z_1 + z_2 \mathbf{j}$ where $z_1, z_2 \in \mathbb{C}^{\mathbf{i}}$. Due to noncommutativity of quaternions product, the order in $z_2 \mathbf{j}$ is important. In this notation, known as the Cayley-Dickson notation, $z_1 = a + \mathbf{i}b$ and $z_2 = c + \mathbf{i}d$.

A quaternion can be seen as a four dimensional real vector, *i.e.* an element of $\mathbb{R}^4$. Also, it is possible to obtain the four real elements (a, b, c, d) from combinations of z_1 and z_2 and their conjugates, or from combinations of q and its three involutions. This allows to introduce for a quaternion random variable q , the three following *vector* representations:

$$\hat{q} = [a \ b \ c \ d]^T, \quad \tilde{q} = [z_1 \ \bar{z}_1 \ z_2 \ \bar{z}_2]^T, \quad \check{q} = [q \ q_{\mathbf{i}} \ q_{\mathbf{j}} \ q_{\mathbf{k}}]^T \quad (2.5)$$

These representations are linked the following way:

$$\tilde{q} = \mathbf{A}^{[\mathbb{RC}]} \hat{q} \text{ and } \check{q} = \mathbf{A}^{[\mathbb{CH}]} \tilde{q} \quad (2.6)$$

with:

$$\mathbf{A}^{[\mathbb{RC}]} = \begin{bmatrix} 1 & \mathbf{i} & 0 & 0 \\ 1 & -\mathbf{i} & 0 & 0 \\ 0 & 0 & 1 & \mathbf{i} \\ 0 & 0 & 1 & -\mathbf{i} \end{bmatrix}; \quad \mathbf{A}^{[\mathbb{CH}]} = \begin{bmatrix} 1 & 0 & 0 & \mathbf{j} \\ 1 & 0 & 0 & -\mathbf{j} \\ 0 & 1 & \mathbf{j} & 0 \\ 0 & 1 & -\mathbf{j} & 0 \end{bmatrix} \quad (2.7)$$

Vector representations are of interest in the study of a quaternion random variable as they allow easier geometrical interpretations in 4D space.

2.2. Quaternion random variables

A quaternion valued random variable is defined unambiguously as a real valued four dimensional random vector. As such, a quaternion random variable q is fully described by the joint probability density function (pdf) of the four components a, b, c, d of its vector representation $\hat{q}$, or equivalently by the characteristic function. However, special features of the pdf, such as symmetries under some transformations, may not be easily revealed by this representation. Precisely, using the complex representation $\tilde{q}$ or the quaternion representation $\check{q}$ may reveal more easily these features. Since $\hat{q}$, $\tilde{q}$ and $\check{q}$ are linked by relations (2.6), using the complex or the quaternion representations amounts to define the pdf of the quaternion variable on $\mathbb{C}^4$ and on $\mathbb{H}^4$. This can be done if one consider that $z_1, \bar{z}_1, z_2, \bar{z}_2$ are algebraically independent variables, and $q, q_{\mathbf{i}}, q_{\mathbf{j}}, q_{\mathbf{k}}$ are algebraically independent variables also. This can be rigorously formalized—as was done for complex variables in (Amblard, P.O. *et. al.* (1996)). However, we restrict in the following to first and second order statistics only.

When considering a quaternion random variable q , the mathematical expectation of q is given as follows:

$$\mathbb{E}[q] = \mathbb{E}[a] + \mathbb{E}[b]\mathbf{i} + \mathbb{E}[c]\mathbf{j} + \mathbb{E}[d]\mathbf{k} \quad (2.8)$$

where the expectation of real valued random variables (a, b, c, d) is taken in the classical sense. Without loss of generality, the considered quaternion random variables are supposed centered (*i.e.* $\mathbb{E}[q] = 0$) in the sequel. Using the *real vector* representation of q , noted $\hat{q}$, its covariance matrix is:

$$\Lambda_{\hat{q}} = \mathbb{E} \left[\hat{q} \hat{q}^T \right] \quad (2.9)$$

that contains second order statistical relationships between a, b, c and d . Using the two other representations, it is possible to define a *complex* and a *quaternion* representation of the covariance matrix given by:

$$\begin{cases} \Lambda_{\tilde{q}} = \mathbb{E}[\tilde{q} \tilde{q}^\dagger] \\ \Lambda_{\check{q}} = \mathbb{E}[\check{q} \check{q}^\dagger] \end{cases} \quad (2.10)$$

$\Lambda_{\tilde{q}}$ contains the second order cross-moments between $z_1, \bar{z}_1, z_2$ and $\bar{z}_2$, and $\Lambda_{\tilde{q}}$ the second order cross-moments between q, q_i, q_j and q_k . Operator † stands for conjugation-transposition.

3. Properness of quaternion random variables

In (Vakhania, N.N. (1998)), it was shown that there exists two levels of properness for a quaternion random variable, namely $\mathbb{C}$ - and $\mathbb{H}$ -properness. Vakhania proposed a definition of properness based on the fact that the real representation of the covariance matrix commutes with either the real matrix representation of $\mathbf{i}$ or with both the real representations of $\mathbf{i}$ and $\mathbf{j}$. This properness can be interpreted as the invariance of the pdf under some specific rotations of angle $\pi/2$. Here, we extend the definition to an arbitrary axis and angle and examine the consequences of so defined properness on the second order statistical relationships between components of the quaternion random variable.

3.1. $\mathbb{C}$ -properness

DEFINITION 1. A quaternion valued random variable q is called $\mathbb{C}^\eta$ -proper if:

$$q \stackrel{d}{=} e^{\eta\varphi} q, \quad \forall \varphi \quad (3.1)$$

for one and only one imaginary unit $\eta = \mathbf{i}, \mathbf{j}$ or $\mathbf{k}$.

Clearly, a $\mathbb{C}^\eta$ -proper quaternion random variable has a distribution that is invariant by *left Clifford translation* of axis η and of any angle φ (i.e. simultaneous rotations in of angle φ in two orthogonal planes of 4D space). As an example, we now study the case of a $\mathbb{C}^{\mathbf{i}}$ -proper quaternion random variable.

PROPERTY 1. The real, complex and quaternion representation covariance matrices of a $\mathbb{C}^{\mathbf{i}}$ -proper quaternion random variable q have the following structures:

$$\Lambda_{\tilde{q}} = \begin{bmatrix} \sigma_1^2 & 0 & \tau_\beta & \tau_\gamma \\ 0 & \sigma_1^2 & \tau_\gamma & -\tau_\beta \\ \tau_\beta & \tau_\gamma & \sigma_2^2 & 0 \\ \tau_\gamma & -\tau_\beta & 0 & \sigma_2^2 \end{bmatrix} \text{ and } \Lambda_{\tilde{q}} = \begin{bmatrix} \Sigma_1^2 & 0 & 0 & \Omega \\ 0 & \Sigma_1^2 & \Omega & 0 \\ 0 & \Omega & \Sigma_2^2 & 0 \\ \Omega & 0 & 0 & \Sigma_2^2 \end{bmatrix} \text{ and } \Lambda_{\tilde{q}} = \begin{bmatrix} \Sigma & \Delta & 0 & 0 \\ \Delta_{\mathbf{i}} & \Sigma & 0 & 0 \\ 0 & 0 & \Sigma & \Delta \\ 0 & 0 & \Delta_{\mathbf{i}} & \Sigma \end{bmatrix} \quad (3.2)$$

where $\sigma_1^2 = \mathbb{E}[a^2] = \mathbb{E}[b^2]$, $\sigma_2^2 = \mathbb{E}[c^2] = \mathbb{E}[d^2]$, $\tau_\beta = \mathbb{E}[ac] = -\mathbb{E}[bd]$ and $\tau_\gamma = \mathbb{E}[ad] = \mathbb{E}[bc]$ are real coefficients corresponding to cross-covariances between pairs of variables (a, b) and (c, d) . We also have $\Sigma_1^2 = \mathbb{E}[|z_1|^2] = 2\sigma_1^2$, $\Sigma_2^2 = \mathbb{E}[|z_2|^2] = 2\sigma_2^2$ and $\Omega = \mathbb{E}[z_1 z_2] = 2(\tau_\beta + \tau_\gamma)$. Finally, we have $\Sigma = \mathbb{E}[|q|^2] = \mathbb{E}[|q_{\mathbf{i}, \mathbf{j}, \mathbf{k}}|^2] = 2\sigma_1^2 + 2\sigma_2^2$ and $\Delta = \mathbb{E}[q \bar{q}_{\mathbf{i}}] = \mathbb{E}[q_{\mathbf{j}} \bar{q}_{\mathbf{k}}] = 2\sigma_1^2 - 2\sigma_2^2 + 2(\tau_\beta + \mathbf{i}\tau_\gamma)\mathbf{j}$.

Proof. $\mathbb{C}^{\mathbf{i}}$ -properness of q involves that $\mathbb{E}[q \bar{q}] = e^{\mathbf{i}\varphi} \mathbb{E}[q \bar{q}] e^{-\mathbf{i}\varphi}$. Using vector representations of q given in (2.5) and transition matrices in (2.7), the structures of $\Lambda_{\tilde{q}}$, $\Lambda_{\tilde{q}}$ and $\Lambda_{\tilde{q}}$ come out by straightforward calculation. $\square$

Thus, a $\mathbb{C}^{\mathbf{i}}$ -proper quaternion random variable q is correlated with the variable $q_{\mathbf{i}}$ while it is decorrelated with variables $q_{\mathbf{j}}$ and $q_{\mathbf{k}}$. $\mathbb{C}^{\mathbf{i}}$ -properness involves an invariance of the distribution under *left Clifford translation*. It means that the distribution is left invariant by simultaneous rotations of angle $\pi/2$ in the planes spanned by $\{1, \mathbf{i}\}$ and $\{\mathbf{j}, \mathbf{k}\}$. Looking at the complex representation of q , $\mathbb{C}^{\mathbf{i}}$ -properness is equivalent to the second order circularity of both z_1 and z_2 . Recall that a complex random variable z is circular if its pdf is invariant under any rotations; second order circularity of z is achieved if the real and imaginary parts of z have same variance and are uncorrelated. Furthermore, $\mathbb{C}^{\mathbf{i}}$ -properness does not require uncorrelation between z_1 and z_2 . As an example, consider a Gaussian $\mathbb{C}^{\mathbf{i}}$ -proper random variable with $\sigma_1^2 = 1, \sigma_2^2 = 1.5, \tau_\beta = 0.7$ and $\tau_\gamma = 0.2$. In figure 1, we plot 10^4 samples of this variable. Looking at planes $\{1, \mathbf{i}\}$ and $\{\mathbf{j}, \mathbf{k}\}$, we can see that z_1 and z_2 are both circular: the distributions are invariant under rotations. However, the correlation between z_1 and z_2 is revealed by looking at the distributions in the planes $\{1, \mathbf{j}\}$ and $\{\mathbf{i}, \mathbf{k}\}$: the ellipsis parameter are governed by τ_β , whereas looking at the distributions in the planes $\{1, \mathbf{k}\}$ and $\{\mathbf{i}, \mathbf{j}\}$ reveals parameter τ_γ .

3.2. $\mathbb{H}$ -properness

DEFINITION 2. A quaternion random variable q is said to be $\mathbb{H}$ -proper if:

$$q \stackrel{d}{=} e^{\eta\varphi} q, \quad \forall \varphi \quad (3.3)$$

and for any pure unit quaternion η .

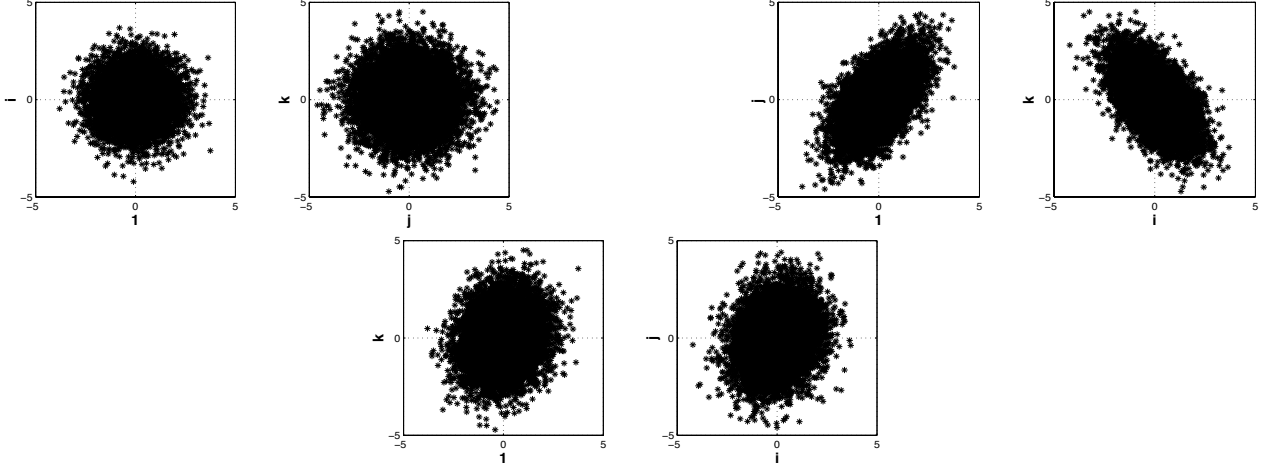


FIGURE 1. $\mathbb{C}$ -proper Gaussian random variable. Three possible representations: 1- planes $\{1, i\}$ and $\{j, k\}$ (upper-left), 2- planes $\{1, j\}$ and $\{i, k\}$ (upper-right), 3- planes $\{1, k\}$ and $\{i, j\}$ (lower-center).

In this case, the real, complex and quaternion representations of the covariance matrix have the following structure:

$$\Lambda_{\mathbf{q}} = \sigma^2 \mathbf{I}_4 ; \Lambda_{\bar{\mathbf{q}}} = 2\sigma^2 \mathbf{I}_4 ; \Lambda_{\mathbf{q}} = 4\sigma^2 \mathbf{I}_4 \quad (3.4)$$

where $\mathbf{I}_4$ is the 4×4 identity matrix and $\sigma^2 = \mathbb{E}[a^2] = \mathbb{E}[b^2] = \mathbb{E}[c^2] = \mathbb{E}[d^2]$. Clearly, the distribution of a $\mathbb{H}$ -proper quaternion random variable is invariant under any four dimensional isometric transformation. In the Gaussian case, the distribution of a $\mathbb{H}$ -proper variable is contained in a 4D hypersphere.

4. Discussion

We have extended a definition of properness for quaternion valued random variables based on the invariance of the pdf under the action of *left Clifford translations*. This allows to go deeper in the study of the symmetries of the pdf. We have however restricted the implications of properness to the second order statistics. Of course, for Gaussian variables the analysis is full since Gaussian are entirely described by the second order statistics. Further work will consist in the use of higher order statistics to characterize some 4D geometrical properties of quaternion valued random variable distributions. For example, we are working on the notion of n -th order properness, for which the rotational invariance are no longer continuous but rather discrete $2\pi/n$. This notion will be of importance in the study of 4D constellations used for communication purposes, see for example (Zetterberg, L. H. & Brändström (1977)).

REFERENCES

- AMBLARD, P.O. & LACOUME J.L. & GAETA M. 1996 Statistics for Complex Random Variables and Signals : Part 1 and 2 *Signal Processing* Vol. 53, pp. 1–25.
- NEESER, F.D. & MASSEY, J.L. 1993 Proper Complex Random Process with applications to information theory *IEEE Trans. on Information Theory* Vol. 39, No. 4, pp. 1293–1302.
- HAMILTON W.R. 1843 On quaternions *Proceeding of the Royal Irish Academy*
- COXETER H.S.M. 1946 Quaternions and reflections *The american mathematical monthly* Vol. 53, pp. 136–146.
- SANGWINE, S. J. 1996 Fourier transforms of colour images using quaternions, or hypercomplex, numbers *Electronics letters* Vol. 32, No. 21, pp. 1979–1980.
- BÜLOW, T. & SOMMER, G. 2001 Hypercomplex Signals– A Novel Extension of the Analytic Signal to the Multidimensional Case *IEEE Trans. on Signal Processing* vol. 49, No. 11, pp. 2844–2852.
- LE BIHAN, N. & MARS, J. 2004 Singular Value Decomposition of matrices of quaternions: a new tool for vector-sensor signal processing *Signal Processing* Vol. 2004.
- VAKHANIA, N.N. 1998 Random vectors with values in quaternions Hilbert spaces *Th. Probab. Appl.* Vol. 43, No. 1, pp. 99–115
- ZETTERBERG, L. H. & BRÄNDSTRÖM, H. 1977 Codes for combined phase and amplitude modulated signals in a four-dimensional space *IEEE Trans. on Communications* Vol. 25, No. 9, pp. 943–950.

Signaux quaternioniques

Sommaire

2.1	Signaux quaternioniques et polarisation	63
2.2	Méthodes par sous-espaces et traitement d'antenne	65
2.3	Signal hyperanalytique	70
2.4	Conclusion	73
2.5	Publications annexées en lien avec ce chapitre	74
2.5.1	"MUSIC algorithm for vector-sensors array using biquaternions" IEEE TSP 2007	75
2.5.2	"The $\mathbb{H}$ -analytic signal" EUSIPCO 2008	86
2.5.3	"Quaternionic independent component analysis using hypercomplex nonlinearities" IMA 2006	91

Dans ce chapitre, nous présentons quelques exemples de contributions à l'utilisation des quaternions en traitement du signal. Un choix arbitraire a été pour ne présenter que quelques contributions. De plus, en fin de chapitre, nous avons inclus un article sur l'analyse en composantes indépendantes quaternionique [Le Bihan 2006b], sujet qui n'est pas détaillé dans le corps de ce chapitre. Cette inclusion a été faite à titre d'illustration des autres travaux que j'ai mené sur le thème du traitement des signaux quaternioniques. Mes autres contributions dans le domaine sont [Le Bihan 2004, Miron 2006b, Buchholz 2008, Menanno 2010, Le Bihan 2001, Bas 2003, Le Bihan 2003, Miron 2005, Miron 2006a, Buchholz 2006, Le Bihan 2006a, Javidi 2011].

Dans ce chapitre, les travaux concernant le traitement d'antenne vectorielle ont été menés dans le cadre des thèses de Sebastian MIRON et Manuel HOBIGER.

2.1 Signaux quaternioniques et polarisation

Il y a plusieurs façons d'obtenir des signaux quaternioniques. On peut construire un signal quaternionique à partir de trois ou quatre signaux réels (ou complexes dans le cas biquaternionique). Une telle construction n'a d'intérêt que si les signaux sont reliés entre eux et si ces relations sont décrites de manière simple sur $\mathbb{H}$. Historiquement, l'utilisation des quaternions pour la modélisation de signaux temporels a commencé avec des séries temporelles dans $\mathbb{R}^3$ [Kuipers 1999]. On peut également mentionner les séries temporelles de rotations (éléments de $SO(3)$) qui peuvent être facilement implémentée sur $\mathbb{H}$. L'étude des séries temporelles de rotations sera abordée dans la partie II.

On peut aussi mentionner la modélisation des signaux issus de capteurs sismiques OBS (Ocean Bottom Seismometer) et OBC (Ocean Bottom Cable) qui sont constitués d'un accéléromètre à trois composantes (qui enregistrent les vibrations dans les trois directions de l'espace) et d'un hydrophone (capteur de pression). Les signaux de ces capteurs fournissent donc un signal vectoriel 3D et un signal scalaire 1D. Modéliser ces signaux 4D comme des signaux quaternioniques permet d'étendre les algorithmes classiques pour ces

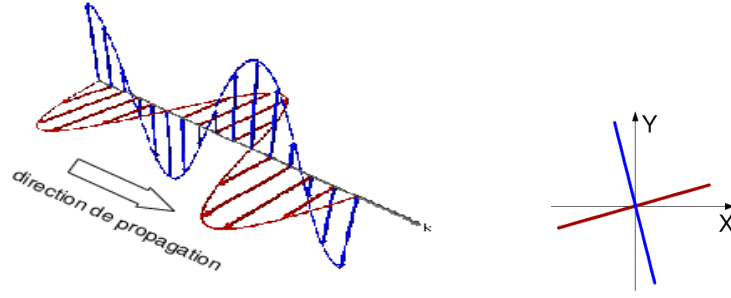


FIGURE 2.1 – Ondes polarisées linéairement se propageant suivant la direction k (gauche). Tracé paramétrique (paramètre t) dans le plan de polarisation à partir des deux composantes $x(t)$ et $y(t)$ pour deux ondes de vibrations orthogonales entre elles (bleu et rouge).

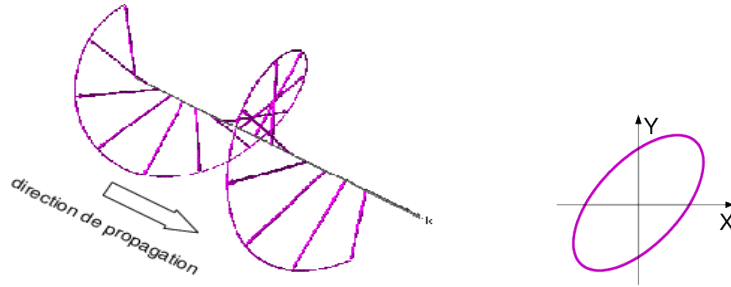


FIGURE 2.2 – Onde polarisée elliptiquement se propageant suivant la direction k (gauche). Tracé paramétrique (paramètre t) dans le plan de polarisation (plan orthogonal à k) à partir des deux composantes $x(t)$ et $y(t)$.

signaux particuliers [Le Bihan 2004]. Toutefois, la construction d’un signal quaternionique paraît artificielle dans ce cas précis, car les règles de composition (addition/multiplication) des quaternions ne correspondent pas à des “transformations” physiques des signaux. La référence [Le Bihan 2004] est toutefois intéressante en ce sens qu’elle donne les bases des méthodes par sous-espaces sur $\mathbb{H}$.

Il est également possible de construire un signal quaternionique à partir de deux signaux complexes. Cette construction peut s’avérer intéressante quand on cherche à traiter les signaux représentant des ondes polarisés par exemple. Si l’on considère une onde polarisée (onde électromagnétique ou onde élastique de cisaillement), on peut la caractériser par sa direction de propagation k et dans le plan orthogonal à k , l’onde est complètement décrite¹ par ses deux composantes dans le plan, souvent notées $E_x(t)$ et $E_y(t)$ dans le cas des ondes électromagnétiques. Nous utiliserons plus simplement $x(t)$ et $y(t)$. Afin de mesurer les rapports de phase et d’amplitude entre ces signaux (paramètres de polarisation), ces deux signaux sont à valeurs complexes (les signaux analytiques associés aux signaux réels mesurés). Ainsi, un couple de signaux complexes décrit un signal polarisé (à condition de connaître le plan de polarisation, *c.à.d.* le plan orthogonal à k).

1. De manière rigoureuse, ceci est vrai à une fréquence donnée. La polarisation, et donc les paramètres de polarisation, sont définis entre deux signaux de même fréquence.

Si on note $x(t)$ et $y(t)$ ces deux signaux, on est alors tenté de construire un signal quaternionique qui décrit le signal total polarisé $q(t)$:

$$q(t) = x(t) + jy(t)$$

Dans le cas de signaux aléatoires représentant une onde purement polarisée², les deux signaux $x(t)$ et $y(t)$ ne diffèrent qu'en module et phase (qui eux sont déterministes). En supposant ici que ce rapport est constant, l'information de polarisation se résume à un nombre complexe $\rho e^{i\theta}$ tel que $y(t) = \rho e^{i\theta} x(t)$. Dans ce cas, en utilisant la notation vectorielle complexe $q_{\mathbb{C}}(t) = [z_1 \ z_1^* \ z_2 \ z_2^*]$ introduite en 1.2.2, on peut voir que la matrice de covariance de q est :

$$\Lambda_q = \begin{bmatrix} \sigma_x^2 & 0 & \rho \cos \theta \sigma_x^2 & -\rho \sin \theta \sigma_x^2 \\ 0 & \sigma_x^2 & -\rho \sin \theta \sigma_x^2 & -\rho \cos \theta \sigma_x^2 \\ \rho \cos \theta \sigma_x^2 & -\rho \sin \theta \sigma_x^2 & \rho^2 \sigma_x^2 & 0 \\ -\rho \sin \theta \sigma_x^2 & -\rho \cos \theta \sigma_x^2 & 0 & \rho^2 \sigma_x^2 \end{bmatrix}$$

avec σ_x^2 la variance de $x(t)$. Cette structure de matrice de covariance est caractéristique d'un signal quaternionique $\mathbb{C}^1$ -circulaire (voir [Amblard 2004, Buchholz 2008, Via 2010b] pour les expressions des matrices de covariance des représentations vectorielles d'un processus quaternionique). On peut de la même façon montrer qu'un signal non polarisé aura, dans son expression quaternionique, une matrice de covariance typique d'un processus $\mathbb{H}$ -circulaire [Buchholz 2006, Buchholz 2008], *c.à.d.* un multiple de $\mathbf{I}_4$, la matrice identité dans $\mathbb{R}^4$.

Cette distinction $\mathbb{C}$ -circulaire / $\mathbb{H}$ -circulaire permet d'élaborer des algorithmes de traitement des signaux polarisés, comme du débruitage, via de l'analyse en composantes indépendantes [Le Bihan 2006b] (article inclus dans la section 2.5) ou des réseaux de neurones quaternioniques [Buchholz 2006, Buchholz 2008] par exemple. Récemment, certains auteurs ont développé des techniques statistiques de détection de circularité afin de discriminer entre $\mathbb{C}$ - et $\mathbb{H}$ -circularité [Via 2010a].

2.2 Méthodes par sous-espaces et traitement d'antenne

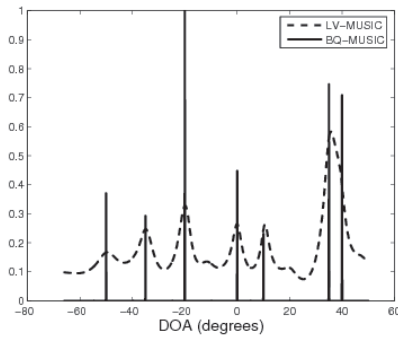
Les travaux que nous avons menés sur les décompositions de matrices quaternioniques et biquaternioniques ont été utilisés pour proposer des algorithmes de séparation de champs d'ondes polarisées [Le Bihan 2004], de traitement d'antenne polarisée Haute Résolution [Miron 2006b, Le Bihan 2007a] ou des algorithmes de déconvolution de signaux polarisés [Le Bihan 2005, Menanno 2010].

Pour les antennes, nous avons proposé des versions quaternioniques et biquaternioniques de l'algorithme MUSIC, défini à l'origine par Schmidt [Schmidt 1986]. La version quaternionique permet de traiter les signaux issus d'antennes à deux composantes, *i.e.* dans lesquelles chaque capteur enregistre les vibrations dans deux directions (orthogonales en général) de l'espace. La version biquaternionique autorise le traitement simultané de trois composantes. Ainsi, on peut construire une matrice spectrale (bi)quaternionique à l'aide des deux (trois) matrices spectrales calculées sur les deux (trois) composantes. L'intérêt d'utiliser les (bi)quaternions vient du fait que cette construction préserve la phase entre les composantes pendant le traitement, *c.à.d.* l'information de polarisation.

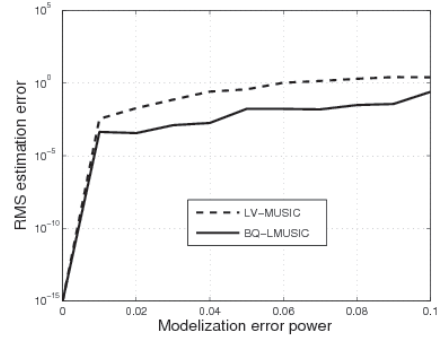
Nous avons montré la supériorité des techniques quaternioniques et biquaternioniques par rapport aux techniques de type *long-vecteur* qui consistent à concaténer les signaux

2. Voir la section 3.1 du chapitre 3, ou la référence [Brosseau 1998], pour plus de détails sur les notions liées à la polarisation.

des différentes composantes en un seul grand vecteur. Cette supériorité est flagrante dans la robustesse aux erreurs de modèle. À titre d'exemple, on reprend ici sur la figure 2.3 la comparaison entre BQ-Music (Biquaternion Music) et LV-Music (Long-Vecteur Music) pour l'estimation des paramètres et l'estimation de la direction d'arrivée dans un cas où les paramètres de polarisation sont biaisés (erreur de modèle). L'erreur est de 5% et on peut remarquer la perte de résolution de de l'approche LV 2.3(a) ainsi que l'accroissement de l'erreur d'estimation des paramètres de polarisation 2.3(b).



(a) Robustesse au biais dans le modèle de polarisation.



(b) Erreur d'estimation du paramètre de polarisation en fonction du rapport signal sur bruit.

FIGURE 2.3 – Comportement de l'algorithme BQ-Music vis à vis de l'approche Long-Vecteur dans le cas où le modèle est biaisé.

Ces figures sont prises de l'article [Le Bihan 2007a] qui est inclus ici à la section 2.5. Comme expliqué dans [Le Bihan 2007a], la supériorité de l'approche biquaternionique est due à la contrainte d'orthogonalité entre vecteurs biquaternioniques, qui permet entre autre de conserver une résolution théoriquement infinie de l'algorithme BQ-MUSIC, même dans des configurations où il y a des erreurs de modélisation.

Dans le cas quaternionique, nous avons montré récemment l'intérêt d'utiliser l'algorithme Quaternion-Music, combiné avec l'algorithme Music classique [Schmidt 1986], pour l'estimation des paramètres des ondes de Rayleigh et leur utilisation dans l'identification des paramètres du sous-sol (vitesses des ondes). Avec cet algorithme [Hobiger 2011], il est possible d'estimer l'ellipticité (paramètre de polarisation) de ces ondes, ainsi que leur sens de rotation (prograde ou rétrograde) de manière directe (expression analytique des estimateurs), ce qui apporte une information précieuse dans l'identification du milieu de propagation local (estimation de la variation de la vitesse des ondes en fonction de la profondeur).

La problématique que nous considérons ici peut être résumée ainsi : à partir des signaux polarisés enregistrés sur une antenne de capteurs vectoriels, estimer la courbe de dispersion des ondes de Love et Rayleigh ainsi que la courbe d'ellipticité des ondes de Rayleigh. Ces courbes³ sont ensuite utilisées pour obtenir une estimation du profil de vitesse du milieu (via un algorithme d'inversion).

Afin d'illustrer les performances de cet algorithme, nous présentons les résultats obtenus par MUSIQUE pour l'estimation des ondes de surface, sur des données synthétiques obtenues par le code de simulation en nombres d'ondes Hisada [Hisada 1995].

Nous considérons un milieu dont le profil est présenté sur la figure 2.4(a). Ce profil donne

3. En fait, seules les courbes des ondes de Rayleigh sont utilisées dans l'algorithme d'inversion. L'information sur les ondes de Love est complémentaire sans être indispensable à l'étape d'inversion du milieu.

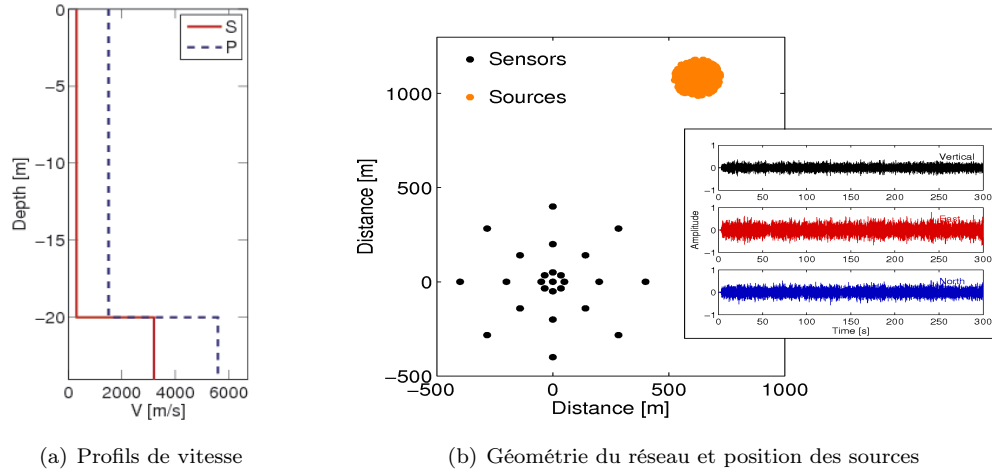


FIGURE 2.4 – Paramètres physiques utilisés pour la simulation et position du réseau de capteurs et des sources sismiques.

la vitesse de propagation des ondes P et S en fonction de la profondeur. La géométrie du réseau de capteurs est présentée sur la figure 2.4(b).

Dans ce modèle de simulation, on connaît les profils des ondes de Love et Rayleigh attendus (cette information est donnée par la connaissance des vitesses des ondes dans le milieu). On a donc la courbe théorique de dispersion (vitesse en fonction de la fréquence) des ondes de Love et Rayleigh, ainsi que la courbe théorique d'ellipticité des ondes de Rayleigh (la courbe d'ellipticité traduit le déphasage entre les composantes verticale et horizontale de l'onde, en fonction de la fréquence).

Sur les figures 2.5(a) et 2.5(b), on présente les courbes de dispersion des ondes de Rayleigh et Love ainsi que la courbe d'ellipticité des ondes de rayleigh. Les différents sens de rotation des ondes (prograde et rétrograde) sont indiqués en fonction de la fréquence.

Les signaux simulés par le code Hisada sont des signaux vectoriels (3 composantes) dont la durée était ici de 300 secondes. L'algorithme MUSIQUE travaille à bande étroite, de qui

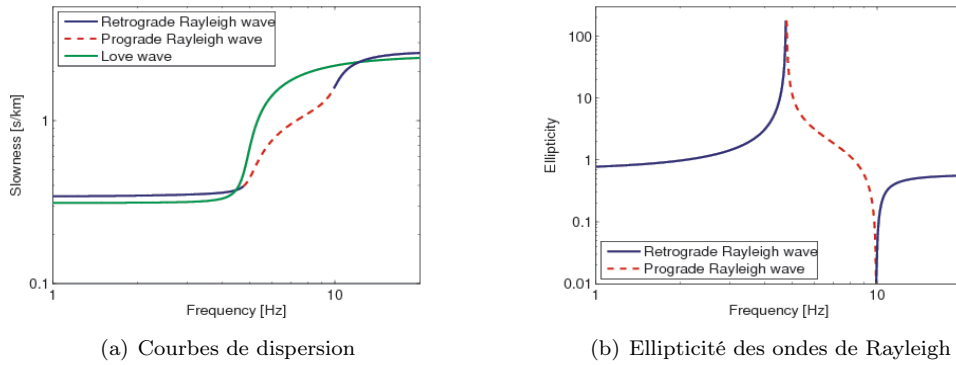
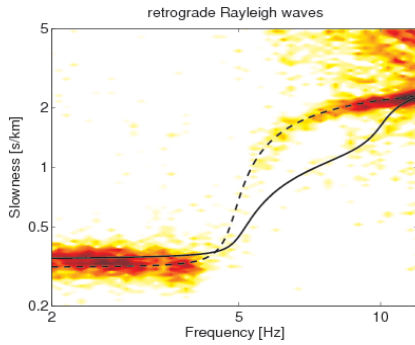


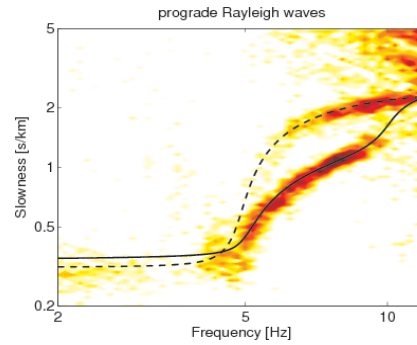
FIGURE 2.5 – Courbes théoriques de dispersion et d'ellipticité issues du modèle 2.4(a).

veut dire que pour chaque fréquence, une première étape d'identification de direction d'arrivée est effectuée (Music classique) en même temps qu'une identification de la vitesse des ondes détectées. Ensuite, la version quaternionique de MUSIC identifie les différents types d'ondes (Love ou Rayleigh, prograde ou rétrograde) et permet d'identifier la contribution de chaque type d'onde.

En travaillant fréquence par fréquence, l'algorithme MUSIQUE permet d'estimer les courbes de dispersion et la courbe d'ellipticité. Des résultats obtenus sur les signaux simulés sont présentés sur les figures 2.6(a), 2.6(b), 2.7, 2.8(a) and 2.8(b).



(a) Courbe de dispersion des ondes de Rayleigh rétrogrades estimée



(b) Courbe de dispersion des ondes de Rayleigh progrades estimée

FIGURE 2.6 – Courbes de dispersion des ondes de Rayleigh estimées par MUSIQUE (à mettre en comparaison avec la figure 2.5(a)).

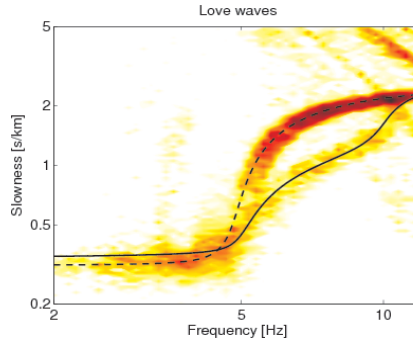


FIGURE 2.7 – Courbe de dispersion des ondes de Love estimée par MUSIQUE (à mettre en comparaison avec la figure 2.5(a)).

Sur les figures 2.6(a) et 2.6(b), on voit les courbes de dispersion des ondes de Rayleigh rétrogrades et progrades estimées par MUSIQUE. Ces courbes sont en accord avec les courbes théoriques 2.5(b), malgré des difficultés d'estimation pour les hautes fréquences (au delà de 10 Hz). En particulier, sur les figures 2.6(a) et 2.6(b), on voit que dans les hautes fréquences, l'algorithme ne parvient pas à distinguer entre les ondes de Rayleigh et les ondes de Love. C'est également le cas sur la figure 2.7 où malgré une estimation correcte de la courbe de dispersion de l'onde de Love, il subsiste des erreurs de classification, l'algorithme ne pouvant pas distinguer entre Love et Rayleigh. Ce problème récurrent est d'ailleurs à la

base de la recherche de nouveaux algorithmes de séparation entre ces deux types d'ondes. En effet, bien identifier les ondes de Love est une garantie d'une bonne estimation des ondes de Rayleigh qui sont ensuite utilisées pour l'inversion et l'estimation des paramètres du milieu.

On voit sur les courbes 2.8(a) et 2.8(b) que l'ellipticité des ondes de Rayleigh est relativement correctement retrouvée. L'estimation obtenue par MUSIQUE est correcte, et ce sur une grande partie de la bande fréquentielle.

Ces résultats sont meilleurs que ceux utilisant la technique dite "H/V"⁴ qui consiste à estimer l'ellipticité par un rapport entre les modules de TF des composantes verticales et horizontales. Cette technique est systématiquement biaisée par la présence d'ondes de Love en plus des Rayleigh. L'algorithme MUSIQUE pallie ce problème en permettant une séparation des ondes de Love et Rayleigh.

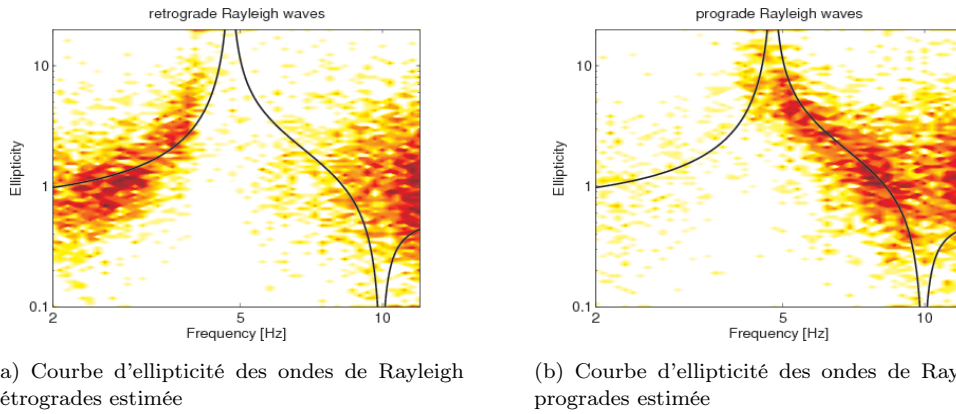


FIGURE 2.8 – Courbes d'ellipticité des ondes rétrogrades et progrades des ondes de Rayleigh estimées par MUSIQUE (à mettre en comparaison avec la figure 2.5(b)).

Les résultats de MUSIQUE sur ces données simulées montre la potentialité des approches quaternioniques à traiter les signaux polarisés. L'intérêt de tels algorithmes est qu'ils généralisent de manière naturelle les algorithmes connus pour les signaux à échantillons scalaires. Ils permettent d'incorporer facilement l'information de polarisation, d'en tirer avantage (rapport signal à bruit) dans l'estimation des paramètres du modèle et donc de tirer avantage simultanément de toutes les diversités de l'antenne (diversité spatiale et polarisation).

L'utilisation de techniques haute résolution pour les antennes polarisées en sismologie est très récente. La spécificité des ondes sismiques, et particulièrement les ondes de surface (Rayleigh et Love), permet d'exploiter les avantages de la version quaternionique MUSIQUE. Une validation sur données réelles provenant d'un réseau sismologique en Californie a été pratiquée. Les résultats, exposés dans la thèse de M. Hobiger [Hobiger 2011], valident l'algorithme pour l'identification d'événements sismiques, mais d'autres types de données sont nécessaires pour tirer avantage des capacités de MUSIQUE à estimer les courbes caractéristiques des ondes de surface.

4. Pour plus de détails sur l'algorithme "H/V", voir [Hobiger 2011].

2.3 Signal hyperanalytique

La notion de signal analytique est connue depuis longtemps [Gabor 1946, Ville 1948] et ses applications en traitement du signal sont nombreuses (voir par exemple l'ouvrage [Schreier 2010] et les références qu'il contient). Ici, nous présentons une extension de cette notion de signal analytique pour les signaux qui sont déjà à valeurs complexes, mais non-circulaires⁵. L'intérêt est de proposer une approche alternative aux *représentations augmentées* classiquement adoptées pour les signaux non-circulaires [Schreier 2010]. Ces représentations utilisent simultanément le signal complexe et son conjugué. Par exemple, dans le cas non-circulaire, le problème d'estimation linéaire devient un problème d'*estimation linéaire au sens-large* [Picinbono 1995], faisant intervenir le signal et son conjugué.

Un constat simple est au centre de l'idée du signal hyperanalytique : si on considère un signal complexe $z(t)$ non-circulaire, sa Transformée de Fourier (TF) ne satisfait plus la symétrie Hermitienne et les parties paires et impaires de $\Re(z(t))$ et $\Im(z(t))$ se retrouvent mélangées dans les parties réelles et imaginaires de la TF. En fait, pour autoriser des symétries pour la TF de $z(t)$, il faut qu'elle soit définie sur un espace de dimension plus grande, en l'occurrence de dimension 4. La construction d'un signal "analytique" passe par l'utilisation d'une TF possédant des symétries. Le signal hyperanalytique doit donc se définir à partir d'une TF quaternionique.

On peut montrer aisement (voir [Le Bihan 2008] par exemple) que les différentes parties paires et impaires de $z(t)$ se retrouvent isolées dans les 4 parties de sa TF quaternionique. Cette TFQ satisfait alors la relation de symétrie évoquée au chapitre précédent (voir 1.4), en fonction de l'axe de la TFQ.

Pour le signal complexe $z(t)$, sa TFQ_j (TFQ d'axe j) est donnée par :

$$Z(\nu) = TFQ_j [z(t)] = \int_{-\infty}^{\infty} z(t) e^{-2j\nu t} dt$$

Le signal hyperanalytique peut être construit *via* la technique de mise à zéros des fréquences négatives. On peut montrer [Le Bihan 2008] que cela revient à construire un signal quaternionique (obtenu par TFQ inverse du spectre droit). À un signal complexe, il correspond un unique signal quaternionique *via* cette opération. Le signal hyperanalytique est ainsi constitué de deux signaux complexes (via la notation de Cayley-Dickson) qui sont orthogonaux entre eux. D'autres propriétés de ce signal sont décrites dans l'article [Le Bihan 2008] qui est inclus dans la section 2.5.

Nous nous intéressons maintenant à l'information apportée par le signal hyperanalytique. Dans le cas classique, il est bien connu que le signal analytique associé à un signal réel permet d'obtenir l'amplitude instantanée, la phase instantanée et la fréquence instantanée du signal original.

Dans le cas du signal hyperanalytique, il existe des équivalents. On peut identifier un module et une phase, ainsi qu'un *axe*. Ces trois objets sont obtenus via les représentations polaires des quaternions (voir le chapitre 1) : la représentation polaire de Cayley-Dickson pour l'enveloppe complexe, et la représentation polaire pour la phase.

Si l'on note $z_h(t)$ le signal hyperanalytique associé à $z(t)$, alors on a l'identité suivante :

$$z_h(t) = TFQ_j^{-1} [(1 + \text{sign}(\nu)) Z(\nu)]$$

On démontre aisement [Le Bihan 2008] que le signal $z_h(t)$ est à valeurs quaternioniques et qu'il peut s'écrire comme la somme du signal $z(t)$ et de sa transformée de Hilbert (calculée avec la TFQ) dans un plan complexe orthogonal au plan complexe de $z(t)$.

5. Les signaux analytiques au sens classique du terme sont circulaires.

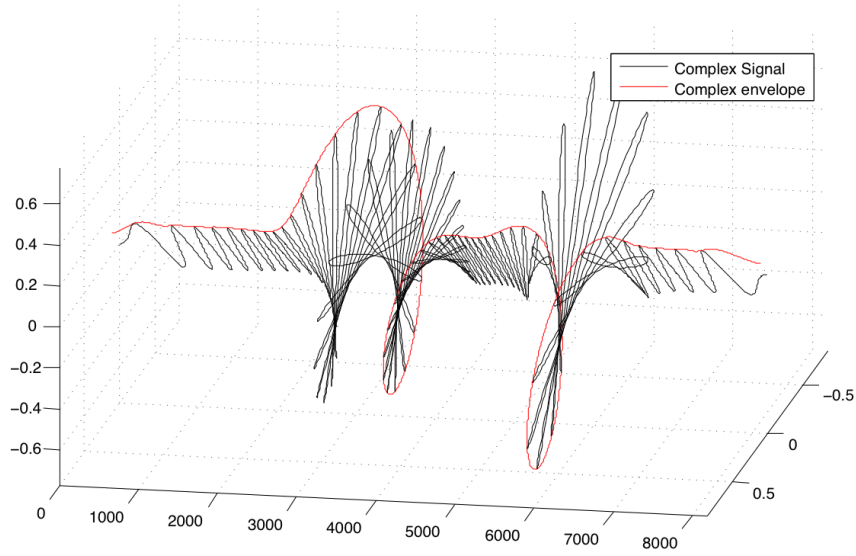


FIGURE 2.9 – Signal complexe non-circulaire (en noir) et son amplitude instantanée (en rouge) donnée par l’enveloppe complexe du signal hyperanalytique associé.

Afin d’illustrer les informations contenues dans le signal hyperanalytique, nous présentons les concepts d’enveloppe complexe et de phase instantanée sur un exemple de signal complexe simple. Le signal $z(t)$ est présenté sur la figure 2.9. C’est un chirp⁶ modulé en amplitude par des gaussiennes réelles (enveloppe) et par une exponentielle complexe non-stationnaire (phase).

L’enveloppe complexe de $z_h(t)$ est donnée par le module de la forme de Cayley-Dickson polaire [Sangwine 2010]. Elle est représentée en rouge sur la figure 2.9. Comme dans le cas classique du signal analytique, elle représente les oscillations lentes du signal, mais ici, elle est à valeurs complexes. L’enveloppe de $z_h(t)$ est l’amplitude instantanée (complexe) de $z(t)$.

On s’intéresse maintenant à la “phase instantanée” ou la “fréquence instantanée” de $z(t)$. Ici, quelques précautions s’imposent du fait que le signal $z_h(t)$ est à valeurs quaternioniques. Si l’on écrit $z_h(t)$ dans sa forme polaire :

$$z_h(t) = \|z_h(t)\| e^{\eta_z(t) \varphi_z(t)}$$

on voit que la “phase instantanée” de $z(t)$ peut être simplement définie comme l’argument de $z_h(t)$. On notera que ce n’est pas un scalaire, mais un bivecteur.

Maintenant, pour en déduire une fréquence instantanée, quelques considérations géométriques simples permettent d’appréhender cette notion de fréquence qui est différente du cas classique du fait de la dimensionalité du signal original.

On travaille sur la version normée de $z_h(t)$, i.e. $\bar{z}_h(t) = z_h(t)/\|z_h(t)\|$. On a alors directement par un calcul de mécanique classique (dont la formulation quaternionique est

6. Le signal consiste en réalité d’une succession de deux chirps, un “montant” et l’autre “descendant”.

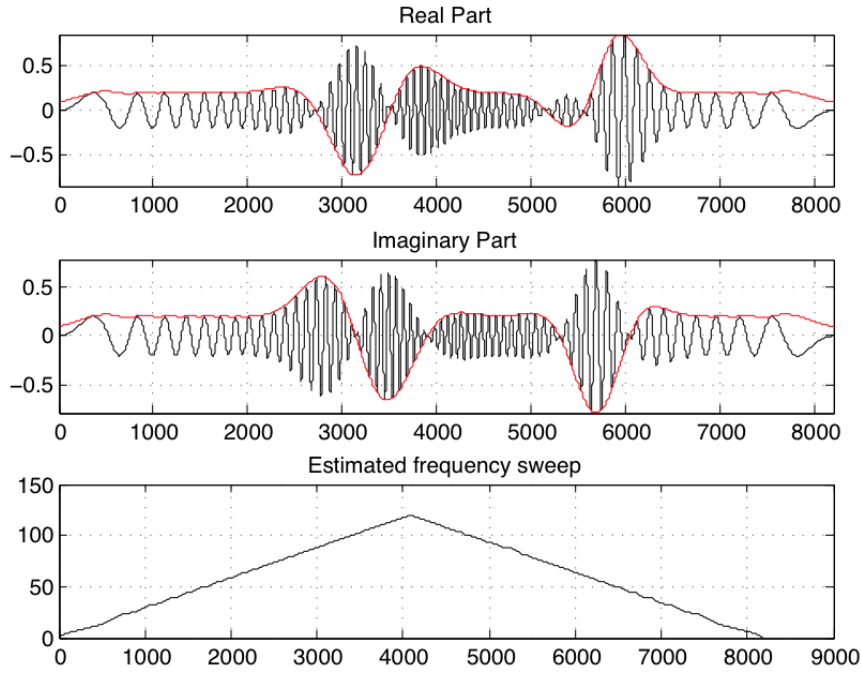


FIGURE 2.10 – Parties réelles et imaginaires du signal $z(t)$ (en noir) et de son enveloppe complexe (en rouge). Fréquence instantanée de $z(t)$ obtenue à partir de la dérivée de la phase instantanée du signal hyperanalytique $z_h(t)$.

détaillée dans [Kuipers 1999, p. 264]) que :

$$\frac{d\varphi_z(t)}{dt} = \arg \left(\bar{z}_h^{-1}(t) \frac{d\bar{z}_h(t)}{dt} \right)$$

Cette fréquence instantanée est en fait l’argument d’une fréquence angulaire. En complément de cette “fréquence”, et contrairement au cas classique du signal analytique, il faut considérer une direction de rotation, qui est représentée par $\eta_z(t)$. Cette direction indique comment la normale au plan osculateur, dans lequel le signal se déplace localement, évolue au cours du temps. La fréquence instantanée donne la vitesse de cette rotation.

Dans le cas du signal présenté sur la figure 2.9, nous présentons sur la figure 2.10 les parties réelles et imaginaires du signal (avec les parties réelles et imaginaires de l’amplitude instantanée), ainsi que la valeur de la fréquence instantanée obtenue par dérivation de la partie unitaire du signal hyperanalytique. On voit que l’on retrouve le chirp du signal complexe que l’on avait construit, c’est à dire la fréquence instantanée. Il est à noter que la description totale du comportement “géométrique” du signal $z(t)$ requiert la connaissance de l’axe de rotation.

Nous avons illustré l’utilité du signal hyperanalytique, défini via la TF Quaternionique, pour la description géométrique des signaux complexes non-circulaires. Dans le cas où $z(t)$ est circulaire, sa partie réelle est en quadrature avec sa partie imaginaire, et donc la connaissance d’une des deux composantes engendre la connaissance de la seconde. Quand ce n’est plus le cas, il est nécessaire de connaître le comportement non-stationnaire du signal $z(t)$

et cela peut être fait via le signal hyperanalytique. Ce dernier nous permet de connaître à tout instant les changements de directions dans le plan complexe via l'argument et les variations d'amplitude via l'enveloppe complexe.

Le signal hyperanalytique est intéressant dans le sens où il permet d'exhiber des caractéristiques géométriques d'un signal complexe via des manipulations "fréquentielles". C'est l'argument principal de l'utilisation des transformations de Fourier Quaternioniques (et plus généralement les transformées de Fourier géométriques). Le signal hyperanalytique souffre des mêmes limitations que le signal analytique en ce qui concerne le nombre de composantes présentes dans le signal : dès que le contenu spectral devient large-bande, l'interprétation en fréquence instantanée a atteint sa limite, et il faut se tourner vers des outils qui ne sont pas développés pour l'heure : les représentations temps-fréquences quaternioniques pour les signaux complexes non-circulaires. C'est un des axes naturels de poursuite de ces travaux. Nous reparlerons de ce point au chapitre 4.

2.4 Conclusion

Ce chapitre a présenté quelques applications des quaternions en traitement du signal. Le formalisme des quaternions permet de développer des traitements naturellement adaptés aux signaux à échantillons 3D ou 4D. Nous avons montré comment, par exemple, en traitement d'antenne, ils permettent d'extraire directement l'information de polarisation, et en particulier de faire la distinction entre ondes progrades et rétrogrades.

Maintenant au-delà de l'extension simple des algorithmes existants du cas complexe au cas quaternionique, les quaternions ont sans doute un autre rôle à jouer. Réécrire systématiquement tous les algorithmes connus pour les adapter aux signaux quaternioniques et les utiliser pour le traitement de signaux à échantillons de dimension 3 ou 4, écrits sous forme de quaternions, peut permettre quelques avancées, mais n'est pas vraiment stimulant à mon sens. La contribution la plus intéressante des quaternions, de mon point de vue, est dans la *géométrie* qu'ils apportent, en particulier *via* leurs représentations polaires et leurs phases. Le concept de signal hyperanalytique me paraît la contribution la plus complète des quaternions en traitement du signal car l'apport de la TF Quaternionique réside également dans l'alliance d'une analyse spectrale et géométrique d'un signal complexe. C'est sans doute la contribution de la TFQ en traitement du signal, et ses extensions aux dimensions supérieures via les algèbres de Clifford, qui me paraît être la plus prometteuse.

2.5 Publications annexées en lien avec ce chapitre

Les articles suivants sont inclus ici :

1. “MUSIC algorithm for vector-sensors array using biquaternions”, S. Miron, N. Le Bihan and J. Mars, IEEE Transactions on Signal Processing, Vol. 55, No. 9, pp. 4523 – 4532, 2007.
2. “The $\mathbb{H}$ -analytic signal”, N. Le Bihan and S.J. Sangwine, Eusipco, Lausanne, 2008.
3. “Quaternionic independent component analysis using hypercomplex nonlinearities”, N. Le Bihan and S. Buchholz, IMA, Cirencester, 2006.

MUSIC Algorithm for Vector-Sensors Array Using Biquaternions

Nicolas Le Bihan, Sebastian Miron, and Jérôme I. Mars

Abstract—In this paper, we use a biquaternion formalism to model vector-sensor signals carrying polarization information. This allows a concise and elegant way of handling signals with eight-dimensional (8-D) vector-valued samples. Using this model, we derive a biquaternionic version of the well-known array processing MUSIC algorithm, and we show its superiority to classically used *long-vector* approach. New results on biquaternion valued matrix spectral analysis are presented. Of particular interest for the biquaternion MUSIC (BQ-MUSIC) algorithm is the decomposition of the spectral matrix of the data into orthogonal subspaces. We propose an effective algorithm to compute such an orthogonal decomposition of the observation space via the eigenvalue decomposition (EVD) of a Hermitian biquaternionic matrix by means of a newly defined quantity, the *quaternion adjoint matrix*. The BQ-MUSIC estimator is derived and simulation results illustrate its performances compared with two other approaches in polarized antenna processing (LV-MUSIC and PSA-MUSIC). The proposed algorithm is shown to be superior in several aspects to the existing approaches. Compared with LV-MUSIC, the BQ-MUSIC algorithm is more robust to modelization errors and coherent noise while it can detect less sources. In comparison with PSA-MUSIC, our approach exhibits more accurate estimation of direction of arrival (DOA) for a small number of sources, while keeping the polarization information accessible.

Index Terms—Biquaternions and biquaternion-valued matrices, Biquaternion MUSIC (BQ-MUSIC), eigenvalue decomposition (EVD) of biquaternionic matrices, vector-sensor array processing.

I. INTRODUCTION

THE vector-sensors are now of common use in different applications such as electromagnetics, communications, seismic sensing, seismology, etc. These sensors record the components of the observed nonisotropic field and allow the recovery of polarization information. Depending on the application and the type of sensors, one can record two (two-component sensors) to six (three components of $\vec{E}$ and three components of $\vec{B}$ for electromagnetic wave fields) signals on a collocated sensor. The use of such sensors has proved its advantages in increasing the performances of classical algorithms (due to the redundancy of signals on the different components) and represents at the same time the only possibility to recover polarization information. There is a large number of studies on

the extensions of classical signal/array processing techniques to the vector-sensor case (see [1] and references therein). Further, high-resolution array processing algorithms were studied for the multicomponent case, mainly by Nehorai [2], [3], Wong and Zoltowski [1], [4]–[7] and Li ([8]–[11]), for different configurations and both for MUSIC- and ESPRIT-like algorithms. Furthermore, the performances of vector-sensor arrays were analyzed and quantified in [3] and [12]. In all these contributions considering arrays of vector-sensors, the *vector* dimension of the recorded signals was unfolded along the *distance* (related to the number of sensors/aperture of the array) dimension, resulting in the so-called “long-vector” approach. This way of processing data originated from vector-sensors has the main advantage of allowing, together with a rather complicated parametrization of the data, the use of well-known matrix algebra techniques over the real or the complex field. However, the “long-vector” approach has the drawback of destroying locally the vector-type of the signal because of the reorganization of the data into a large vector.

In this paper, we propose an alternative way to process signals from vector-sensor arrays. Instead of reorganizing data into long vectors, we introduce a hypercomplex model for multicomponent signals impinging on vector-sensors. This model is based on biquaternions (quaternions with complex coefficients) and allows the processing of multicomponent signals using linear algebra algorithms over the biquaternions. Consequently, the derivation of high-resolution techniques for vector-sensors array is possible. We illustrate our approach by deriving a Biquaternion MUSIC (BQ-MUSIC)-like algorithm for this type of arrays. The use of biquaternions allows us to skip the parametrization step used in long-vector techniques [3] as it intrinsically includes the vector dimension in the process. The authors previously proposed the use of quaternions to process vector-sensor signals [13], [14]. In [13], a quaternion model for three-components vector-sensor signals was used and a subspace method was derived in the time domain, allowing denoising of polarized waves. In [14], only two-component vector-sensors arrays were considered. A quaternion modelization of the output signals was used and a MUSIC algorithm derived for direction-of-arrival (DOA) and polarization parameters estimation. The proposed technique in this paper is a generalization of the one presented in [14] to the case of three-component vector-sensor arrays. The use of biquaternions for signal modelization leads to new problems, such as the diagonalization of the biquaternionic sample covariance matrix. An original technique is proposed for this task.

Since biquaternions have not been widely studied in literature, there is a lack of known results on matrices with biquater-

Manuscript received February 6, 2006; revised December 25, 2006. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. A. Rahim Leyman.

N. Le Bihan and J. I. Mars are with the GIPSA-Lab, INPG, CNRS, Département Signal-Images, Domaine Universitaire, 38402 Saint Martin d'Hères Cedex, France (e-mail: nicolas.le-bihan@lis.inpg.fr; jerome.mars@lis.inpg.fr).

S. Miron is with the Centre de Recherche en Automatique de Nancy (CRAN), Faculté des Sciences et Techniques, 54506 Vandœuvre-les-nancy Cedex, France (e-mail: sebastien.miron@cran.uhp-nancy.fr).

Digital Object Identifier 10.1109/TSP.2007.896067

nion coefficients. We present here some novel results about such matrices with emphasis on eigenvalue decomposition. We provide a way to compute the eigenvalue decomposition of a Hermitian biquaternion valued matrix and show its application in a biquaternion MUSIC algorithm. The proposed biquaternion MUSIC algorithm is then compared with the classical long-vector MUSIC and to the Polarization Smoothing Algorithm (PSA)-MUSIC algorithm [15]; its superiority in computational/robustness/resolution issues is demonstrated on numerical examples.

The biquaternion approach developed here is part of a new way of considering vector-sensor signals, whose global underlying philosophy consists in considering that these signals evolve on extended algebraic structures, rather than trying to make the signal fit the already existing algorithms/concepts.¹

The paper is organized as follows. In Section II, we introduce biquaternions and their basic properties. Then in Section III, we present a detailed study of biquaternion valued matrices with particular attention to the eigenvalue decomposition problem. This decomposition is introduced, and the link with orthogonal decomposition and rank properties are illustrated. In Section IV, the biquaternion model for polarized waves recorded on three-component vector-sensor arrays is introduced. This model, together with the eigenvalue decomposition (EVD) allows the definition of a BQ-MUSIC algorithm described in Section V. Simulation results and comparisons with the long-vector approach and PSA-MUSIC are enlightened in Section VI. Concluding remarks about this work are presented in Section VII.

II. BIQUATERNIONS

Biquaternions, also known as “complexified quaternions,” are an eight-dimensional (8-D) algebra and consist of quaternion numbers with complex coefficients. They were discovered by Hamilton in 1853 [17]. While Hamilton’s (real) quaternions [18] are noted $\mathbb{H}$, the set of complex quaternions is noted $\mathbb{H}_{\mathbb{C}}$ [19].

Definition 1: A complexified quaternion $q \in \mathbb{H}_{\mathbb{C}}$ is given by

$$q = q_0 + q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k} \quad (1)$$

where $q_0, q_1, q_2, q_3 \in \mathbb{C}_I$ and with elements of $\mathbb{C}_I$ defined as

$$z = \Re(z) + I\Im(z) \Leftrightarrow z \in \mathbb{C}_I \quad (2)$$

with $I = \sqrt{-1}$ and $\Re(z), \Im(z) \in \mathbb{R}$. The following *standard* relations between imaginary quaternion units hold:

$$\begin{aligned} \mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = \mathbf{i}\mathbf{j}\mathbf{k} &= -1 \\ \mathbf{i}\mathbf{j} &= -\mathbf{j}\mathbf{i} = \mathbf{k} \\ \mathbf{k}\mathbf{i} &= -\mathbf{i}\mathbf{k} = \mathbf{j} \\ \mathbf{j}\mathbf{k} &= -\mathbf{k}\mathbf{j} = \mathbf{i} \end{aligned} \quad (3)$$

with, in addition, the following relations between complex imaginary unit I and quaternion imaginary units:

$$\mathbf{i}I = I\mathbf{i}; \mathbf{j}I = I\mathbf{j}; \mathbf{k}I = I\mathbf{k} \quad (4)$$

¹This approach has to be put in parallel with the one developed by Manton [16], who developed the processing of signals evolving on manifolds.

meaning that any complex coefficient commutes with any quaternion imaginary unit.

Thus, biquaternions form an 8-D vector space over $\mathbb{R}$ with basis:

$$\{1, \mathbf{i}, \mathbf{j}, \mathbf{k}, I, \mathbf{i}I, \mathbf{j}I, \mathbf{k}I\}. \quad (5)$$

Biquaternions form an *associative algebra* but *not a normed division algebra*. The only 8-D normed division algebra are the ones isomorphic to Cayley’s octonions (this is known as generalized Frobenius and Hurwitz theorems, see [20] for details).

Biquaternions are isomorphic to Clifford algebra $\mathcal{C}_3$ (the Clifford algebra built over $\mathbb{R}^3$ with basis $\{e_1, e_2, e_3\}$ and such that $e_m e_n + e_n e_m = 2\delta_{nm}$), with identifications, as follows:

$$\begin{cases} \mathbf{i}I \leftrightarrow -e_3 \\ \mathbf{j}I \leftrightarrow -e_1 \\ \mathbf{k}I \leftrightarrow e_2 \end{cases} \quad \text{and} \quad \begin{cases} \mathbf{i} \leftrightarrow e_1 e_2 \\ \mathbf{j} \leftrightarrow e_2 e_3 \\ \mathbf{k} \leftrightarrow e_1 e_3 \\ I \leftrightarrow e_1 e_2 e_3 \end{cases} \quad (6)$$

where $e_n e_m$ are bivectors and $e_1 e_2 e_3$ is a pseudoscalar [21], [22].

Next, we present a nonexhaustive list of properties for biquaternions. The interested reader will find more material in [19]. Note that (real) Hamilton’s quaternions are a special case of biquaternions. As in the case of quaternions, any biquaternion q can be seen as the sum of a scalar and a vector part, both with complex valued coefficients, as follows:

$$q = \mathcal{S}(q) + \mathcal{V}(q) \quad (7)$$

where

$$\begin{cases} \mathcal{S}(q) = q_0 \\ \mathcal{V}(q) = q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k}. \end{cases} \quad (8)$$

At the same time, q can be seen as the sum of a real and an imaginary part, both being quaternion valued, as follows:

$$q = \Re(q) + I\Im(q) \quad (9)$$

where

$$\begin{cases} \Re(q) = \Re(q_0) + \Re(q_1) + \Re(q_2) + \Re(q_3) \\ \Im(q) = \Im(q_0) + \Im(q_1) + \Im(q_2) + \Im(q_3). \end{cases} \quad (10)$$

This notation of a biquaternion can be seen as an equivalent of the Cayley–Dickson notation for real quaternions [20], and it will be useful in the study of biquaternion valued matrices. Note that a biquaternion with zero scalar part ($\mathcal{S}(q) = 0$) is called *pure*.

Some known properties of complex and quaternions numbers, such as the multiplication and the addition, extend naturally to biquaternions. For some others, the extension is not trivial.

Definition 2: There exist three different conjugations over $\mathbb{H}_{\mathbb{C}}$. Thus, given a complex quaternion q , it is possible to define its conjugations, as follows:

- $\mathbb{C}$ -conjugate: $q^* = q_0^* + q_1^* \mathbf{i} + q_2^* \mathbf{j} + q_3^* \mathbf{k} = \Re(q) - I\Im(q);$
- $\mathbb{H}$ -conjugate: $q^* = q_0 - q_1 \mathbf{i} - q_2 \mathbf{j} - q_3 \mathbf{k} = \mathcal{S}(q) - \mathcal{V}(q);$
- (Total) $\mathbb{H}_{\mathbb{C}}$ -conjugate: $\bar{q} = (q^*)^* = q_0^* - q_1^* \mathbf{i} - q_2^* \mathbf{j} - q_3^* \mathbf{k}.$

These definitions induce different possible definitions for norms. We mention here a norm and a pseudonorm.

Definition 3: The norm of a biquaternion $q \in \mathbb{H}_{\mathbb{C}}$, noted $|q|$, is given by

$$|q| = \sqrt{|q_0|^2 + |q_1|^2 + |q_2|^2 + |q_3|^2}. \quad (11)$$

Note that $|q| \geq 0 \forall q \in \mathbb{H}_{\mathbb{C}}$, and $|q| = 0 \Rightarrow q = 0$; the biquaternions are not a normed algebra under this norm, so in general $|qp| \neq |q||p|$ for $q, p \in \mathbb{H}_{\mathbb{C}}$.

It is possible to define a pseudo-norm satisfying the property that the pseudo-norm of a product of biquaternions is equal to the product of the pseudo-norms of the individuals.

Definition 4: The pseudo-norm of a biquaternion $q \in \mathbb{H}_{\mathbb{C}}$, noted $|q|_p$, is given by

$$|q|_p = q_0^2 + q_1^2 + q_2^2 + q_3^2 \quad (12)$$

and it satisfies the following equality: $|rq|_p = |r|_p|q|_p$ for $r, q \in \mathbb{H}_{\mathbb{C}}$. It has the drawback of being complex valued in general. This involves that the pseudo-norm of a nonzero biquaternion can vanish. For example, for the biquaternion $q = e^{I(\pi/4)} + e^{-I(\pi/4)}\mathbf{i} + e^{I(\pi/4)}\mathbf{j} + e^{-I(\pi/4)}\mathbf{k}$, its norm is $|q| = 2$ while its pseudo-norm is $|q|_p = 0$. This problem forbids a systematic use of this pseudo-norm in biquaternion valued signal processing for obvious reasons (problems in estimating the magnitude or the energy of a signal for example). We also give the following property that will be useful in the sequel.

Property 1: Any complex number $z = z_1 + z_2I$ with $z_1, z_2 \in \mathbb{R}$, (i.e., $z \in \mathbb{C}_I$) commutes with any biquaternion $q \in \mathbb{H}_{\mathbb{C}}$ defined as in (1), as follows:

$$qz = zq. \quad (13)$$

The proof is straightforward from the multiplication rules in (3) and (4). We now introduce some material on matrices with biquaternion valued coefficients and on their decomposition.

III. MATRICES WITH BIQUATERNION COEFFICIENTS

In this section, we present definitions and properties of biquaternion valued matrices. The study of these matrices was not paid much attention to in literature. In [23], Tian proved the existence of the eigenvalues and the eigenvectors for biquaternion matrices as well as a few other properties. We present in this section the definitions necessary for our purpose and we concentrate mainly on Hermitian biquaternion matrices as they will be of interest in Section V.

A. Vectors and Matrices of Biquaternions

Biquaternions have mainly been used in formulations of electromagnetics [24] and special relativity [19], [25]. However, in such studies, the case of matrices with biquaternions coefficients has not been considered. We present here some results of a study on such matrices with particular attention to the eigendecomposition of Hermitian biquaternion matrices.

1) Biquaternion Valued Vectors: A biquaternion valued vector is an element of $\mathbb{H}_{\mathbb{C}}^N$. Equipped with the classical addition of vectors and the multiplication with a biquaternionic

scalar, $\mathbb{H}_{\mathbb{C}}^N$ is a $\mathbb{H}_{\mathbb{C}}$ -module (vector space over the ring $\mathbb{H}_{\mathbb{C}}$). The scalar product of two biquaternion valued vectors $\mathbf{a}, \mathbf{b} \in \mathbb{H}_{\mathbb{C}}^N$ is defined the following way:

$$\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{H}_{\mathbb{C}}} = \mathbf{a}^\dagger \mathbf{b} = \sum_{n=1}^N \bar{a}_n b_n \quad (14)$$

where $\dagger$ stands for total conjugation-transposition. With this definition, two biquaternion valued vectors $\mathbf{a}, \mathbf{b} \in \mathbb{H}_{\mathbb{C}}^N$ are said *orthogonal* iff

$$\langle \mathbf{a}, \mathbf{b} \rangle_{\mathbb{H}_{\mathbb{C}}} = 0. \quad (15)$$

Based on the scalar product definition, the norm of a biquaternion valued vector $\mathbf{b} \in \mathbb{H}_{\mathbb{C}}^N$ is given by

$$\|\mathbf{b}\| = \sqrt{S(\langle \mathbf{b}, \mathbf{b} \rangle_{\mathbb{H}_{\mathbb{C}}})} \quad (16)$$

where $S(\cdot)$ is the scalar part defined in (8). We now turn to matrices with biquaternion coefficients.

2) Matrices of Biquaternions: A biquaternion valued matrix with M rows and N columns is an element of $\mathbb{H}_{\mathbb{C}}^{M \times N}$. Given a biquaternion valued matrix $\mathbf{B} = (b_{st}) \in \mathbb{H}_{\mathbb{C}}^{M \times N}$, one can define the following [23], [26]:

- the *dual* matrix of $\mathbf{B}$: $\mathbf{B}^\diamond = (b_{ts}^*) \in \mathbb{H}_{\mathbb{C}}^{N \times M}$;
- the transpose-conjugate of $\mathbf{B}$: $\mathbf{B}^\dagger = (\bar{b}_{ts}) \in \mathbb{H}_{\mathbb{C}}^{N \times M}$.

A matrix $\mathbf{B} \in \mathbb{H}_{\mathbb{C}}^{N \times N}$ is then called *Hermitian* if $\mathbf{B} = \mathbf{B}^\dagger$ and unitary if $\mathbf{B}\mathbf{B}^\dagger = \mathbf{B}^\dagger\mathbf{B} = \mathbf{I}_N$. Invertibility and the definition of the inverse of a biquaternion valued matrix are defined similarly to the real or complex case. Given two matrices $\mathbf{A} \in \mathbb{H}_{\mathbb{C}}^{M \times N}$ and $\mathbf{B} \in \mathbb{H}_{\mathbb{C}}^{N \times P}$, then the following equalities stand [23], [26]:

- 1) $(\mathbf{A}^\diamond)^\diamond = \mathbf{A}$, $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$;
- 2) $(\mathbf{AB})^\diamond = \mathbf{B}^\diamond \mathbf{A}^\diamond$, $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$;
- 3) $(\mathbf{AB})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1}$, if $\mathbf{A}$ and $\mathbf{B}$ are invertible;
- 4) $(\mathbf{A}^\diamond)^{-1} = (\mathbf{A}^{-1})^\diamond$, $(\mathbf{A}^\dagger)^{-1} = (\mathbf{A}^{-1})^\dagger$ if $\mathbf{A}$ is invertible.

These properties will be of use in the sequel.

3) Quaternionic Adjoint Matrix of a Biquaternion Valued Matrix: In order to compute the eigenvalue decomposition of a biquaternion valued matrix, we now introduce the *quaternionic adjoint matrix* of a given biquaternionic matrix. A similar technique was employed by Lee and Brenner [27] in the study of quaternion matrices. The use of such an “equivalent” quaternion matrix is possible because any Clifford algebra is isomorphic to a complex matrix algebra [28]. Consequently, any biquaternion (and by extension any matrix of biquaternions) is isomorphic to a complex matrix (by extension to a tensor product of complex matrices). (For more details on isomorphisms between complex matrices algebras and Clifford algebras, see [28, Ch. 11].)

Given a biquaternion valued matrix $\mathbf{B} \in \mathbb{H}_{\mathbb{C}}^{M \times N}$ written as $\mathbf{B}_1 + I\mathbf{B}_2$, where $\mathbf{B}_1, \mathbf{B}_2 \in \mathbb{H}^{M \times N}$, then its *quaternionic adjoint matrix*, noted $\gamma_{\mathbf{B}}$, takes values in $\mathbb{H}^{2M \times 2N}$ and has the following expression:

$$\gamma_{\mathbf{B}} = \begin{pmatrix} \mathbf{B}_1 & \mathbf{B}_2 \\ -\mathbf{B}_2 & \mathbf{B}_1 \end{pmatrix}. \quad (17)$$

Consider now the complex matrix $\Psi_M \in \mathbb{C}_I^{M \times 2M}$ defined as

$$\Psi_M = (\mathbf{I}_M, -I\mathbf{I}_M) \quad (18)$$

where $\mathbf{I}_M$ is the identity matrix of dimension $M \times M$. It is straightforward that the following equality holds:

$$\mathbf{B} = \frac{1}{2} \Psi_M \gamma_B \Psi_N^\dagger. \quad (19)$$

It is also important to notice the two following properties of the matrix Ψ_M that will be of use in the forthcoming calculations:

$$\Psi_M \Psi_M^\dagger = 2\mathbf{I}_M \quad (20)$$

$$\gamma_B \Psi_N^\dagger \Psi_N = \Psi_M^\dagger \Psi_M \gamma_B. \quad (21)$$

Property (20) can be demonstrated by direct calculation while, for the equality (21), it can be proved by multiplication on the left by Ψ_M and on the right by $\Psi_N^\dagger$. Property (20) is then used to fulfill the demonstration.

Lemma 1: The *quaternion adjoint matrix* of a Hermitian biquaternion matrix is also Hermitian.

Proof: Consider a biquaternion valued Hermitian matrix $\mathbf{B} \in \mathbb{H}_\mathbb{C}^{N \times N}$

$$\mathbf{B} = \mathbf{B}^\dagger \quad (22)$$

and its quaternion adjoint matrix $\gamma_B \in \mathbb{H}^{2N \times 2N}$. Substituting (19) in (22), one can write

$$\Psi_N \gamma_B \Psi_N^\dagger = \left(\Psi_N \gamma_B \Psi_N^\dagger \right)^\dagger. \quad (23)$$

Using the fact that for biquaternion valued matrices $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$, (23) becomes

$$\Psi_N \gamma_B \Psi_N^\dagger = \Psi_N \gamma_B^\dagger \Psi_N^\dagger \quad (24)$$

leading to

$$\gamma_B = \gamma_B^\dagger. \quad (25)$$

Thus γ_B is Hermitian. ■

In a similar way, using definition (17) and properties (20) and (21), it is possible to prove that the *quaternion adjoint matrix* conserves the unitary property of a biquaternion valued matrix.

Next, we make use of the *quaternion adjoint matrix* for the computation of the eigenvalue decomposition of a biquaternion valued matrix.

B. Eigendecomposition of a Biquaternion Valued Matrix

As in the quaternion case [29], the noncommutativity of biquaternion multiplication leads to two possible eigenvalues, namely the *left* and the *right* eigenvalues. However, in the sequel, we will only consider right eigenvalues. This choice is motivated by the link between biquaternionic right eigenvalues and quaternionic eigenvalues of the *quaternion adjoint matrix*. In the quaternion case, the theory of left eigenvalues is still not complete [30], and this motivates our choice to consider only right eigenvalues, which have been well understood for several years now [31].

After a definition of (right) EVD for biquaternion valued matrices, we present several lemmas and corollaries that are helpful

for effective computation of the eigenvalues of a biquaternion matrix.

Definition 5: Given a biquaternion valued matrix $\mathbf{B} \in \mathbb{H}_\mathbb{C}^{N \times N}$, then its eigenvalue decomposition is given by

$$\mathbf{B} = \mathbf{U} \mathbf{D} \mathbf{U}^\dagger \quad (26)$$

where $\mathbf{U} \in \mathbb{H}_\mathbb{C}^{N \times 2N}$ is a biquaternion valued matrix containing the eigenvectors of $\mathbf{B}$ and $\mathbf{D} \in \mathbb{H}_\mathbb{C}^{2N \times 2N}$ is a diagonal matrix containing eigenvalues of $\mathbf{B}$ on its diagonal.

Next, we present some results showing how the eigenvalues of a biquaternion matrix can be obtained from the eigenvalue decomposition of its *quaternion adjoint matrix*. First, the (right) eigenvectors of a square biquaternionic matrix $\mathbf{B} \in \mathbb{H}_\mathbb{C}^{N \times N}$ can be obtained using the following lemma.

Lemma 2: Given a square biquaternionic matrix $\mathbf{B} \in \mathbb{H}_\mathbb{C}^{N \times N}$, then if $\mathbf{u}_q \in \mathbb{H}^{2N}$ is a right eigenvector of its *quaternion adjoint matrix* γ_B , then $\mathbf{u}_b \in \mathbb{H}_\mathbb{C}^N$, defined as

$$\mathbf{u}_b = \Psi_N \mathbf{u}_q \quad (27)$$

is a right eigenvector of $\mathbf{B}$.

Proof: Assume $\mathbf{u}_q$ is a right eigenvector of γ_B , then the following equality holds:

$$\gamma_B \mathbf{u}_q = \mathbf{u}_q \lambda. \quad (28)$$

Using (19) and (27), one can write

$$\begin{aligned} \mathbf{B} \mathbf{u}_b &= \frac{1}{2} \Psi_N \gamma_B \Psi_N^\dagger \Psi_N \mathbf{u}_q \\ &= \frac{1}{2} \Psi_N 2\mathbf{I}_N \gamma_B \mathbf{u}_q \\ &= \Psi_N \gamma_B \mathbf{u}_q. \end{aligned} \quad (29)$$

Substituting (28) in (29) results in

$$\mathbf{B} \mathbf{u}_b = \Psi_N \mathbf{u}_q \lambda = \mathbf{u}_b \lambda \quad (30)$$

so $\mathbf{u}_b = \Psi_N \mathbf{u}_q$ is a right eigenvalue of $\mathbf{B}$. ■

As a result, the eigenvalue decomposition of a biquaternion valued matrix can be obtained from the eigendecomposition of a double size quaternion valued matrix, the *quaternion adjoint matrix*. As a consequence, it is possible to use algorithms developed for quaternion valued matrices for this calculation [13]. The following corollary states this fact.

Corollary 1: Consider a biquaternion valued matrix $\mathbf{B} \in \mathbb{H}_\mathbb{C}^{N \times N}$ and assume that its *quaternion adjoint matrix* γ_B has the following EVD: $\gamma_B = \mathbf{U} \mathbf{D} \mathbf{U}^\dagger$, where $\mathbf{U} \in \mathbb{H}^{2N \times 2N}$ and $\mathbf{D} \in \mathbb{C}_j^{2N \times 2N}$ ($\mathbb{C}_j$ is a subset of $\mathbb{H}$, isomorphic to $\mathbb{C}$, for which the coefficients of the imaginary units $\mathbf{i}$ and $\mathbf{k}$ are null). The eigendecomposition of $\mathbf{B}$ is then given by

$$\mathbf{B} = \mathbf{U}_b \mathbf{D} \mathbf{U}_b^\dagger \quad (31)$$

where $\mathbf{U}_b = (1/\sqrt{2}) \Psi_N \mathbf{U} \in \mathbb{H}_\mathbb{C}^{N \times 2N}$ and $\mathbf{D}$ is the diagonal matrix with the eigenvalues of γ_B as diagonal elements.

Proof: Assuming the EVD of γ_B can be written as

$$\gamma_B = \mathbf{U} \mathbf{D} \mathbf{U}^\dagger \quad (32)$$

with $\mathbf{U} \in \mathbb{H}^{2N \times 2N}$ and $\mathbf{D} \in \mathbb{C}_j^{2N \times 2N}$, then, by replacing (32) in (19), one gets

$$\mathbf{B} = \frac{1}{2} \Psi_N \mathbf{U} \mathbf{D} \mathbf{U}^\dagger \Psi_N^\dagger = \frac{1}{2} \Psi_N \mathbf{U} \mathbf{D} (\Psi_N \mathbf{U})^\dagger. \quad (33)$$

Knowing that $(1/\sqrt{2})\Psi_N \mathbf{U} = \mathbf{U}_b \in \mathbb{H}_\mathbb{C}^{N \times 2N}$, then it is possible to write

$$\mathbf{B} = \mathbf{U}_b \mathbf{D} \mathbf{U}_b^\dagger \quad (34)$$

where $\mathbf{D}$ is a diagonal matrix, and $\mathbf{U}_b \in \mathbb{H}_\mathbb{C}^{N \times 2N}$ contains the eigenvectors of $\mathbf{B}$ on its columns, as previously shown. ■

The eigenvalues of $\gamma_\mathbf{B}$ are also the eigenvalues of $\mathbf{B}$. In the general case, the eigenvalues of a biquaternion valued matrix are quaternion valued. However, the possible values taken by the eigenvalues are either in $\mathbb{C}_i$, $\mathbb{C}_j$, or $\mathbb{C}_k$, which are, in the biquaternion case, degenerate quaternions.² This means that in the biquaternion case, the eigenvalues are generally quaternions with two or three null components.

Nevertheless, it is important to notice that the eigenvalues of the *quaternion adjoint matrix* do not appear in conjugate pairs along the diagonal of $\mathbf{D}$, as opposed to the quaternion case where this happens for the eigenvalues of the *complex adjoint matrix* [29], [31]. As a consequence, it is necessary to consider all the $2N$ eigenvectors and their associated eigenvalues to rebuild a whole biquaternionic matrix $\mathbf{B} \in \mathbb{H}_\mathbb{C}^{N \times N}$.

Note that in the case of symmetric octonion³ valued matrices, it has been demonstrated that a 3×3 matrix has six independent eigenvalues [32].

An interpretation to this large number of eigenvalues can be given using isomorphisms. It has been shown that the algebra of complexified quaternions is identical to that generated by Pauli matrices (elements of $\mathbb{C}^{2 \times 2}$) [19], [28]. The space of biquaternion valued matrices $\mathbb{H}_\mathbb{C}^{N \times N}$ is then isomorphic to $\mathbb{C}^{2 \times 2} \otimes \mathbb{R}^{N \times N}$, where $\otimes$ denotes the tensor product of two vector-spaces. As a consequence the dimension of the column vector space of $\mathbb{H}_\mathbb{C}^{N \times N}$ is given by

$$\dim(\mathbb{H}_\mathbb{C}^{N \times N}) = \dim(\mathbb{C}^{2 \times 2}) \dim(\mathbb{R}^{N \times N}) = 2N. \quad (35)$$

1) EVD of a Hermitian Biquaternionic Matrix: The high-resolution vector-sensor array processing algorithm presented in Section V is based on the decomposition of the covariance matrix of the observations into orthogonal subspaces, using a biquaternion model. This covariance matrix is biquaternionic

²Note that the biquaternion case is different from the quaternion case; as for the latter, the eigenvalues of quaternion matrices are isomorphic to complex eigenvalues.

³Octonions are the only 8-D normed division algebra [20]. They form a nonassociative and noncommutative algebra.

Hermitian. Consequently, we now pay attention to the EVD of a Hermitian biquaternion valued matrix.

A matrix $\mathbf{B} \in \mathbb{H}_\mathbb{C}^{N \times N}$ is called Hermitian if $\mathbf{B} = \mathbf{B}^\dagger$. We have already demonstrated (Lemma 1) that the *quaternion adjoint matrix* $\gamma_\mathbf{B} \in \mathbb{H}^{2N \times 2N}$ of a Hermitian biquaternionic matrix is also Hermitian. Thus, $\gamma_\mathbf{B} = \gamma_\mathbf{B}^\dagger$.

As the eigenvalues of $\mathbf{B}$ are the same as the ones of $\gamma_\mathbf{B}$, and due to the fact that the eigenvalues of a Hermitian quaternion valued matrix are real valued [29], then the eigenvalues of a Hermitian biquaternion valued matrix are real as well. It is easy to demonstrate (see [33] for the quaternion case) that for Hermitian matrices, the right and left eigenvalues (and associated eigenvectors) are the same. We now prove that an important lemma, well known for the real, complex, and quaternionic case, extends to biquaternions. This is fundamental for the construction of any algorithm based on orthogonal decomposition of the observed data.

Lemma 3: Given a Hermitian biquaternion valued matrix $\mathbf{B}$, then any two of its eigenvectors corresponding to two different eigenvalues are orthogonal.

Proof: Consider two eigenvalues of $\mathbf{B} \in \mathbb{H}_\mathbb{C}^{N \times N}$, $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 \neq \lambda_2$ and their associated eigenvectors $\mathbf{u}_1, \mathbf{u}_2 \in \mathbb{H}_\mathbb{C}^N$. Then, one can write

$$\begin{aligned} \lambda_1 (\mathbf{u}_1^\dagger \mathbf{u}_2) &= (\mathbf{u}_1 \lambda_1)^\dagger \mathbf{u}_2 = (\mathbf{B} \mathbf{u}_1)^\dagger \mathbf{u}_2 = \mathbf{u}_1^\dagger \mathbf{B}^\dagger \mathbf{u}_2 \\ &= \mathbf{u}_1^\dagger (\mathbf{B} \mathbf{u}_2) = \mathbf{u}_1^\dagger (\lambda_2 \mathbf{u}_2) = (\mathbf{u}_1^\dagger \mathbf{u}_2) \lambda_2. \end{aligned} \quad (36)$$

As $\lambda_1, \lambda_2 \in \mathbb{R}$ and $\lambda_1 \neq \lambda_2$, then equality (36) involves $\mathbf{u}_1^\dagger \mathbf{u}_2 = 0$, which means that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal. ■

The following numerical example illustrates the link between the rank of a biquaternion valued matrix and its eigenvalue decomposition. Consider a biquaternion valued vector of dimension 3, $\mathbf{s} \in \mathbb{H}_\mathbb{C}^3$, given as (37), shown at the bottom of the page.

Then, the following matrix is Hermitian:

$$\mathbf{S} = \mathbf{s} \mathbf{s}^\dagger \in \mathbb{H}_\mathbb{C}^{3 \times 3}. \quad (38)$$

Using the classical definition for the rank of a matrix, by construction $\mathbf{S}$ has a rank equal to 1. The eigendecomposition of $\mathbf{S}$ gives two different non-null real eigenvalues: $\lambda_1 = 5.918$ and $\lambda_2 = 4.166$. The remaining four other eigenvalues are null. The eigenvectors associated to the non-null eigenvalues are $\mathbf{u}_1 \in \mathbb{H}_\mathbb{C}^N$ and $\mathbf{u}_2 \in \mathbb{H}_\mathbb{C}^N$ and have the numerical values of (39) and (40), shown at the bottom of the next page.

It can be directly verified by calculation that $\mathbf{u}_1$ and $\mathbf{u}_2$ are orthogonal. Thus, the eigendecomposition of $\mathbf{S}$ can be written as

$$\mathbf{S} = \sum_{k=1}^2 \lambda_k \mathbf{u}_k \mathbf{u}_k^\dagger \quad (41)$$

$$\mathbf{s} = \begin{pmatrix} 0.950 + 0.486I & 0.456 + 0.444I & 0.921 + 0.405I & 0.410 + 0.352I \\ 0.231 + 0.891I & +i & 0.018 + 0.615I & +j \\ 0.606 + 0.762I & 0.821 + 0.791I & 0.176 + 0.916I & 0.057 + 0.009I \end{pmatrix}. \quad (37)$$

Now, comparing (41) and (38), one remarks that in order to recover the information contained in $\mathbf{s}$, it is necessary to consider two eigenvalues and their associated biquaternion eigenvectors. This result will be used in the vector-sensor HR array processing algorithm derived in Section V.

We saw that in order to be consistent with the real, complex, and quaternion valued matrix theory, the classical definition of the rank of a matrix needs to be revisited. Thus, the following definition stands for rank definition for biquaternion valued matrices.

Definition 6: The rank of a biquaternionic matrix $\mathbf{B} \in \mathbb{H}_{\mathbb{C}}^{N \times M}$ is given by

$$\text{rank}(\mathbf{B}) = \frac{1}{2} \text{rank}(\gamma_{\mathbf{B}}). \quad (42)$$

Now, with the presented material on spectral decomposition of biquaternion matrices and the matrix algebra tools over $\mathbb{H}_{\mathbb{C}}$, we are ready for developing our biquaternionic model and the algorithm for vector-sensor array processing purpose.

IV. POLARIZED SIGNAL MODEL USING BIQUATERNIONS

Following the approach proposed in [14] for the processing of signals recorded on two-component vector-sensors, we introduce a biquaternionic model for polarized signals recorded on three-component vector-sensors.

A. Three-Component Vector-Sensor Signals

Consider a three-component vector-sensor, recording the three orthogonal components of an incident vector wave field, yielding the output signals $s_{c_1}(t)$, $s_{c_2}(t)$, and $s_{c_3}(t)$. The three components of the vector-sensor define an orthogonal basis in the Euclidean 3-D space. If $(O, \vec{x}, \vec{y}, \vec{z})$ is the orthonormal basis associated to the vector-sensor, the vector product relations between the unit vectors $\vec{x}, \vec{y}, \vec{z}$ fit perfectly the relationships between the quaternionic units $\mathbf{i}, \mathbf{j}, \mathbf{k}$ [see (3)]. Thus, the idea of using quaternions/biquaternions to model the signals recorded on the three components of a vector-sensor comes naturally.

The associated three-components pure quaternion valued signal $s(t)$ is then given by

$$s(t) = s_{c_1}(t)\mathbf{i} + s_{c_2}(t)\mathbf{j} + s_{c_3}(t)\mathbf{k}. \quad (43)$$

Defining the Fourier transform of $\mathbf{s}(t)$ as a triplet of complex Fourier transforms applied separately on each of the three components, one gets

$$s(\nu) = s_{c_1}(\nu)\mathbf{i} + s_{c_2}(\nu)\mathbf{j} + s_{c_3}(\nu)\mathbf{k} \quad (44)$$

where $s_{c_\alpha}(\nu) = \text{TF}[s_{c_\alpha}(t)]$, with $\alpha = 1, 2, 3$ and with the Fourier transform taking values in $\mathbb{C}_I$. Using the modulus-phase representation, (44) can be rewritten as

$$s(\nu) = \eta_1(\nu)e^{I\chi_1(\nu)}\mathbf{i} + \eta_2(\nu)e^{I\chi_2(\nu)}\mathbf{j} + \eta_3(\nu)e^{I\chi_3(\nu)}\mathbf{k} \quad (45)$$

where $\eta_1, \eta_2, \eta_3 \in \mathbb{R}$ are the amplitudes and $\chi_1, \chi_2, \chi_3 \in \mathbb{R}$ are the phases of the signals recorded on the three components. In the following, the frequency argument ν is omitted for clarity as the proposed algorithm is derived for narrowband signals, or it is applied at different frequencies independently. Considering the first component as reference, one can rewrite the biquaternion signal as the product between a pure biquaternion containing the relative amplitude ratios and the phase shifts of the second and the third components with respect to the first component, and a complex number representing the absolute amplitude and phase of the signal on the first component, as follows:

$$s = p(\rho_1, \varphi_1, \rho_2, \varphi_2)\eta_1 e^{I\chi_1}. \quad (46)$$

The expression for $p(\rho_1, \varphi_1, \rho_2, \varphi_2)$ is given by

$$p(\rho_1, \varphi_1, \rho_2, \varphi_2) = \mathbf{i} + \rho_1 e^{I\varphi_1}\mathbf{j} + \rho_2 e^{I\varphi_2}\mathbf{k} \quad (47)$$

with $\rho_1 = \eta_2/\eta_1$, $\rho_2 = \eta_3/\eta_1$ and $\varphi_1 = \chi_2 - \chi_1$, $\varphi_2 = \chi_3 - \chi_1$. In this model, $p(\rho_1, \varphi_1, \rho_2, \varphi_2)$ contains the polarization information of the signal, if we consider the first component as reference.

B. Polarized Plane Waves

Now, given a set of N_x equally spaced three-component vector-sensors, recording the contributions of L polarized plane waves, using the biquaternion model, the recorded signal $\mathbf{x} \in \mathbb{H}_{\mathbb{C}}^{N_x}$ is given by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{N_x} \end{bmatrix} = \sum_{l=1}^L p_l(\rho_{1l}, \varphi_{1l}, \rho_{2l}, \varphi_{2l}) \mathbf{a}(\theta_l) s_l + \mathbf{n} \quad (48)$$

where $p_l(\rho_{1l}, \varphi_{1l}, \rho_{2l}, \varphi_{2l})$ is the biquaternion valued polarization coefficient of the l^{th} wave, containing its polarization parameters, $\mathbf{a}(\theta_l)$ is the propagation vector of the l^{th} wave on the array and is given (assuming plane waves contributions only) by

$$\mathbf{a}(\theta_l) = \begin{bmatrix} 1 & e^{-I\theta_l} & \dots & e^{-I(N_x-1)\theta_l} \end{bmatrix}^T. \quad (49)$$

The vector $\mathbf{n} \in \mathbb{H}_{\mathbb{C}}^{N_x}$ contains unpolarized noise contributions on the vector-sensor array. Also, the s_l coefficients correspond

$$\mathbf{u}_1 = \begin{pmatrix} 0.333 & 0.095I & 0.031I & 0.317I \\ 0.347 + 0.366I & +\mathbf{i} & 0.125 - 0.087I & +\mathbf{j} & 0.021 - 0.011I & +\mathbf{k} & 0.344 - 0.337I \\ 0.287 + 0.036I & 0.1899 - 0.196I & 0.116 + 0.162I & -0.030 - 0.258I \end{pmatrix} \quad (39)$$

$$\mathbf{u}_2 = \begin{pmatrix} -0.429 & -0.306I & -0.011I & -0.300I \\ -0.356 - 0.039I & +\mathbf{i} & 0.045 - 0.245I & +\mathbf{j} & 0.013 - 0.033I & +\mathbf{k} & 0.010 - 0.257I \\ -0.307 - 0.290I & 0.150 - 0.201I & 0.035 + 0.071I & 0.259 - 0.236I \end{pmatrix}. \quad (40)$$

to the magnitude contribution of the l^{th} wave (at a fixed frequency). In the following, we use the notation

$$\mathbf{d}_l(\theta_l, \rho_{1l}, \varphi_{1l}, \rho_{2l}, \varphi_{2l}) = p_l(\rho_{1l}, \varphi_{1l}, \rho_{2l}, \varphi_{2l})\mathbf{a}(\theta_l) \quad (50)$$

where $\mathbf{d}_l$ is called the *polarized steering vector* of the l^{th} wave and so that the observations can be written as

$$\mathbf{x} = \sum_{l=1}^L \mathbf{d}_l s_l + \mathbf{n}. \quad (51)$$

The biquaternion observation vector $\mathbf{x} \in \mathbb{H}_{\mathbb{C}}^{N_x}$ is built from the observations (in frequency domain) $\mathbf{x}_{c_1}, \mathbf{x}_{c_2}, \mathbf{x}_{c_3}$ on the three components as

$$\mathbf{x} = \mathbf{x}_{c_1} \mathbf{i} + \mathbf{x}_{c_2} \mathbf{j} + \mathbf{x}_{c_3} \mathbf{k}. \quad (52)$$

C. Long-Vector Approach

As a comparison, the long-vector approach classically used in vector-sensor array processing [2], [3] makes use of the concatenated vector $\mathbf{x}_{lv}$ built the following way:

$$\mathbf{x}_{lv} = [\mathbf{x}_{c_1}^T | \mathbf{x}_{c_2}^T | \mathbf{x}_{c_3}^T]^T \quad (53)$$

with $\mathbf{x}_{lv} \in \mathbb{C}^{3N_x}$. The long-vector approach allows, with additional parametrization, the use of classical matrix algebra algorithms and was used to define MUSIC- and ESPRIT-like algorithms for vector-sensor arrays [3], [8]. However, the use of long vectors has some drawbacks, such as leading to “over computation” and breaking the local polarized structure of the data. This last point has no deep consequences in the presented algorithm but could be of importance in more complicated ones, for example, if higher order statistics (HOS) are used. The use of long vectors in a processing involving HOS would lead to (highly) complicated structures in tensor valued cumulants or cost functions. We claim here that the use of hypercomplex numbers (and more generally the use of geometric numbers/algebras) can lead to easier manipulation of vector valued signals.

V. BIQUATERNION MUSIC ESTIMATOR

The BQ-MUSIC algorithm is based on the decomposition of the biquaternionic spectral matrix of the observation data vector $\mathbf{x}$ into signal and noise orthogonal subspaces. Using the modelization and linear algebra tools previously presented, we derive in the sequel an expression for this new BQ-MUSIC estimator.

A. Biquaternionic Spectral Matrix

Since second-order statistics of the observed data are used in the BQ-MUSIC, we now introduce the biquaternionic spectral matrix. All the biquaternion valued signals are considered centered here.

1) *Definition:* Considering that the output of the vector-sensor array is $\mathbf{x} \in \mathbb{H}_{\mathbb{C}}^{N_x}$ given in (51), then the spectral matrix is defined as

$$\mathbf{\Lambda} = \mathbb{E}[\mathbf{x}\mathbf{x}^\dagger] \in \mathbb{H}_{\mathbb{C}}^{N_x \times N_x} \quad (54)$$

The mathematical expectation $\mathbb{E}[\cdot]$ is defined naturally over $\mathbb{H}_{\mathbb{C}}$, just like it is done over $\mathbb{C}$ or $\mathbb{H}$ [34]. Substituting (51) in (54) and assuming decorrelation between the different sources (i.e.,

$\mathbb{E}[s_f \overline{s_v}] = 0$ for $f \neq v$) and between sources and noise (i.e., $\mathbb{E}[s_f \overline{n_v}] = 0 \forall f, v$), the biquaternionic spectral matrix takes the following form:

$$\mathbf{\Lambda} = \sum_{l=1}^L \sigma_l^2 \mathbf{d}_l \mathbf{d}_l^\dagger + \mathbf{\Lambda}_n \quad (55)$$

where σ_l^2 are the powers of the L sources on the antenna and $\mathbf{d}_l \in \mathbb{H}_{\mathbb{C}}^{N_x}$ are the *biquaternionic source vectors* describing source contributions on the antenna. The matrix $\mathbf{\Lambda}_n$ is given by $\mathbf{\Lambda}_n = \mathbb{E}[\mathbf{n}\mathbf{n}^\dagger] = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_{N_x}^2)$, where $\sigma_f^2 = \mathbb{E}[n_f \overline{n_f}]$ is the power of the noise on the f^{th} sensor. In order to build a MUSIC estimator, it is necessary to decompose the observation data spectral matrix into orthogonal subspaces, using the algorithm derived in Section III-B-1).

B. BQ-MUSIC Estimator

As presented in (50), every polarized wave impinging on the vector-sensor array is parametrized by five parameters, and the proposed version of MUSIC aims to estimate the five of them simultaneously. In order to do so, and as usual in MUSIC approach, a parametrized steering vector is projected onto the *noise* subspace built using the last eigenvectors of the spectral matrix of the observations. The biquaternionic steering vector has the following expression:

$$\mathbf{f}(\Omega) = \frac{1}{\mathcal{N}} \begin{bmatrix} p & p e^{-I\theta} & \dots & p e^{-I(N_x-1)\theta} \end{bmatrix}^T \quad (56)$$

where

$$\begin{aligned} \Omega &= \{\theta, \rho_1, \rho_2, \varphi_1, \varphi_2\} \\ p &= \mathbf{i} + \rho_1 e^{I\varphi_1} \mathbf{j} + \rho_2 e^{I\varphi_2} \mathbf{k} \\ \mathcal{N} &= \sqrt{N_x(1 + \rho_1^2 + \rho_2^2)}. \end{aligned}$$

Then, the BQ-MUSIC consists of finding the set of parameters Ω that maximizes the following functional:

$$\mathcal{F}(\Omega) = \frac{1}{\mathbf{f}^\dagger(\Omega) \mathbf{\Pi}_N \mathbf{f}(\Omega)} \quad (57)$$

where $\mathbf{\Pi}_N = \sum_{p=2L+1}^{2N_x} \mathbf{u}_p \mathbf{u}_p^\dagger$, built with the last $2(N_x - L)$ eigenvectors of $\mathbf{\Lambda}$ is the orthogonal biquaternionic projector on noise subspace. One can see that the use of hypercomplex numbers allows an estimator expression very similar to scalar-valued signal, without any “additional” structure in the projector, except the algebra on which it is expressed. The functional $\mathcal{F}(\Omega)$ has maxima for the values of the parameters corresponding to polarized plane waves that have impinged on the local vector-sensor array. In the case where those parameters are unknown, finding these maxima will consist in finding the extrema of a 5-D surface. The use of a biquaternionic formulation for polarized MUSIC estimator has not been studied for this optimization problem. Consequently, the presented study does not allow to conclude on possible advantages of the proposed approach among others on this aspect of the algorithm. We present next some results for the *long-vector* and the *biquaternion* approaches regarding computational and orthogonality issues.

C. Computational Issues

If the three-component *long-vector* model (53) is used, the spectral matrix is complex of size $3N_x \times 3N_x$. Compared

to this long-vector matrix having $9N_x^2$ complex entries, the spectral matrix in the biquaternionic approach has N_x^2 biquaternion-valued coefficients. As a biquaternion is composed of four complex numbers, the biquaternion spectral matrix can thus be represented on $4N_x^2$ complex values. This way, the memory requirements for data covariance representation are reduced by a factor of 4/9, provided that a biquaternion model is used.

D. Orthogonality Issues

As we saw in Section V-B, BQ-MUSIC algorithm is based on the orthogonality between biquaternion-vectors. We show next that this orthogonality constraint implies stronger relationships between the three components of the signal than the long-vector approach does. Consider two biquaternionic vectors $\mathbf{x}, \mathbf{y} \in \mathbb{H}_{\mathbb{C}}^{N_x}$, with their expressions given by

$$\begin{cases} \mathbf{x} = \mathbf{x}_{c1} \mathbf{i} + \mathbf{x}_{c2} \mathbf{j} + \mathbf{x}_{c3} \mathbf{k} \\ \mathbf{y} = \mathbf{y}_{c1} \mathbf{i} + \mathbf{y}_{c2} \mathbf{j} + \mathbf{y}_{c3} \mathbf{k}. \end{cases} \quad (58)$$

The corresponding long-vector representations [see (53)] are $\mathbf{x}_{lv}, \mathbf{y}_{lv} \in \mathbb{C}^{3N_x}$, as follows:

$$\mathbf{x}_{lv} = \begin{pmatrix} \mathbf{x}_{c1} \\ \mathbf{x}_{c2} \\ \mathbf{x}_{c3} \end{pmatrix} \quad \text{and} \quad \mathbf{y}_{lv} = \begin{pmatrix} \mathbf{y}_{c1} \\ \mathbf{y}_{c2} \\ \mathbf{y}_{c3} \end{pmatrix}. \quad (59)$$

By imposing the orthogonality constraint for the biquaternion vectors

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{H}_{\mathbb{C}}} = 0 \quad (60)$$

one gets the following relationships between the complex components:

$$\mathbf{y}_{c1}^\dagger \mathbf{x}_{c1} + \mathbf{y}_{c2}^\dagger \mathbf{x}_{c2} + \mathbf{y}_{c3}^\dagger \mathbf{x}_{c3} = 0 \quad (61)$$

$$\mathbf{y}_{c3}^\dagger \mathbf{x}_{c2} = \mathbf{y}_{c2}^\dagger \mathbf{x}_{c3} \quad (62)$$

$$\mathbf{y}_{c1}^\dagger \mathbf{x}_{c3} = \mathbf{y}_{c3}^\dagger \mathbf{x}_{c1} \quad (63)$$

$$\mathbf{y}_{c2}^\dagger \mathbf{x}_{c1} = \mathbf{y}_{c1}^\dagger \mathbf{x}_{c2}. \quad (64)$$

The orthogonality constraint for the long-vector approach

$$\langle \mathbf{x}_{lv}, \mathbf{y}_{lv} \rangle_{\mathbb{C}} = 0 \quad (65)$$

yields only (61), implying that

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{H}_{\mathbb{C}}} = 0 \quad \Rightarrow \quad \langle \mathbf{x}_{lv}, \mathbf{y}_{lv} \rangle_{\mathbb{C}} = 0. \quad (66)$$

The reciprocal is not always true meaning that the biquaternionic orthogonality imposes stronger constraints between the components of the vector-sensor array, and implicitly between the signal and noise subspaces. This affects in a positive way the robustness of BQ-MUSIC algorithm to different kinds of errors as we show in the next section.

The following section compares some simulation results on the resolution and robustness of the BQ-MUSIC estimator to the long-vector approach and to PSA-MUSIC proposed by Rahmim [15], which uses the polarization information to improve the spectral matrix conditioning.

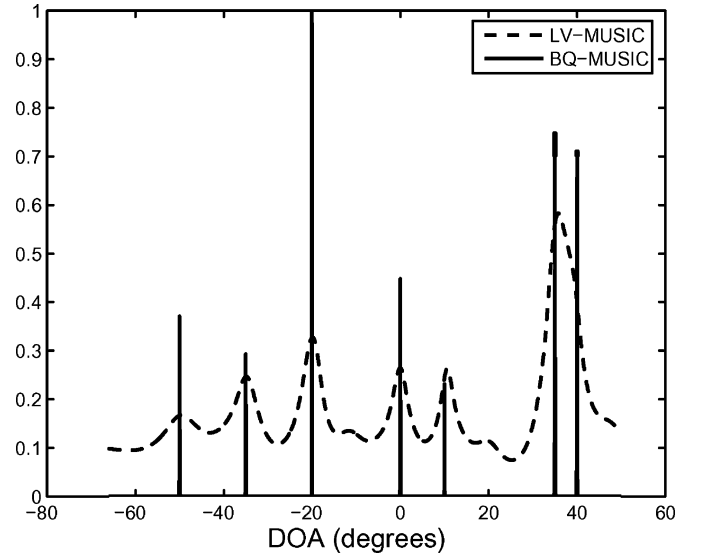


Fig. 1. Robustness to polarization parameters errors.

VI. SIMULATION RESULTS

By maximizing the functional (57) over the five parameters simultaneously, it is possible to jointly estimate the DOA and the polarization parameters for the sources impinging on the antenna. For computational power reasons, we supposed in this section that the polarization parameters were known or they have been estimated previously and we focused only on the estimation of the direction of arrival parameter θ . In practice, this situation corresponds to DOAs estimation for sources of known polarization, as it is often the case in electromagnetics. Before presenting the simulation results, notice that on an array of N_x three-component vector-sensors, the BQ-MUSIC algorithm allows detection of maximum $N_x - 1$ sources while the long-vector approach (LV-MUSIC) detects a maximum number of $3N_x - 1$ sources. This reduction of the signal subspace dimension is directly related to the fact that a stronger orthogonality constraint is imposed between signal and noise subspaces (as shown in Section V-D). On the other hand, this stronger constraint increases the algorithm robustness to noise, model errors and polarization parameters estimation errors as we show in simulations.

First, we consider an array of 20 vector-sensors and seven sources of known polarization parameters $(\rho_1^l, \rho_2^l, \varphi_1^l, \varphi_2^l)$, $l = 1 \dots 7$ impinging on the antenna. The simulated DOAs for the sources are as follows: -50° , -35° , -20° , 0° , 10° , 35° , 40° , and the SNR = 30 dB. If the polarization parameters are correctly estimated, the two algorithms (BQ-MUSIC and LV-MUSIC) perform identically well. For the plots in Fig. 1, we supposed that the estimated polarization parameters were slightly biased (the perturbation bias has a equal to 5% of the norm of the original vector). The DOA detection results for the two algorithms are presented. The detection curves corresponding to each of the seven sources were superposed in order to have all results on the same plot (Fig. 1). The long-vector approach undergoes a serious loss in resolution power, failing to discriminate sources 6 and 7, while BQ-MUSIC performs a

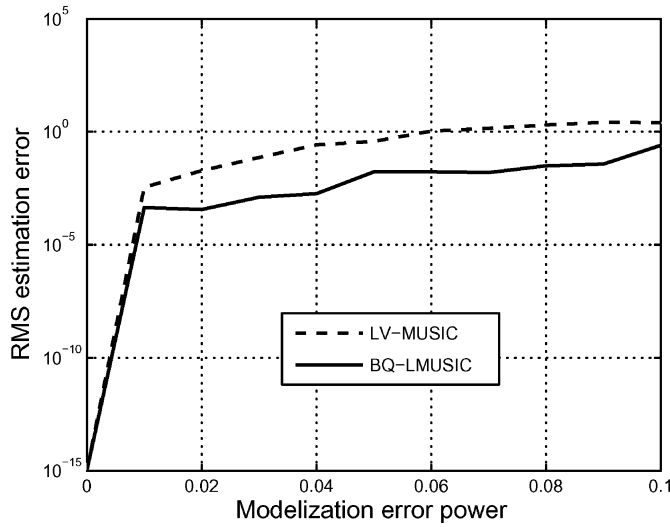


Fig. 2. RMS estimation error for modelization errors.

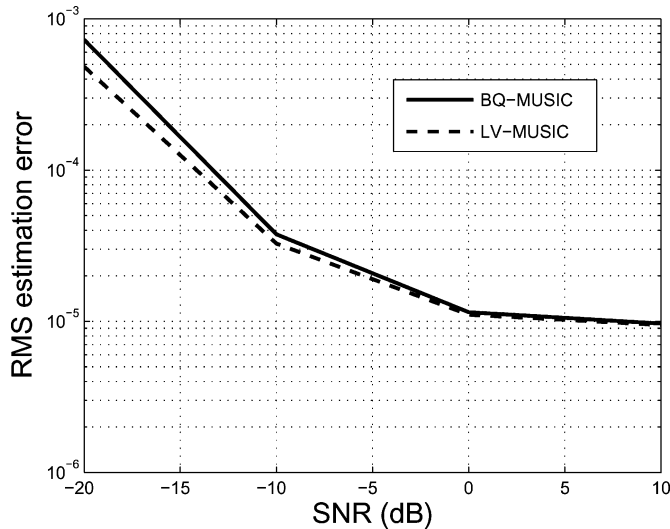


Fig. 3. RMS estimation error for one source in the presence of noncoherent noise.

very accurate detection, proving to be more robust to polarization parameters estimation errors.

The second simulation tests the robustness of the algorithms to modelization errors. The same set of sources is considered as before. We assume that the model used for the source vectors $\mathbf{d}_l$ [(50)] is not accurate and we modeled this lack of knowledge by an additive Gaussian noise of variable power. Fig. 2 plots the root-mean-square (RMS) error for the estimation of the DOA of source number 4 ($\theta = 0^\circ$), versus the energy of the noise corrupting the model. For each point on the image, 100 runs were used. As expected, for a perfectly fitting model, the errors for the two methods approach zero. As the error increases, BQ-MUSIC overperforms the classical approach and seems to be more robust to modelization errors.

Fig. 3 illustrates the behavior of the two algorithms to noncoherent noise on the sensors. A scenario with one source of DOA 10° , impinging on a ten-vector-sensors array was considered. We supposed that the snapshots were corrupted by additive Gaussian, nonpolarized, spatially white noise. The polar-

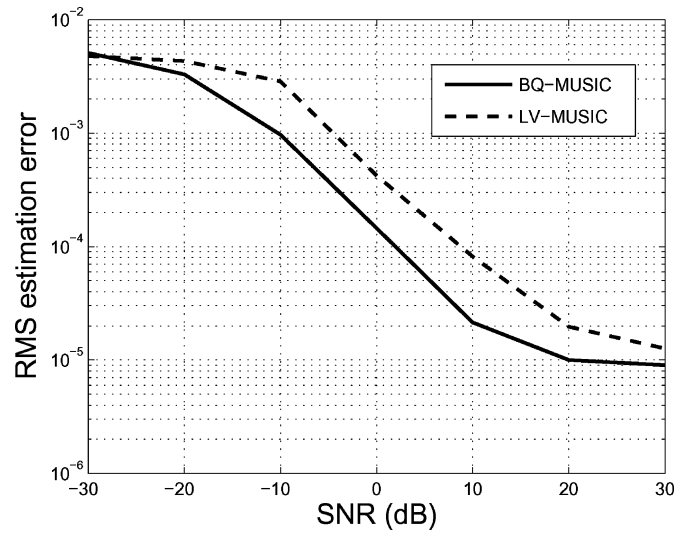


Fig. 4. RMS estimation error for one source in the presence of coherent noise.

ization parameters of the source are supposed perfectly known. For each point, 100 runs were used. We plotted the RMS estimation error for the source DOA estimation versus the SNR. The proposed algorithm performs fairly well compared with the long-vector approach, with only a slight loss of accuracy for very low SNR.

In addition, we tested the robustness of the BQ-MUSIC algorithm to coherent noise as it is well known that this is the weak point of MUSIC-like algorithms. We considered the same configuration as before, but this time, the additive noise is coherent along the array and on the three components. Noncoherent noise was also injected with a signal-to-noise ratio of 0 dB. The results of the simulation are presented in Fig. 4 which plots the estimation error for the DOA of the source versus the signal-to-coherent-noise ratio. The BQ-MUSIC algorithm proves to be more robust to this kind of errors than its long-vector version. The strange form of the detection curves for low SNR (< 10 dB) can be explained by the fact that when the coherent noise becomes important, it behaves as an interfering source, biasing the signal subspace estimation and strongly perturbing the detection of the targeted source. For high values of SNR, noncoherent noise becomes more important than the coherent one, and we fall into the configuration previously studied.

As we mentioned at the beginning of this section, the “long-vector” approach allows the detection of maximum number of sources almost three times larger than BQ-MUSIC; therefore, the comparison between algorithms is not completely fair. In the sequel, we compare in simulations BQ-MUSIC with PSA-MUSIC [15] a high-resolution technique based on PSA. The idea behind this algorithm is to use the polarization information to improve the estimation of the spectral matrix, by averaging over the three components of the antenna. As a result, the information on the polarization parameters is lost, which is not the case for LV-MUSIC and BQ-MUSIC; the maximum number of detectable sources is $N_x - 1$ (the same as BQ-MUSIC).

We considered two scenarios, the first with six sources impinging on a seven-vector-sensor array (Fig. 5) and the second with only one source (Fig. 6). In the first case, the sources have

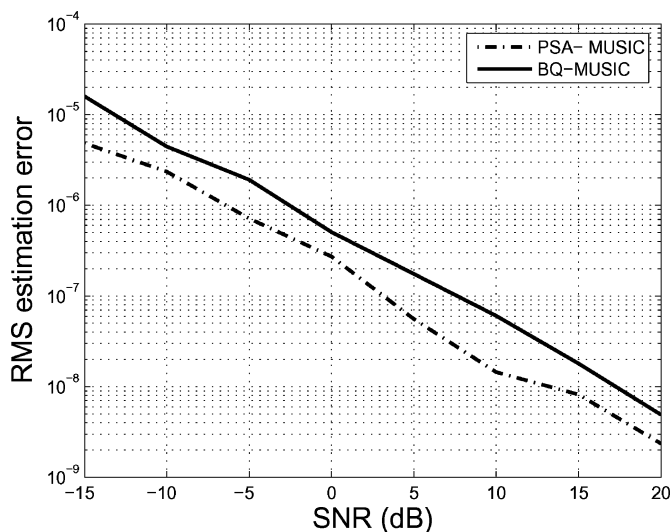


Fig. 5. RMS estimation error in the presence of six sources recorded on seven sensors, for different SNR.

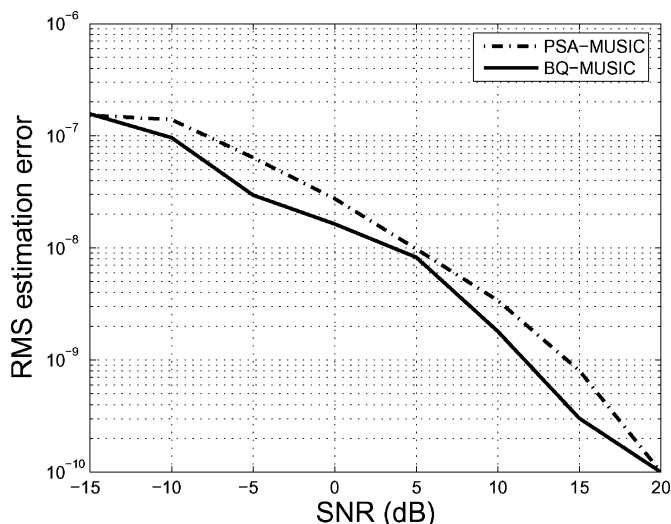


Fig. 6. RMS estimation error in the presence of one source recorded on seven sensors, for different SNR.

different polarizations and their DOAs are -50° , -35° , -20° , 0° , 10° , 35° , 40° . In Fig. 5, we plotted the RMS estimation error for the forth source with respect to the SNR (in decibels). In the second, the same curve was plotted (Fig. 6), assuming the presence of only one source of DOA equal to -20° in the recorded data. For each point on the figures, 100 runs were used, and the spectral matrix was estimated with 50 samples. One can remark that when the number of sources is large (equal to the limit of MUSIC algorithm), PSA-MUSIC performs better than the biquaternion algorithm. This can be explained by the fact that the estimation of the spectral matrix is more accurate in the case of PSA-MUSIC, because the number of samples used for estimation is three times larger than for BQ-MUSIC algorithm (in the case of PSA the three components can be assimilated to three snapshots). However, when the noise subspace dimension grows, the biquaternion orthogonality constraint prevails and BQ-MUSIC behaves better than PSA-MUSIC (Fig. 6). The main advantage of our algorithm over PSA-MUSIC is the preservation of the polarization information of the sources.

Meanwhile, if the polarization parameters are unknown, the performance of BQ-MUSIC is expected to degrade. A version of BQ-MUSIC including the estimation of polarization information will be the focus of future work.

VII. CONCLUSION

In this paper, we proposed a MUSIC-like algorithm (BQ-MUSIC) for three-component vector-sensor array processing, based on *biquaternions*. The performances of this algorithm are compared in simulations to the classical approach (LV-MUSIC) based on the concatenation of the three components in a *long vector* and with PSA-MUSIC, which performs an average over the three components. Furthermore, we present a technique for the decomposition of biquaternion-valued matrices into eigenelements.

The BQ-MUSIC algorithm is based on a quaternionic model of a polarized source, and it is well adapted to the acquisition geometry. The use of this model preserves the polarization information and imposes a stronger orthogonality constraint between the signal and noise subspaces. As a result, the proposed method proves to be more robust to coherent noise, modelization errors, and polarization parameters estimation errors. Nevertheless, the use of biquaternions provides a more compact and elegant way of handling multicomponent signals.

Also, this paper illustrates the high potentiality of high-dimensional algebras (and especially geometric algebras) to model complex-structured data in signal processing.

REFERENCES

- [1] K. T. Wong and M. D. Zoltowski, "Uni-vector-sensor ESPRIT for multi-source azimuth, elevation and polarization estimation," *IEEE Trans. Antennas Propag.*, vol. 45, no. 10, pp. 1467–1474, 1997.
- [2] A. Nehorai and E. Paldi, "Vector-sensor array processing for electromagnetic source localization," in *Proc. 25th Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, 1991, pp. 566–572.
- [3] A. Nehorai and E. Paldi, "Vector-sensor array processing for electromagnetic source localization," *IEEE Trans. Signal Process.*, vol. 42, no. 2, pp. 376–398, Feb. 1994.
- [4] K. T. Wong, L. Li, and M. D. Zoltowski, "Root-MUSIC-based direction-finding & polarization-estimation using diversely-polarized possibly-collocated antennas," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, no. 8, pp. 129–132, 2004.
- [5] K. T. Wong and M. D. Zoltowski, "Root-MUSIC-based azimuth-elevation angle of arrival estimation with uniformly spaced but arbitrarily oriented velocity hydrophones," *IEEE Trans. Signal Process.*, vol. 47, no. 12, pp. 3250–3260, Dec. 1999.
- [6] K. T. Wong and M. D. Zoltowski, "Self-initiating MUSIC direction finding and polarization estimation in spatio-polarizational beam space," *IEEE Trans. Antennas Propag.*, vol. 48, no. 8, pp. 1235–1245, 2000.
- [7] M. D. Zoltowski and K. T. Wong, "ESPRIT-based 2-D direction finding with a sparse array of electromagnetic vector-sensors," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2195–2204, Aug. 2000.
- [8] J. Li and T. Compton, "Angle estimation using a polarization sensitive array," *IEEE Trans. Antennas Propag.*, vol. 39, pp. 1539–1543, 1991.
- [9] J. Li and T. Compton, "Performance analysis for angle and polarization estimation using ESPRIT," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Apr. 1992, pp. V417–V420.
- [10] J. Li and T. Compton, "Two-dimensional angle and polarization estimation using the ESPRIT algorithm," *IEEE Trans. Antennas Propag.*, vol. 40, no. 5, pp. 550–555, May 1992.
- [11] J. Li and T. Compton, "Angle and polarization estimation in a coherent signal environment," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 3, pp. 706–716, Jul. 1993.
- [12] A. J. Weiss and B. Friedlander, "Performance analysis of diversely polarized antenna arrays," *IEEE Trans. Signal Process.*, vol. 39, no. 7, pp. 1589–1603, Jul. 1991.

- [13] N. Le Bihan and J. Mars, "Singular value decomposition of quaternion matrices: A new tool for vector-sensor signal processing," *Signal Process.*, vol. 84, no. 7, pp. 1177–1199, 2004.
- [14] S. Miron, N. Le Bihan, and J. Mars, "Quaternion-music for vector-sensor array processing," *IEEE Trans. Signal Process.*, vol. 4, no. 54, pp. 1218–1229, 2006.
- [15] D. Rahamim, J. Tabrikian, and R. Shavit, "Source localization using vector sensor array in multipath environment," *IEEE Trans. Signal Process.*, vol. 52, no. 11, pp. 3096–3103, Nov. 2004.
- [16] J. Xavier and J. Manton, "On the generalization of AR processes to Riemannian manifolds," in *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Toulouse, France, 2006, pp. V–V.
- [17] W. Hamilton, "On the geometrical interpretation of some results obtained by calculation with biquaternions," in *Proc. Roy. Irish Acad.*, 1853, vol. V, pp. 388–390.
- [18] W. Hamilton, "On quaternions," in *Proc. Roy. Irish Acad.*, 1843, pp. 1–16.
- [19] J. Ward, *Quaternions and Cayley Numbers, Algebra and Applications*. New York: Kluwer Academic, 1997.
- [20] I. Kantor and A. Solodovnikov, *Hypercomplex Numbers, an Elementary Introduction to Algebras*. New York: Springer-Verlag, 1989.
- [21] P. Lounesto, *Clifford Algebra and Spinors*, ser. Electrical and computer engineering: Digital signal processing. Cambridge, U.K.: Cambridge Univ. Press, 1997.
- [22] D. Hestenes and G. Sobczyk, *Clifford Algebra to Geometric Calculus: A Unified Language for Physics and Mathematics*. Amsterdam, The Netherlands: Reidel, 1984.
- [23] Y. Tian, "Matrix theory over the complex quaternion algebra," *ArXiv Mathematics e-prints* 2000 [Online]. Available: <http://arxiv.org/abs/math/0004005>
- [24] V. Majernick, "Quaternionic formulation of the classical fields," *Appl. Clifford Algebras*, vol. 9, no. 1, pp. 119–130, 1999.
- [25] J. Edmond, "Nature's natural numbers: Relativistic quantum theory over the ring of complex quaternions," *Int. J. Theoret. Phys.*, vol. 6, no. 3, pp. 205–224, 1972.
- [26] M. Mehta, *Matrix Theory, Selected Topics and Useful Results*. Delhi, India: Hindustan Publishing Corp., 1989, 2nd ed.
- [27] L. Wolf, "Similarity of matrices in which the elements are real quaternions," *Bull. Amer. Math. Soc.*, vol. 42, pp. 737–743, 1936.
- [28] I. Porteous, *Clifford Algebra and Classical Groups*. Cambridge, U.K.: Cambridge Univ. Press, 1995.
- [29] F. Zhang, "Quaternions and Matrices of Quaternions," *Linear Algebra Its Appl.*, vol. 251, pp. 21–57, 1997.
- [30] L. Huang and W. So, "On left eigenvalues of a quaternionic matrix," *Linear Algebra Its Appl.*, vol. 323, pp. 105–116, 2001.
- [31] H. Lee, "Eigenvalues and canonical forms of matrices with quaternions coefficients," *Proc. Roy. Irish Acad.*, vol. 52A, no. 2, pp. 253–260, 1949.
- [32] S. Okubo, "Eigenvalue problem for symmetric 3×3 octonionic matrix," *Adv. Appl. Clifford Algebras*, vol. 1, no. 9, pp. 131–176, 1999.
- [33] R. Serôdio, E. Pereira, and J. Vitória, "Computing the zeros of quaternion polynomials," *Comput. Math. With Appl.*, vol. 42, no. 8, pp. 1229–1237, 2001.
- [34] P.-O. Amblard and N. Le Bihan, "On properness of quaternion valued random variables," presented at the 6th IMA Conf. Mathematics in Signal Processing, Cirencester, U.K., 2004.



His research interests include polarized signal processing using multilinear, geometric algebra and group theory techniques, and applications of signal processing in geophysics.



array processing, multilinear algebra, and hypercomplex numbers.



Berkeley. He is currently Professor of signal processing at the Institut National Polytechnique of Grenoble and is with the Laboratoire des Images et des Signaux, Grenoble, France. His research interests include seismic and acoustic signal processing, wave-field separation methods, time-frequency time-scale characterization, and applied geophysics.

Prof. Mars is a Member of the Society of Exploration Geophysics (SEG) and the European Association of Geoscientists and Engineers (EAGE).

Nicolas Le Bihan was born in Morlaix, France, in 1974. He received the B.Sc. degree in physics from the Université de Bretagne Occidentale (UBO), Brest, France, in 1997 and the M.Sc. degree and the Ph.D. degree in signal processing from the Institut National Polytechnique de Grenoble (INPG), Grenoble, France, in 1998 and 2001, respectively.

Since 2002, he is Chargé de Recherche at the Centre National de la Recherche Scientifique (CNRS) and is working with the Laboratoire des Images et des Signaux (UMR 5083) of Grenoble, France.

Sebastian Miron was born in Suceava, Romania, in 1977. He graduated from "Gh. Asachi" Technical University of Iasi, Romania, in 2001 and received the M.Sc. and Ph.D. degree in signal, image, and speech processing from the Institut National Polytechnique de Grenoble, France, in 2002 and 2005, respectively.

He is currently a Maître de Conférence at the University of Nancy, France, and he is conducting research at the Centre de Recherche en Automatique de Nancy (CRAN), Nancy. His current research interests include seismic data processing, vector-sensor

Jérôme I. Mars received the M.Sc. degree in geophysics from Joseph Fourier University, Grenoble, France, in 1986 and the Ph.D. degree in signal processing from the Institut National Polytechnique de Grenoble, France, in 1988.

From 1989 to 1992, he was a Postdoctoral Researcher at the Centre d'Étude des Phénomènes Aléatoires et Géophysiques, Grenoble, France. From 1992 to 1995, he was a visiting Lecturer and a Scientist in the Materials Sciences and Mineral Engineering Department, University of California,

THE $\mathbb{H}$ -ANALYTIC SIGNAL

Nicolas Le Bihan⁽¹⁾, Stephen J. Sangwine⁽²⁾

(1): GIPSA-Lab, Dept. Images and Signal, Grenoble, France,

(2): Dept. of Computing and Electronic Systems, Univ. of Essex, Colchester, United Kingdom
S.sangwine@ieee.org, Nicolas.le-bihan@gipsa-lab.inpg.fr

ABSTRACT

We consider the extension of the analytic signal concept known for real valued signals to the case of complex signals. This extension is based on the Quaternion Fourier Transform (QFT) and leads to the so-called $\mathbb{H}$ -analytic signal. After defining the $\mathbb{H}$ -analytic signal and giving some of its properties, we present a new notation for quaternions, named the polar Cayley-Dickson form, which allows the extension of instantaneous phase and amplitude for the $\mathbb{H}$ -analytic signal. Identification of the components of a complex signal are then performed through the analysis of its $\mathbb{H}$ -analytic signal. We illustrate these new ideas on simulations.

1. INTRODUCTION

The definition of an analytic signal for general complex signals is still an open question. When considering complex signals, the class of *proper* (or analytic, in the sense originally stated by [1]) signals contain the signals with real and imaginary parts having the same amplitude and being decorrelated, while the *improper* class contains the remaining complex signals. While the *proper* signal can be identified as the analytic signal (in the sense defined by Ville [1]) of a real signal (in fact, its real part), the *improper* signal has no such link with real signals. However, *improper* signals arise in different areas in signal processing such as communications, for example [2, 3, 4]. The aim of this paper is to propose an extension of the analytic signal concept for *improper* complex signals, and this requires the use of a Quaternion Fourier Transform. It must be noticed that previous extensions of the analytic signal concept already exist [5, 6], some based on Quaternion Fourier Transforms as well, but they all considered multidimensional real signals, while our approach here is about complex signals.

In previous work [7], Sangwine and Le Bihan proposed the use of the biquaternion Fourier Transform [8] to define a hyperanalytic signal. This previous approach was motivated by the definition of the complex envelope which had most of the “classical” properties and thus was an obvious candidate for the extension of the analytic signal to complex signals. In this paper, we demonstrate that the Quaternion Fourier Transform, as defined in [9], is “sufficient” to construct the so-called $\mathbb{H}$ -analytic signal. Provided that the axis of the Quaternion Fourier Transform is correctly chosen, it is possible to construct the $\mathbb{H}$ -analytic signal which exhibits the same properties as the “classical” analytic signal. In order to extend the concept of instantaneous phase and amplitude to *improper* complex signals, we also introduce a new quaternion representation, named the *polar Cayley-Dickson* form, which is helpful to interpret the $\mathbb{H}$ -analytic signal. Simulations illustrate the concept introduced in this paper.

2. PRELIMINARY CONCEPTS

We present here some useful concepts used in the definition of the $\mathbb{H}$ -analytic signal.

2.1 Quaternions

We review shortly some facts about quaternions. Details can be found for example in [10]. A quaternion q is a 4D hypercomplex number classically written in its Cartesian form as: $q = a + bi + jc + kd$, where $a, b, c, d \in \mathbb{R}$ are its components and where i, j and k are roots of -1 and multiply together like: $ij = -ji$ and $ijk = -1$. The norm of q is $|q| = (a^2 + b^2 + c^2 + d^2)^{\frac{1}{2}}$, its conjugate is $\bar{q} = a - bi - jc - kd$ and its inverse is $q^{-1} = \bar{q}/|q|^2$. Any quaternion q can be expressed in the polar form: $q = |q|(\sin(\theta) + \mu \cos(\theta)) = |q|\exp(\mu\theta)$. Another notation, called Cayley-Dickson notation, represents a quaternion as a complex number with complexified components (with a different imaginary unit), the following way: $q = s + rj$ where $s = a + ib$ and $r = c + id$. A quaternion is called unitary if $|q| = 1$ and any unitary quaternion can be written as: $\exp(\mu\theta)$. A pure quaternion q is such that $a = 0$. A pure unit quaternion is a square root of -1 . A quaternion basis is a 4D basis such as $\{1, \mu, \xi, \mu\xi\}$ where μ, ξ are two orthogonal pure unit quaternions¹. Over the set of quaternions $\mathbb{H}$, it is possible to define some involutions and we present here one of use in the sequel. Given a quaternion $q \in \mathbb{H}$ and a pure unit quaternion $p \in \mathbb{H}$, then $q_p = -pqp$ is an involution². Such involutions are useful in quaternion components identification, see for example [5]. More on quaternion involutions can be found in [11].

2.2 Generalized Cayley-Dickson form

Consider a quaternion valued signal $q(t)$ that can be expressed in a (generalized) Cayley-Dickson form:

$$q(t) = z_1(t) + z_2(t)\mu \quad (1)$$

where $z_1(t) = \Re(z_1(t)) + \xi\Im(z_1(t))$, $z_2(t) = \Re(z_2(t)) + \xi\Im(z_2(t))$ are complex signals and $\{1, \mu, \xi, \mu\xi\}$ is a quaternion basis. Such signals are representative of polarized signals for example [12, 13]. The two components of the Cayley-Dickson decomposition can be expressed as follows:

$$\begin{cases} z_1(t) = \frac{1}{2}(q_\xi(t) + q(t)) \\ z_2(t) = \frac{1}{2}(q_\xi(t) - q(t))\mu \end{cases} \quad (2)$$

¹Among all the possible quaternion basis, the most widely used is $\{1, i, j, k\}$

²Involution means here that $(q_p)_p = q$ and $(qm)_p = q_pm_p$ where $q, m \in \mathbb{H}$

where $q_\xi(t) = -\xi q(t)\xi$. In this notation, $z_1(t)$ is called the simplex part of $q(t)$ while $z_2(t)$ is called the perplex part of $q(t)$ (see [9] for details). Thus, any quaternion signal can be seen as a pair of complex signals in any quaternion basis.

2.3 Quaternion Fourier transform

Here, we present an important property of the Quaternion Fourier Transform (QFT). We make use of the right QFT definition given by Sangwine and Ell in [9].

Consider the complex signal $z(t) = \Re(z(t)) + \xi \Im(z(t))$, i.e. $z(t)$ takes values in $\mathbb{C}^\xi$. Before trying to build its $\mathbb{H}$ -analytic signal, we examine how it is transformed using a QFT of axis μ (noted as QFT_μ in the sequel), when $(1, \xi, \mu, \xi\mu)$ is a quaternion basis. The QFT_μ of $z(t)$ is thus:

$$\begin{aligned} Z(v) &= QFT_\mu[z(t)] = \int_{-\infty}^{+\infty} z(t)e^{-\mu 2\pi vt} dt \\ &= \int_{-\infty}^{+\infty} \Re(z(t)) [\cos(2\pi vt) - \mu \sin(2\pi vt)] dt \\ &\quad + \xi \int_{-\infty}^{+\infty} \Im(z(t)) [\cos(2\pi vt) - \mu \sin(2\pi vt)] dt \\ &= \int_{-\infty}^{+\infty} \Re(z(t)) \cos(2\pi vt) dt \\ &\quad - \mu \int_{-\infty}^{+\infty} \Re(z(t)) \sin(2\pi vt) dt \\ &\quad + \xi \int_{-\infty}^{+\infty} \Im(z(t)) \cos(2\pi vt) dt \\ &\quad - \xi \mu \int_{-\infty}^{+\infty} \Im(z(t)) \sin(2\pi vt) dt \end{aligned} \quad (3)$$

This last equality shows that the QFT_μ of $z(t)$ naturally makes the following decomposition/association:

- Even part of $\Re(z(t)) \rightarrow \Re(Z(v))$.
- Odd part of $\Re(z(t)) \rightarrow \Im_\mu(Z(v))$
- Even part of $\Im(z(t)) \rightarrow \Im_\xi(Z(v))$.
- Odd part of $\Im(z(t)) \rightarrow \Im_{\xi\mu}(Z(v))$

where $\Im_\eta$ (when η is a pure unit quaternion) stands for the η imaginary component of the quaternion. So, the QFT_μ of a complex signal $z(t) = \Re(z(t)) + \xi \Im(z(t))$ allows us to isolate the odd and even parts of its real and imaginary parts in the four different components of its $Z(v)$. This point guarantees that the symmetries of the real and imaginary parts are not mixed.

Now, consider two functions g and f such that: $g: \mathbb{R} \rightarrow \mathbb{C}$ and $f: \mathbb{R} \rightarrow \mathbb{R}$. Then, consider the QFT of their convolution:

$$\begin{aligned} QFT_\mu[g * f(t)] &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(\tau)f(t-\tau)d\tau e^{-\mu 2\pi vt} dt \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(\tau)e^{-\mu 2\pi v(t'+\tau)} f(t')d\tau dt' \\ &= \int_{-\infty}^{+\infty} g(\tau)e^{-\mu 2\pi v\tau}d\tau \int_{-\infty}^{+\infty} f(t')e^{-\mu 2\pi vt'}dt' \end{aligned} \quad (4)$$

and so:

$$\begin{aligned} QFT_\mu[g * f(t)] &= QFT_\mu[g(t)]QFT_\mu[f(t)] \\ &= QFT_\mu[f(t)]QFT_\mu[g(t)] \end{aligned} \quad (5)$$

Thus, the definition of the QFT we use here has the property of “verifying” the convolution theorem in the considered case of functions g and f . This will be of use for the extension of the analytic signal (definition of the Hilbert transform). Furthermore, the QFT of $f(t) = \frac{1}{\pi t}$ is given by:

$$F(v) = -\mu \text{sign}(v) \quad (6)$$

This is obvious from the possibility of calculating the QFT with axis μ from two complex Fourier transforms in an appropriate basis [9]. Here, as f is real valued, the change of basis has no effect.

3. THE $\mathbb{H}$ -ANALYTIC SIGNAL

We now give the definition and properties of the $\mathbb{H}$ -analytic signal based on the QFT.

3.1 Definition and properties

The $\mathbb{H}$ -analytic signal of $z(t)$ presented here has been worked out with an approach similar to the one originally developed by Ville [1]. The following definitions give the details of the construction of this signal. Note that the signal $z(t)$ is considered to be an *improper* complex signal, i.e. for example $\Re(z(t))$ and $\Im(z(t))$ are not orthogonal.

Definition 1. Consider a complex signal $z(t) = \Re(z(t)) + \xi \Im(z(t))$ and its Quaternion Fourier Transform $Z(v)$ given by:

$$Z(v) = QFT_\mu[z(t)] = \int_{-\infty}^{+\infty} z(t)e^{-\mu 2\pi vt} dt \quad (7)$$

where μ , the axis of the transform, is taken such that $(1, \xi, \mu, \xi\mu)$ is a quaternion basis. Then, the “Hilbert transform” of $z(t)$, noted $z_h(t)$, has the following QFT_μ :

$$Z_h(v) = -\mu \text{sign}(v)Z(v) \quad (8)$$

where the Hilbert transform is defined as:

$$HT[z(t)] = p.v. \left(z * \frac{1}{\pi t} \right)$$

The principal value (p.v.) is understood in its classical sense here (see [1] for example).

This definition of the Hilbert transform based on QFT_μ is derived thanks to the convolution property given in Section 2.3.

Definition 2. Given a complex valued signal $z(t)$ that can be expressed as $z(t) = \Re(z(t)) + \xi \Im(z(t))$, and given a pure unit quaternion μ such that $(1, \xi, \mu, \xi\mu)$ is a quaternion basis, then the $\mathbb{H}$ -analytic signal of $z(t)$, noted $z_a(t)$ is given by:

$$z_a(t) = z(t) + z_h(t)\mu \quad (9)$$

where $z_h(t)$ is the “Hilbert transform” of $z(t)$ given in Definition 1. The QFT of the $\mathbb{H}$ -analytic signal is thus:

$$Z_a(v) = Z(v) - \mu \text{sign}(v)Z(v)\mu \quad (10)$$

which is a direct extension of the “classical” analytic signal.

With this definition of the $\mathbb{H}$ -analytic signal given in Definition 2, we now investigate some of its properties.

Property 1. The spectrum of the $\mathbb{H}$ -analytic signal is right-sided, i.e. $Z_a(v) = 0, \forall v < 0$.

Proof. The QFT of $z_a(t)$ is given by:

$$\begin{aligned}
Z_a(v) &= Z(v) - \mu \operatorname{sign}(v) Z(v) \mu \\
&= \int_{-\infty}^{+\infty} z(t) e^{-\mu 2\pi v t} dt \\
&\quad - \mu \operatorname{sign}(v) \left(\int_{-\infty}^{+\infty} z(t) e^{-\mu 2\pi v t} dt \right) \mu \\
&= \int_{-\infty}^{+\infty} \Re(z(t)) \cos(2\pi v t) dt - \\
&\quad \mu \int_{-\infty}^{+\infty} \Re(z(t)) \sin(2\pi v t) dt + \\
&\quad \xi \int_{-\infty}^{+\infty} \Im(z(t)) \cos(2\pi v t) dt \\
&\quad - \xi \mu \int_{-\infty}^{+\infty} \Im(z(t)) \sin(2\pi v t) dt \\
&\quad - \mu \operatorname{sign}(v) \left(\int_{-\infty}^{+\infty} \Re(z(t)) \cos(2\pi v t) dt \right) \mu \\
&\quad + \mu \operatorname{sign}(v) \left(\mu \int_{-\infty}^{+\infty} \Re(z(t)) \sin(2\pi v t) dt \right) \mu \\
&\quad - \mu \operatorname{sign}(v) \left(\xi \int_{-\infty}^{+\infty} \Im(z(t)) \cos(2\pi v t) dt \right) \mu \\
&\quad + \mu \operatorname{sign}(v) \left(\xi \mu \int_{-\infty}^{+\infty} \Im(z(t)) \sin(2\pi v t) dt \right) \mu
\end{aligned}$$

Noting that μ and ξ commute with all the other terms (sin and cos, sign, $\Re$ and $\Im$) and remembering that $\xi \mu = -\mu \xi$, then the QFT of $z_a(t)$ takes the following simple expression:

$$Z_a(v) = (1 + \operatorname{sign}(v)) Z(v)$$

which completes the proof. $\square$

Property 1 together with the property of the QFT_μ given in section 2.3 show that the $\mathbb{H}$ -analytic signal is right-sided and at the same time, keeps the different part of the original signal in different imaginary components of the transform.

Property 2. The original signal $z(t)$ is the simplex part of its corresponding $\mathbb{H}$ -analytic signal $z_a(t)$. It is obtained by:

$$z(t) = \frac{1}{2} (z_a(t) - \xi z_a(t) \xi)$$

Note that if the original signal $z(t)$ is expressed in the classical complex basis $\{1, i\}$ and if the axis of the QFT is taken as j , then the $\Im_j$ and $\Im_k$ parts of $z_a(t)$ contain the Hilbert transform of $z(t)$. This property is a direct consequence of the way we have defined the $\mathbb{H}$ -analytic signal given in eq. (9), and this allows us to recover the original complex signal from its quaternion valued $\mathbb{H}$ -analytic signal. Note that this is the counterpart of the fact that, in the “classical” case, the original real signal is the real part of the analytic signal [1].

Also, note that our definition of the $\mathbb{H}$ -analytic signal includes the classical definition of Ville [1] as a special case. If the signal $z(t)$ is real, then the $\mathbb{H}$ -analytic signal is simply complex with the imaginary axis being the one chosen for the QFT.

4. THE POLAR CAYLEY-DICKSON FORM

We now look at the definition of the amplitude and phase concepts for the $\mathbb{H}$ -analytic signal. In order to so, we introduce a new notation for quaternions. It is different from the classical polar form and the polar form introduced in [5]. It is based on the Cayley-Dickson notation. Details about this new quaternion representation can be found in [14].

Definition 3. Any quaternion $q \in \mathbb{H}$ with Cartesian form as: $q = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{k}$ can be expressed in a polar Cayley-Dickson form:

$$q = A e^{Bj} \quad (11)$$

where $A = \Re(A) + i\Im(A) \in \mathbb{C}$ and $B = \Re(B) + i\Im(B) \in \mathbb{C}$.

This form of a quaternion q is the counterpart of the polar form of complex numbers. Here, the modulus and phase are complex valued. A method for finding A and B is detailed in [14].

Now, in the case of the $\mathbb{H}$ -analytic signal $z(t)$, its polar Cayley-Dickson form is given as:

$$z_a(t) = A_a(t) e^{B_a(t)j} \quad (12)$$

The values of the components (as well as the information they provide on the original signal) of this polar Cayley-Dickson form of the $\mathbb{H}$ -analytic signal are illustrated in the following section.

5. SIMULATIONS

We illustrate here the $\mathbb{H}$ -signal concept on a simple simulation example. Consider a complex signal $z(t)$ made up in the following way:

$$z(t) = f(t) \cdot (s_1(t) + i s_2(t)) \quad (13)$$

where $s_1(t) = \sin(2\pi v_1 t)$; and $s_2(t) = \sin(2\pi v_2 t + \xi)$ and $f(t) = \sin(2\pi v_f t)$ and with $v_f > v_1 > v_2$. The $\mathbb{H}$ -analytic signal of $z(t)$, i.e. $z_a(t)$, is computed, using j as the axis of the QFT, and expressed in its polar Cayley-Dickson form as in (12).

Then, from the polar Cayley-Dickson form of the $\mathbb{H}$ -analytic signal, and remembering that $A_a(t)$ and $B_a(t)$ are complex valued and can be expressed as $A_a(t) = |A_a(t)| \exp(\Psi_{A_a(t)})$ and $A_a(t) = |B_a(t)| \exp(\Psi_{B_a(t)})$, the following information is available:

$$\left\{ \begin{array}{ll} s_1(t) &= \Re(A_a(t)) \\ s_2(t) &= \Im(A_a(t)) \\ \Phi_2(t) - \Phi_1(t) &= \tan(\Psi_{A_a(t)}) \\ \frac{|z(t)|}{|f(t)|} &= |A_a(t)| \\ f(t) &= \mp \cos(|B_a(t)|) \end{array} \right. \quad (14)$$

where $\Phi_2(t)$ and $\Phi_1(t)$ are the instantaneous amplitudes of $s_2(t)$ and $s_1(t)$ respectively. In figure (1) we present, as a function of time, the complex amplitude $A_{z_a}(t)$ as well as the original signal $z(t)$. It shows that the complex envelope of the $\mathbb{H}$ -analytic signal, namely $A_{z_a}(t)$ is covering the original signal.

As presented above, $A_{z_a}(t)$ allows to recover parts of the original signal: $\Re(A_{z_a}(t)) = s_1(t)$ and $\Im(A_{z_a}(t)) = s_2(t)$. This interesting property could be of interest for example in

finding the modulation frequency of an improper complex signal (or a common component shared by the real and imaginary parts of an improper signal), as it allows a simple way of identifying the real and imaginary base band signals (here $s_1(t)$ and $s_2(t)$). This is a consequence of theorem 1 in [14].

In figure (2) the modulus of the original is compared to $|A_{z_a}(t)|$. It can be seen that the modulus of $A_{z_a}(t)$ is the envelope of the modulus of the original signal $z(t)$, which illustrates the concept of instantaneous amplitude to the case of improper complex signals. In Figure (3), the signal $f(t)$ is compared with the cosine of the modulus of the instantaneous complex phase $B_{z_a}(t)$. It can be seen that there is an ambiguity sign on some cycles, however, from an estimation point of view it can be seen that estimation of the frequency of $f(t)$ directly from $\cos(|B_{z_a}(t)|)$ is an easy task. Note that this could be used as a easy estimator of the correlation between real and imaginary components of an improper complex signal $z(t)$. Future work could investigate a comparison with the work proposed in [3]. Finally, in figure (4), the difference between the instantaneous frequencies of $s_1(t)$ and $s_2(t)$ (computed using the classical analytic signal) is compared to the tangent of the phase of the complex envelope, *i.e.* $\tan(\Psi_{A_a}(t))$. The perfect match between the two curves also suggest that it is possible to estimate the *relative instantaneous phase* between the real and imaginary components of an improper complex signal by inspection of the phase of the modulus of its $\mathbb{H}$ -analytic signal.

6. DISCUSSION AND CONCLUSIONS

We have introduced a new extension of the concept of analytic signal to the case of improper complex signals. The $\mathbb{H}$ -analytic signal is based on the use of the Quaternion Fourier transform. Some of its properties have been presented, that generalize in a straightforward manner the known results in the “classical” case. In order to access the information provided by the $\mathbb{H}$ -analytic signal, we have introduced a new representation for quaternions and linked the components of this representation to useful information on the improper original signal. In particular, the $\mathbb{H}$ -analytic signal allows direct access to common parts, relative instantaneous frequencies and uncorrelated components of the complex original signal. Applications of the $\mathbb{H}$ -analytic signal may be expected in the numerous applications dealing with improper

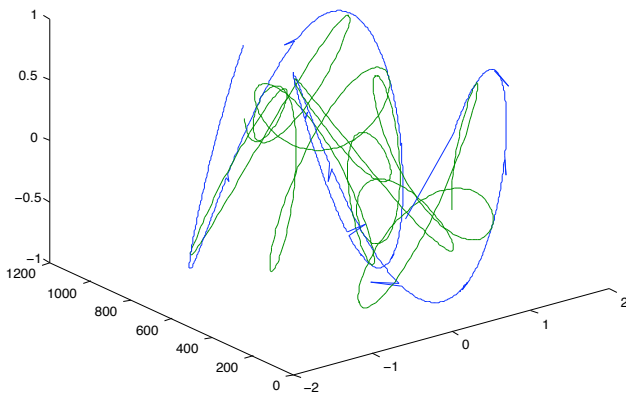


Figure 1: Original complex signal $z(t) = f(t) \cdot (s_1(t) + i s_2(t))$ (green) and complex envelope $A_{z_a}(t)$ (blue).

complex signals. In particular, some estimators of the mentioned characteristics of the improper signal could be based on the $\mathbb{H}$ -analytic signal and allow fast identification of, for example, the parameters of an unknown improper complex signal. Such estimators should be compared to existing work. Also, the possible definition of a $\mathbb{H}$ -analytic signal for improper complex signals suggests the possibility of defining some time-frequency representations for such signals, based on the quaternion Fourier transform.

REFERENCES

- [1] J. Ville, “Théorie et applications de la notion de signal analytique,” *Cables et Transmission*, vol. 2A, pp. 61–74, 1948.
- [2] P. Schreier and L. Scharf, “Second-order analysis of improper complex random vectors and processes,” *IEEE Trans. Signal Processing*, vol. 51, no. 3, pp. 714–725, March 2003.
- [3] P. Schreier, L. Scharf, and A. Hanssen, “A generalized likelihood ratio test for impropriety of complex signals,” *Signal Processing Letters, IEEE*, vol. 13, no. 7, pp. 433–436, July 2006.
- [4] P. Schreier, L. Scharf, and C. Mullis, “Detection and estimation of improper complex random signals,” *Information Theory, IEEE Transactions on*, vol. 51, no. 1, pp. 306–312, Jan. 2005.
- [5] T. Bulow and G. Sommer, “Hypercomplex signals—a novel extension of the analytic signal to the multidimensional case,” *IEEE Trans. Signal Processing*, vol. 49, no. 11, pp. 2844–2852, Nov 2001.
- [6] M. Felsberg and G. Sommer, “The monogenic signal,” *IEEE Trans. Signal Processing*, vol. 49, no. 12, pp. 3136–3144, Dec 2001.
- [7] S. J. Sangwine and N. Le Bihan, “Hypercomplex analytic signals : extension of the analytic signal concept to complex signals,” *XV European Signal Processing Conference (EUSIPCO), Poznan, Poland*, Sept. 2007.

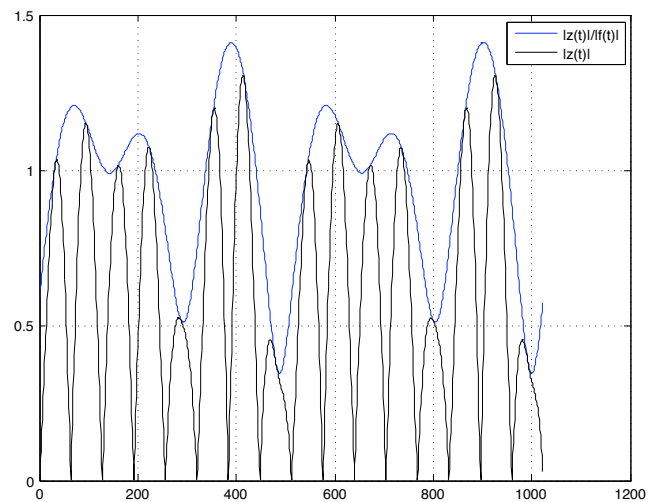


Figure 2: Modulus of the original signal $z(t)$ (black) and modulus of the complex envelope $A_{z_a}(t)$ (blue).

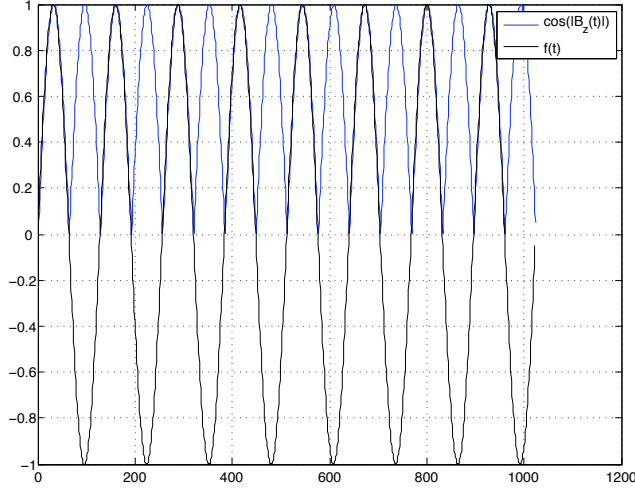


Figure 3: Cosine of the modulus of the complex phase of the $\mathbb{H}$ -analytic signal $|B_{za}(t)|$ (blue) and the $f(t)$ signal (black).

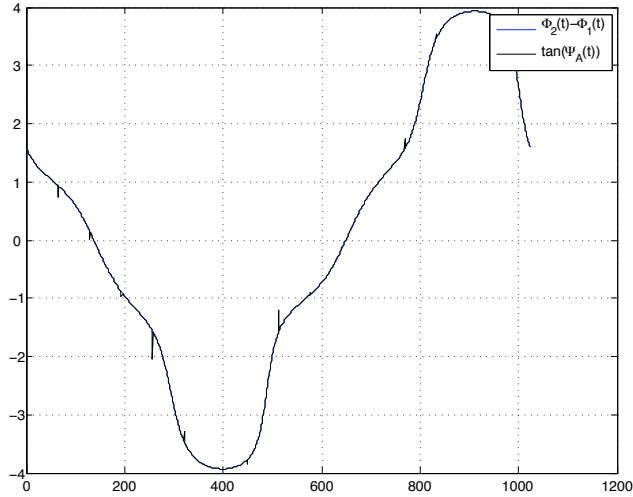


Figure 4: Tangent of the phase of the complex modulus of the $\mathbb{H}$ -analytic signal $A_{za}(t)$ (blue) and the difference of instantaneous phases between original signals $s_1(t)$ and $s_2(t)$, namely $f\Phi_2(t) - \Phi_1(t)$ (black).

orthogonal designs,” *IEEE Trans. Signal Processing*, vol. 56, no. 1, pp. 256–265, Jan. 2008.

- [13] O. Isaeva and V. Sarytchev, “Quaternion presentations polarization state,” *Combined Optical-Microwave Earth and Atmosphere Sensing, 1995. Conference Proceedings., Second Topical Symposium on*, pp. 195–196, 3–6 Apr 1995.
- [14] S. J. Sangwine and N. Le Bihan, “Quaternion polar representation with a complex modulus and complex argument inspired by the Cayley-Dickson form,” Preprint, Feb. 2008. [Online]. Available: <http://arxiv.org/abs/arxiv:0802.0852>

- [8] S. Said, N. Le Bihan, and S. J. Sangwine, “Fast complexified quaternion Fourier transform,” *IEEE Trans. Signal Processing*, vol. 56, no. 4, pp. 1522–1531, Apr. 2008.
- [9] T. A. Ell and S. J. Sangwine, “Hypercomplex Fourier transforms of color images,” *IEEE Trans. Image Processing*, vol. 16, no. 1, pp. 22–35, Jan. 2007.
- [10] J. P. Ward, *Quaternions and Cayley Numbers: Algebra and Applications*, ser. Mathematics and Its Applications. Dordrecht: Kluwer, 1997, vol. 403.
- [11] T. A. Ell and S. J. Sangwine, “Quaternion involutions and anti-involutions,” *Comput. Math. Appl.*, vol. 53, no. 1, pp. 137–143, 2007.
- [12] J. Seberry, K. Finlayson, S. S. Adams, T. A. Wysocki, T. Xia, and B. J. Wysocki, “The theory of quaternion

Quaternionic Independent Component Analysis using hypercomplex nonlinearities

By N. Le Bihan[†] and S. Buchholz[‡]

[†]: Laboratoire des Images et des Signaux, CNRS
ENSIEG, BP46, 38402 Saint-Martin d'Hères, FRANCE

[‡]: Cognitive Systems Group, Dept. of Computer Science,
CAU Kiel, 24098 Kiel, GERMANY

Nicolas.Le-Bihan@lis.inpg.fr, sbh@ks.informatik.uni-kiel.de

Abstract

We propose a quaternionic version of the Infomax algorithm to perform ICA on quaternion valued data. We introduce the three possible types of nonlinearities that can be used as activation functions and derive their differentiability properties. We show that only hypercomplex nonlinearity can lead to the estimation of all possible classes of proper quaternion random variable. We finally illustrate this on a simulation where a comparison of the results of separation obtained on a mixture of different proper signals with Infomax using different types of nonlinearities is presented. The proposed *fully quaternionic* Infomax algorithm is shown to be the only one to perform separation of polarized signal corrupted by $\mathbb{H}$ -proper (non-polarized) noise.

1. Quaternion random variables

Quaternions are 4D hypercomplex numbers. They form a non-commutative algebra and a quaternion q is given by $q = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k}$, where $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are pure imaginary numbers obeying the well-known multiplication rules: $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$ and $\mathbf{ijk} = -1$. For a introduction to quaternions, see Ward, J.P. (1997). Quaternion valued random variables (r.v.) have been studied in (Vakhania, N.N. (1998)) and (Amblard, P.O. & Le Bihan N. (2004)). As presented in (Amblard, P.O. & Le Bihan N. (2004)), it is possible to consider a quaternion random variable as a 4D random variable over $\mathbb{R}$, $\mathbb{C}$ or $\mathbb{H}$. Mathematical expectation is defined in a very natural way for quaternions, so that for a quaternion random variable q , the mean is simply $\mu_q = \mathbb{E}[q]$ and the variance is $\sigma_q^2 \mathbb{E} = [q\bar{q}]$, where $\bar{q}$ is the conjugate of q .

1.1. Properness

The concept of properness was first introduced in (Vakhania, N.N. (1998)) and generalized in (Amblard, P.O. & Le Bihan N. (2004)). There exists two levels of properness:

- A quaternion r.v. q is called $\mathbb{C}$ -proper if: $q \stackrel{d}{=} e^{\eta\varphi}q$, $\forall\varphi$ for one and only one imaginary unit $\eta = \mathbf{i}, \mathbf{j}$ or $\mathbf{k}$.
- A quaternion r.v. q is called $\mathbb{H}$ -proper if: $q \stackrel{d}{=} e^{\eta\varphi}q$, $\forall\varphi$ for any pure unit quaternion η .

It is possible to link such properties of quaternion random variables to statistical description of polarized signals.

1.2. Proper random variables and polarization

It is a well known fact in Physics (and more precisely in Optics) that polarized signal can be described using a Jones vector, which is in fact a spinor and isomorphic to a quaternion. Quaternion models for polarized signals have recently been used in Polarization Mode Dispersion (PMD) and Polarization Dispersion Losses (PDL) treatment (Karlsson, M. & Petersson M. (2004)). Thus, a polarized signal $s[m]$ with m samples can be written like $s[m] = s_1[m] + s_2[m]\mathbf{j}$ where s_1 and s_2 are complex valued and correspond to the components of the Electrical field (E_x, E_y) if, for example, propagation is in the z direction. If the signal is *purely* polarized, then $s[m]$ is deterministic. But if the signal is partially polarized, then $s[m]$ is random (Brosseau, C. (1998)).

As demonstrated in (Buchholz, S. & Le Bihan N. (2006)), if the two components of the signal are linked by a complex number (which phase and amplitude are linked with the polarization ellipsis), then its associated quaternion signal is $\mathbb{C}$ -proper. If the two components are decorrelated, then the signal is $\mathbb{H}$ -proper. Properness can be seen as a way to distinguish between polarized and non polarized signals. More precisely, unpolarized signals will be considered as noise in the sequel. The aim of the algorithm now proposed is to perform separation between different polarized signals or between polarized signals and noise.

2. Quaternionic Infomax

The Infomax algorithm was first proposed by Bell and Sejnowski (Bell, A.J., & Sejnowski, T.J. (1995)) and complex versions were proposed by Calhoun *et al.* (Calhoun, V. & Adalı, T. (2002), Adalı, T. & Kim, T. & Calhoun, V. (2004)). We present here a quaternionic version of this algorithm. Special attention is paid on hypercomplex nonlinearities that allow the separation achievement in section 3.

We consider that the observation datas $\mathbf{x}$ are a linear mixture of some quaternion valued sources $\mathbf{s}$:

$$\mathbf{x}[m] = \mathbf{A}\mathbf{s}[m] \quad m = 0, 1, \dots, M-1 \quad (2.1)$$

where $\mathbf{x} \in \mathbb{H}^N$, $\mathbf{s} \in \mathbb{H}^N$ and $\mathbf{A} \in \mathbb{H}^{N \times N}$. m is the time index. The mixing matrix is supposed constant in time and the sources (and consequently the observations) are supposed stationnary. The infomax algorithm maximizes the entropy of the output $\mathbf{y}[m]$ of a single layer neural network:

$$\mathbf{y}[m] = g(\mathbf{u}[m]) \quad (2.2)$$

where $g(\cdot)$ is a nonlinear function taking values in $\mathbb{H}$: $g : \mathbb{H} \rightarrow \mathbb{H}$. The vector $\mathbf{u}[m]$ is a weighted version of the input data $\mathbf{x}$:

$$\mathbf{u} = \mathbf{W}\mathbf{x} \quad (2.3)$$

where $\mathbf{W} \in \mathbb{H}^{N \times N}$ is the weighting matrix.

The entropy of the output vector $\mathbf{y} \in \mathbb{H}^N$ is given as:

$$H(\mathbf{y}) = -\mathbb{E}[\ln p(\mathbf{y})] = -\int_{-\infty}^{\infty} p(\mathbf{y}) \ln p(\mathbf{y}) d\mathbf{y} \quad (2.4)$$

The definition of the mathematical expectation is taken in the classical sense, such that for a quaternion valued random variable $q = q_0 + \mathbf{i}q_1 + \mathbf{j}q_2 + \mathbf{k}q_3$, we have $\mathbb{E}[q] = \mathbb{E}[q_0] + \mathbf{i}\mathbb{E}[q_1] + \mathbf{j}\mathbb{E}[q_2] + \mathbf{k}\mathbb{E}[q_3]$. The pdf of a quaternion random vector is in fact the joint pdf of its four vector components. As explained, it is possible to use real, complex or quaternion representation to handle quaternion random vectors and variables. So, the pdf of a quaternion random vector is: $p(\mathbf{y}) = p(\mathbf{q}_0, \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \simeq p(\mathbf{q}', \mathbf{q}'^*, \mathbf{q}'', \mathbf{q}''^*) \simeq p(\mathbf{q}, \mathbf{q}_i, \mathbf{q}_j, \mathbf{q}_k)$, where $\mathbf{q} = \mathbf{q}' + \mathbf{q}''\mathbf{j}$ ($\mathbf{q}' = \mathbf{q}_0 + \mathbf{q}_1\mathbf{i} \in \mathbb{C}^N$ and $\mathbf{q}'' = \mathbf{q}_2 + \mathbf{q}_3\mathbf{i} \in \mathbb{C}^N$) and $\mathbf{q}_\eta = -\eta\mathbf{q}\eta$ with $\sqrt{-1} = \eta$ and $\eta \in \mathbb{H}$.

In a similar way as for complex valued random vectors, the entropy of a quaternion valued random vector is the joint entropy of its components (and there is also three possible ways to represent it):

$$H(\mathbf{q}) \triangleq H(\mathbf{q}_0, \mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3) \quad (2.5)$$

The calculations for the weight update are the same as in the complex case, and so the update expression:

$$\Delta \mathbf{W} = \frac{\partial H(\mathbf{y})}{\partial \mathbf{W}} \mathbf{W}^\dagger \mathbf{W} = \mu [\mathbf{I} + \varphi(\mathbf{u})\mathbf{u}^\dagger] \mathbf{W} \quad (2.6)$$

where φ will be $-2 \tanh(\mathbf{u})$ in the case where $g(\mathbf{u}) = \tanh(\mathbf{u})$. Here, we only consider the tanh nonlinearity and turn now to the possible definitions of this nonlinearity and their differentiation properties.

3. Quaternionic nonlinearities

In the work of Calhoun *et al.*, it is emphasized that it is possible to use two different definitions for the nonlinearity: the full complex and split. In the tanh case, the two possible definitions, namely the *split* $g_s(u) = g_s(u_{Re} + \mathbf{i}u_{Im}) = \tanh(u_{Re}) + \mathbf{i} \tanh(u_{Im})$, and the *full* $g_f(u) = g_f(u_{Re} + \mathbf{i}u_{Im}) = \tanh(u_{Re} + \mathbf{i}u_{Im})$. We now propose the extension of this work to the quaternion case.

3.1. Nonlinearities definition

In the quaternionic case there is three possible cases for the nonlinearity definition:

- $\mathbb{H}$ -split: $g(u = u_0 + u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}) \triangleq \tanh(u_0) + \tanh(u_1)\mathbf{i} + \tanh(u_2)\mathbf{j} + \tanh(u_3)\mathbf{k}$
- $\mathbb{C}$ -split: $g(u_0 + u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}) \triangleq \tanh(u_0 + u_1\mathbf{i}) + \mathbf{j} \tanh(u_2 + u_3\mathbf{i})$
- $\mathbb{H}$ -full: $g(u_0 + u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}) \triangleq \tanh(u_0 + u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k})$

Clearly, the $\mathbb{H}$ -split nonlinearity process each component separately, while the $\mathbb{C}$ -split one process the two complex component independently. Those two types of nonlinearities are not able to take into account possible correlation between the components. Only the $\mathbb{H}$ -full one can do this. We show in Section 4 how this ability to take into account the relationship between components allow the identification of $\mathbb{C}$ -proper random variables.

3.2. Differentiability

It is possible to define three types of differentiability for quaternion valued functions. Consider such a function $f : \mathbb{H} \rightarrow \mathbb{H}$ for which $u \rightarrow f(u)$ when $u = u_0 + u_1\mathbf{i} + u_2\mathbf{j} + u_3\mathbf{k}$. Then, it is possible to define:

- $\mathbb{R}$ -differentiability: $[f]_{\mathbb{R}}' = \frac{\partial f}{\partial u} \triangleq \frac{\partial f}{\partial u_0} + \frac{\partial f}{\partial u_1}\mathbf{i} + \frac{\partial f}{\partial u_2}\mathbf{j} + \frac{\partial f}{\partial u_3}\mathbf{k}$
- $\mathbb{C}$ -differentiability: $[f]_{\mathbb{C}}' = \frac{\partial f^{(1)}}{\partial v} + \mathbf{j} \frac{\partial f^{(2)}}{\partial w}$
- $\mathbb{H}$ -differentiability: $[f]_{\mathbb{H}}' = \frac{df}{du}$

In the infomax context, where it is necessary to compute, for the nonlinear activation function f , the ratio $\frac{f''}{f'}$, using $\mathbb{R}$ -differentiation is equivalent to process separately the four components. Choosing an approach using $\mathbb{C}$ - or $\mathbb{H}$ -differentiation allow to take into account the possible links between the components. Note that in the $\mathbb{C}$ -differentiability, both complex components satisfy Cauchy-Riemann equations. In the $\mathbb{H}$ -differentiability case, as explained in Sudbery (Sudbery, A. (1979)), the condition imposed on f to be $\mathbb{H}$ -differentiable lead to the so-called Cauchy-Riemann-Fueter equation: $\frac{df}{du} = \frac{\partial f}{\partial u_0} = -\mathbf{i} \frac{\partial f}{\partial u_1} = -\mathbf{j} \frac{\partial f}{\partial u_2} = -\mathbf{k} \frac{\partial f}{\partial u_3}$. A direct consequence of this definition is that the only existing $\mathbb{H}$ -differentiable functions are of the form: $f(q) = aq + b$ where $a, b \in \mathbb{H}$ are constant quaternions. However, just like in the complex case (Calhoun, V. & Adalı, T. (2002)), it is possible to overcome this problem while using full hypercomplex nonlinearities by assuming that singularities of a hypercomplex function (which is not $\mathbb{H}$ -differentiable) have measure zero, which ensure that the update will not end up at such point.

4. Approximation and estimation properties

The purpose here is to see how well split- and fully- quaternionic approaches can "match" a certain distribution. We look here at the case where there is only one weight and that we consider only a random variable (random vector with dimension 1), then $y = g(wx)$. Then, remembering that it is possible to use vector representations for quaternion random variables, we choose the real representation from here. So a quaternion random variable q is completely described by the joint distribution of its four components, *i.e.* by the distribution of the vector: $\hat{\mathbf{q}} = [q_0 \ q_1 \ q_2 \ q_3]^T$. So, from now we consider the following:

$$\hat{\mathbf{y}} = g(w\hat{\mathbf{x}}) \quad (4.1)$$

Then, back to our problem, the distribution of the output y can be linked to the distribution of the input x the following way:

$$p(\hat{\mathbf{y}}) = \frac{p(\hat{\mathbf{x}})}{|\mathbf{J}|} \Big|_{\hat{\mathbf{x}}=w^{-1}g^{-1}(\hat{\mathbf{y}})} \quad (4.2)$$

where the terms of the Jacobian matrix $\mathbf{J}$ is given as $\mathbf{J}|_{ij} = \partial g_i / \partial x_j$ for $i, j = 0, 1, 2, 3$. We give the Jacobian expression for the three types of nonlinearities, namely the $\mathbb{H}$ -split ($\mathbf{J}_{Hs}$), $\mathbb{C}$ -split ($\mathbf{J}_{Cs}$) and $\mathbb{H}$ -full ($\mathbf{J}_{Hf}$):

$$\mathbf{J}_{Hs} = \begin{pmatrix} \partial_0 g_0 & 0 & 0 & 0 \\ 0 & \partial_1 g_1 & 0 & 0 \\ 0 & 0 & \partial_2 g_2 & 0 \\ 0 & 0 & 0 & \partial_3 g_3 \end{pmatrix}; \mathbf{J}_{Cs} = \begin{pmatrix} \partial_0 g_0 & \partial_1 g_0 & 0 & 0 \\ \partial_0 g_1 & \partial_1 g_1 & 0 & 0 \\ 0 & 0 & \partial_2 g_2 & \partial_3 g_2 \\ 0 & 0 & \partial_2 g_3 & \partial_3 g_3 \end{pmatrix}; \mathbf{J}_{Hf} = \begin{pmatrix} \partial_0 g_0 & \partial_1 g_0 & \partial_2 g_0 & \partial_3 g_0 \\ \partial_0 g_1 & \partial_1 g_1 & \partial_2 g_1 & \partial_3 g_1 \\ \partial_0 g_2 & \partial_1 g_2 & \partial_2 g_2 & \partial_3 g_2 \\ \partial_0 g_3 & \partial_1 g_3 & \partial_2 g_3 & \partial_3 g_3 \end{pmatrix} \quad (4.3)$$

where ∂_α stands for $\partial / \partial x_\alpha$. The pattern of the Jacobian matrices allows to conclude that only the $\mathbb{H}$ -full approach can take into account the largest set of possible relationships between components. More specifically, in the case of a $\mathbb{C}$ -proper r.v., the first component is correlated with the third and fourth, which induces off-diagonal terms in the covariance matrix (see Amblard, P.O. & Le Bihan N. (2004) for details). As a consequence, only the $\mathbb{H}$ -full nonlinearity in quaternionic Infomax can perform a thorough recovering of a $\mathbb{C}$ -proper r.v.. We illustrate this now on simulated signals.

5. Simulation results

We consider a simple example where two vector-sensors record a linear and instantaneous mixture of two random signals. The proposed model for the recorded mixture is: $\mathbf{x}[m] = \mathbf{A}\mathbf{s}[m]$, where $\mathbf{x}[m] \in \mathbb{H}^2$, $\mathbf{s}[m] \in \mathbb{H}^2$

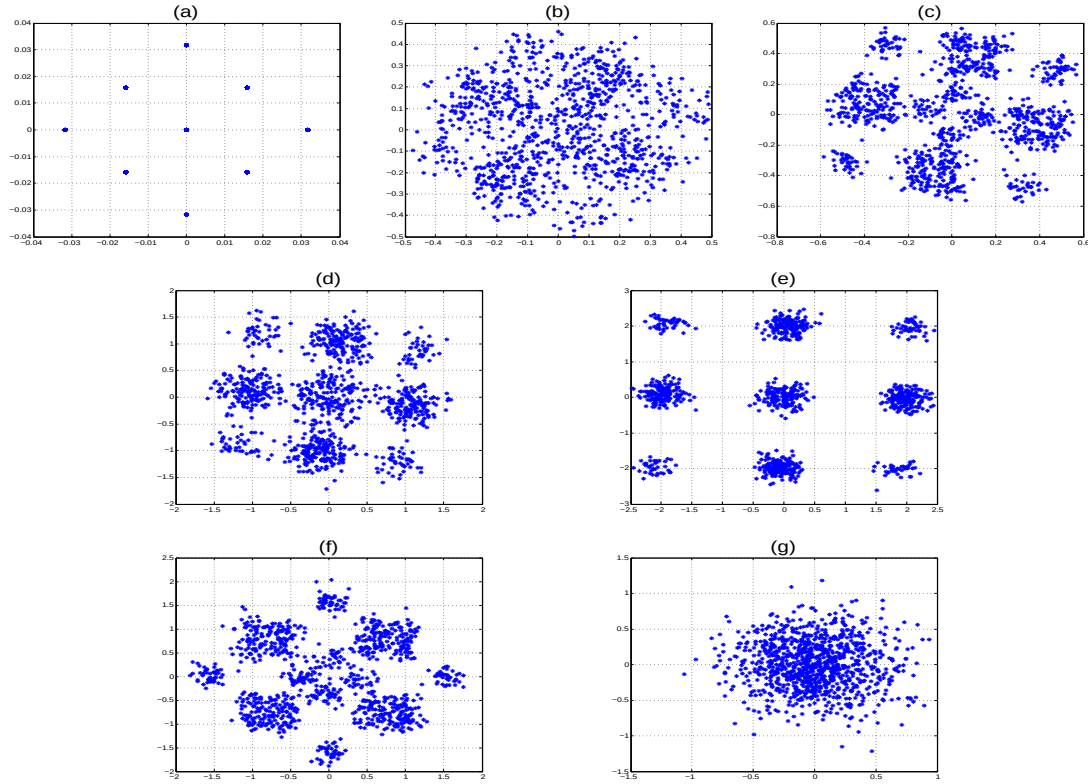


FIGURE 1. Original signal (a), mixture recorded on sensor 1 (b) and 2 (c), estimated signal with $\mathbb{C}$ -split tanh (d) and $\mathbb{H}$ -full (e) nonlinearities, estimated $\mathbb{H}$ -proper noise with $\mathbb{C}$ -split tanh (f) and $\mathbb{H}$ -full (g) nonlinearities

and $\mathbf{A} \in \mathbb{H}^{2 \times 2}$. Source $s_1[m]$ is a $\mathbb{H}$ -proper signal (non polarized, assumed as noise), Gaussian and i.i.d.. Signal $s_2[m]$ is a random signal taking values on the edges of a 4D polytope (see Zetterberg, L. H. & Brändström (1977) for details). The Quaternionic Infomax results are presented in Figure 1. One can see that the $\mathbb{C}$ -split approach is not able to separate the signal from noise while the $\mathbb{H}$ -full approach lead to a better estimation of the signal and furthermore to the recovery of the $\mathbb{H}$ -properness property for the noise.

6. Conclusion

We have proposed an extension of the Infomax algorithm to the quaternionic signals case. The choice of the non-linearity has been demonstrated to be determinant in the separation result. Despite the lack of $\mathbb{H}$ -differentiability, $\mathbb{H}$ -full nonlinearities are the best choice to achieve the separation between polarized signals and noise. The proposed algorithm could be of interest in applications such as Optics, Electromagnetism or Seismic where polarized signals corrupted by noise are encountered and where ICA can help to recover the wavefield sources.

REFERENCES

- ADALI, T. & KIM, T. & CALHOUN, V. 2004 Independent component analysis by complex nonlinearities, *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* Montreal, Quebec, Canada.
- AMBLARD, P.O. & LE BIHAN N. 2004 On properness of quaternion valued random variables, *IMA Conference on Mathematics in Signal Processing* Cirencester, UK.
- BELL, A.J., & SEJNOWSKI, T.J. 1995 An information maximisation approach to blind separation and blind deconvolution, *Neural Computation* Vol. 7, No. 6, pp. 1129–1159
- BROSSEAU, C. 1998 Fundamentals of Polarized Light: A Statistical Approach, *John Wiley and Sons*.

- BUCHHOLZ, S. & LE BIHAN N. 2006 Optimal Separation of Polarized Signals by Quaternionic Neural Networks, *European Signal and Image Processing Conference (EUSIPCO)* Florencia, Italy.
- CALHOUN, V. & ADALI, T. 2002 Complex infomax: convergence and approximation with complex nonlinearities, *IEEE Workshop on Neural Networks for Signal Processing (NNSP)* Martigny, Switzerland.
- KARLSSON, M. & PETERSSON M. 2004 Quaternion approach to PMD and PDL phenomena in optical fiber systems, *IEEE Journal of Lightwave technology* Vol. 22, No. 4, pp. 1137–1146.
- SUDBERY, A. 1979 Quaternionic analysis, *Math. Proc. Camb. Phil. Soc.* Vol. 85, pp. 199–225.
- VAKHANIA, N.N. 1998 Random vectors with values in quaternions Hilbert spaces, *Th. Probab. Appl.* Vol. 43, No. 1, pp. 99–115.
- WARD, J.P. 1997 Quaternions and Cayley numbers, *Kluwer Academic Publisher*.
- ZETTERBERG, L. H. & BRÄNDSTRÖM, H. 1977 Codes for combined phase and amplitude modulated signals in a four-dimensional space *IEEE Trans. on Communications* Vol. 25, No. 9, pp. 943–950.

Deuxième partie

Signaux à valeurs sur $SO(3)$ et $\mathcal{S}^2$

Signaux à valeurs sur $SO(3)$ et $\mathcal{S}^2$

Sommaire

3.1	Polarisation, sphère de Poincaré et $SO(3)$	99
3.2	Variables aléatoires sur $SO(3)$	101
3.2.1	Théorème de Peter-Weyl	102
3.2.2	Fonctions caractéristiques	103
3.3	Processus de Lévy sur $SO(3)$	104
3.3.1	Dépolarisation	104
3.3.2	Diffusion multiple et Processus de Poisson composé sur $SO(3)$	105
3.4	Phase géométrique des ondes polarisées	109
3.4.1	Phase géométrique non-adiabatique : mise en évidence expérimentale pour les ondes élastiques	111
3.4.2	Phase géométrique des ondes élastiques en diffusion multiple	113
3.5	Conclusion	115
3.6	Publications annexées en lien avec ce chapitre	116
3.6.1	"Higher-order statistics of Stokes parameters in a random birefringent medium" Waves in random and Complex media 2008	117
3.6.2	"Decompounding on compact Lie groups" IEEE Transactions on Information Theory 2010	136
3.6.3	"Non-adiabatic geometric phase of elastic waves" Soumis à Journal of the Acoustical Society of America 2011	148

Dans ce chapitre, nous présentons quelques résultats concernant des signaux à valeurs sur les groupes de Lie et les variétés, en particulier la sphère unité dans $\mathbb{R}^3$, *i.e.* $\mathcal{S}^2$, et le groupe des rotations $SO(3)$. Les résultats présentés s'appuient sur la théorie de la représentation, les processus aléatoires à valeurs sur les variétés et quelques résultats de géométrie différentielle. Une des préoccupations centrale des travaux qui sont exposés ici est le traitement des signaux polarisés.

Les travaux présentés dans ce chapitre concernant les processus de Lévy à valeurs sur les groupes de Lie ont été menés dans le cadre de la thèse de Salem SAID. Les résultats concernant la phase géométrique sont liés aux travaux menés actuellement dans la thèse de Jérémie BOULANGER.

3.1 Polarisation, sphère de Poincaré et $SO(3)$

La polarisation des ondes est une propriété partagée par les ondes électromagnétiques et élastiques. Dans ce qui suit, nous nous intéresserons particulièrement aux ondes dont la polarisation est confinée au plan orthogonal à leur direction de propagation, ce qui est le cas pour la lumière et les ondes élastiques de cisaillement (ondes S) [Aki 2002].

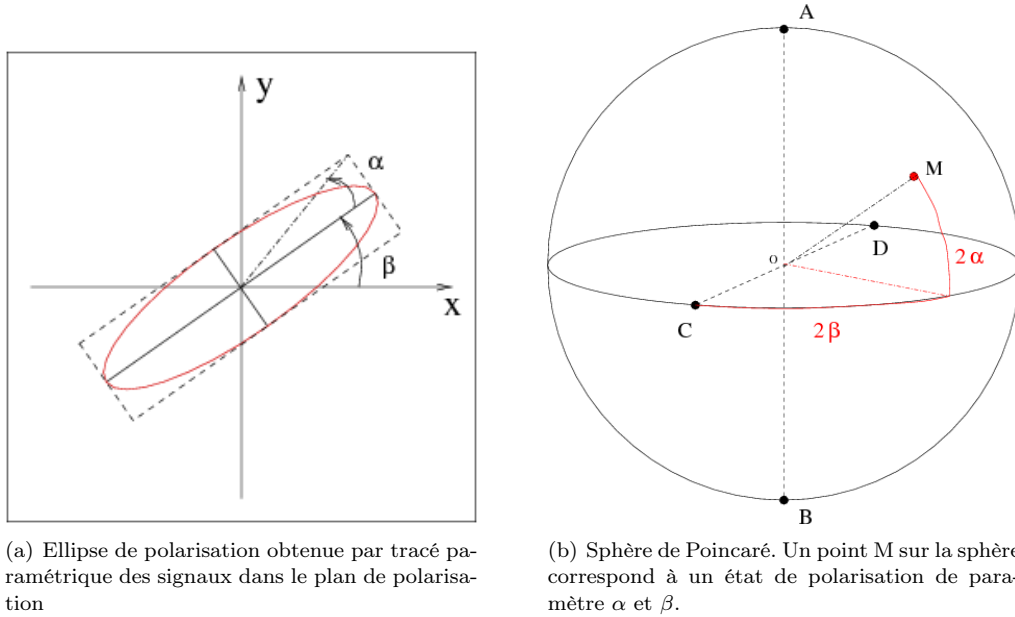


FIGURE 3.1 – Représentation de Poincaré pour la polarisation des ondes. À gauche : dans le plan de polarisation, le tracé au cours du temps d’une composante en fonction de l’autre (un signal en fonction de l’autre), à une fréquence donnée, donne la figure elliptique rouge, appelée parfois *ellipse de polarisation*. Il faut deux paramètres α et β pour décrire cette ellipse. À droite : la représentation de Poincaré assigne à une ellipse de polarisation un point sur la sphère unité S^2 . Un état de polarisation *pur* est repéré par ses deux paramètres, correspondant alors à son azimuth et son élévation.

Comme expliqué au chapitre 2, on décrit une onde polarisée par ses deux composantes dans le plan orthogonal à la direction de propagation¹. Les deux composantes sont des signaux complexes. La polarisation est la relation entre ces deux signaux, et peut donc être décrite par deux grandeurs : une amplitude et une phase. Ces deux paramètres décrivent l’ellipse de polarisation. On représente classiquement une onde totalement polarisée par un point sur la sphère de Poincaré [Brosseau 1998], comme expliqué sur la figure 3.1.

Lors de la propagation dans un milieu, il est possible que la polarisation soit affectée : par la PMD (Polarization Mode Dispersion) ou la PDL (Polarization Dispersion Loss) dans les fibres optiques, par la biréfringence (différence de vitesse de propagation des composantes) ou par la diffusion multiple dans les milieux diffusants (nuages, sang, milieux désordonnés, croûte terrestre, etc.)². Par exemple, le long d’une fibre optique, l’onde peut être affectée par de la biréfringence aléatoire [Gordon 2000]. L’aspect aléatoire induit par le milieu lors de la propagation fait qu’en sortie, l’onde se retrouve *partiellement* polarisée, voire complètement *dépolarisée*. La polarisation partielle se caractérise par non plus un point sur la sphère, mais une distribution. Cette distribution devient uniforme sur S^2 lorsque l’onde est *dépolarisée*. Sur la figure 3.2, on montre schématiquement le passage d’une polarisation totale à un

1. On rappelle que la polarisation n’est définie qu’à une fréquence commune donnée pour les deux signaux/composantes évoluant dans le plan de polarisation.

2. Ces phénomènes linéaires pouvant affecter la polarisation sont les plus courants observés. Il existe évidemment des phénomènes non-linéaires pouvant apparaître dans certains milieux [Brosseau 1998], mais nous ne les considérons pas ici et nous restreignons au cas linéaire.

état de dépolarisation presque total. Ce genre de transition peut apparaître à cause de la biréfringence aléatoire dans une fibre optique par exemple [Yang 2001].

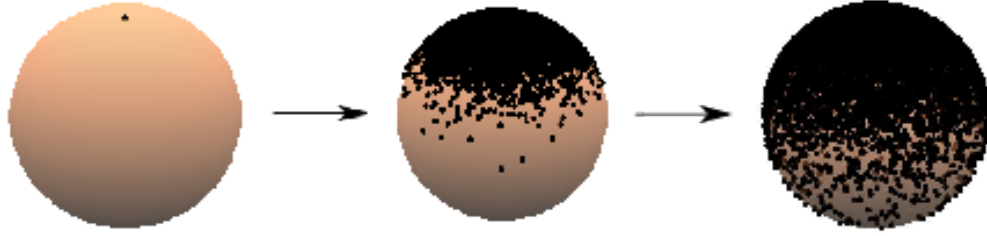


FIGURE 3.2 – Passage d’un état totalement polarisé (gauche) vers un état partiellement polarisé (milieu) puis un état pratiquement dépolarisé (droite) sur la sphère de Poincaré.

Une description statistique est donc nécessaire pour le phénomène de dépolarisation. Notons l’état de polarisation d’une onde à une distance d de l’origine, dans un milieu donné, par $\mathbf{S}(d)$. Nous avons proposé dans [Said 2008a] (article inclus dans la section 3.6) de considérer la polarisation en entrée $\mathbf{S}(d=0)$ et à une position z , $\mathbf{S}(d=z)$ d’un milieu biréfringent comme une variable aléatoire sur la sphère de Poincaré. L’effet du milieu est alors une rotation aléatoire. Cela veut dire que l’effet du milieu biréfringent est l’action d’une variable aléatoire du groupe $SO(3)$ sur $\mathbf{S}(0)$. Ainsi, on peut exprimer la polarisation de l’onde à une certaine distance dans le milieu, $\mathbf{S}(z)$, comme :

$$\mathbf{S}(z) = \mathbf{R}(z)\mathbf{S}(0)$$

avec $\mathbf{R}(z) \in SO(3)$ représentant l’action du milieu sur une distance z . Dans les problèmes que nous aborderons dans la suite, la position z pourra être confondue avec le temps t , en normalisant par la vitesse de propagation et en considérant une propagation “proche”³ de la ligne droite. On pourra ainsi écrire avec le même modèle :

$$\mathbf{S}(t) = \mathbf{R}(t)\mathbf{S}(0)$$

C’est ce que nous ferons dans la Section 3.3.2. On peut donc voir l’évolution de la polarisation comme un processus à valeurs sur la sphère, *i.e.* considérer $\mathbf{S}(z) \in \mathcal{S}^2$, ou un processus à valeurs sur $SO(3)$, *i.e.* considérer $\mathbf{R}(z)$. Ces deux visions sont équivalentes du fait de l’action transitive du groupe $SO(3)$ sur l’espace homogène $\mathcal{S}^2$. On peut noter que ce modèle d’action du canal de propagation est **multiplicatif**, mais également **non-commutatif**. Il est donc nécessaire d’utiliser des outils d’analyse harmonique non-commutative pour décrire les variables et processus à valeurs sur les groupes non-commutatifs comme $SO(3)$. Nous présentons quelques résultats dans la Section 3.2 afin de pouvoir étudier les processus qui nous intéressent dans la Section 3.3.

3.2 Variables aléatoires sur $SO(3)$

Il existe de nombreux ouvrages traitant de l’Analyse Harmonique Non-Commutative (AHNC). Pour la partie théorique, nous conseillons le livre de Taylor [Taylor 1986] et le

3. Cette restriction sur les variations autour de la ligne droite se comprennent naturellement dans le cas d’une propagation dans un guide comme la fibre optique. En milieu aléatoire, ce régime correspond à de la diffusion vers l’avant.

cours de Dieudonné [Dieudonné 1980]. Le livre de Chirikjian donne un nombre d'exemples important d'applications de l'AHNC en robotique ou mécanique [Chirikjian 2000]. Récemment, il est à noter qu'une activité importante de recherche s'est également développée dans la communauté signal autour de l'utilisation de l'AHNC, par exemple sur le développement d'algorithmes rapides de Transformation de Fourier (TF) sur $SO(3)$ [Kostelec 2008] ou d'analyse en ondelettes sur la sphère [Bogdanova 2005]. Pour ce qui est des variables aléatoires sur les espaces "non-standards", on pourra consulter [Grenander 1963].

Pour ce qui nous intéresse, nous reprenons quelques résultats d'AHNC, en particulier le théorème de Peter-Weyl. Le cas général des groupes de lie compacts est traité dans [Dieudonné 1980]. Nous présentons uniquement ici le cas du groupe des rotations $SO(3)$.

3.2.1 Théorème de Peter-Weyl

On note $L^2(SO(3), \mathbb{C})$ l'espace de Hilbert des fonctions de carré sommable sur $SO(3)$. La mesure de Haar sur $SO(3)$ est notée $d\mu(g)$ avec $g \in SO(3)$. En utilisant la convention ZYZ des angles d'Euler⁴ pour $SO(3)$, on a la formule explicite $d\mu(\varphi, \theta, \psi) = \frac{1}{8\pi^2} \sin \theta d\varphi d\theta d\psi$. On note $Irr(G)$ l'ensemble des classes d'équivalence des représentations irréductibles de G . C'est un ensemble dénombrable. Maintenant si $l \in Irr(G)$, toutes les représentations de la classe l sont de même dimension et il existe une représentation unitaire U^l dans cette classe. En choisissant cette représentation, on peut supposer que $U^l : G \rightarrow SU(\mathbb{C}^{d_l})$, avec $SU(\mathbb{C}^{d_l})$ le groupe spécial unitaire⁵ de dimension d_l [Dieudonné 1980].

Dans le cas où $G = SO(3)$, le théorème de Peter-Weyl assure qu'il existe des fonctions $d_l^{1/2} U_{mn}^l$, pour $l \in Irr(SO(3))$ et $m, n = 1, \dots, d_l$, qui forment une base orthonormée de $L^2(SO(3), \mathbb{C})$. Dans le cas de $SO(3)$, $Irr(SO(3)) = \mathbb{N}$ et donc $d^l = 2l + 1$.

Soit $f : SO(3) \rightarrow \mathbb{R}$ une fonction de carré sommable, $f \in L^2(SO(3), \mathbb{R})$, par rapport à la mesure de Haar. Le théorème de Peter-Weyl nous permet de donner une série de Fourier de f :

$$f(g) = \sum_{l \geq 0} \sum_{|m,n| \leq l} (2l+1) A_{mn}^l U_{mn}^l(g)$$

avec $g \in SO(3)$ et les $U_{mn}^l(g)$ donnés dans le cas de la convention ZYZ pour $SO(3)$ par :

$$U_{mn}^l(\varphi, \theta, \psi) = e^{-im\varphi} d_{mn}^l(\cos \theta) e^{-in\psi}$$

où les fonctions $d_{mn}^l(\cos \theta)$ sont les fonctions de Wigner-d pouvant s'exprimer via les polynômes de Jacobi [Dieudonné 1980]. Les coefficients de Fourier A_{mn}^l sont donnés par la formule inverse :

$$A_{mn}^l = \frac{1}{8\pi^2} \int_0^{2\pi} \int_0^\pi \int_0^{2\pi} p(\varphi, \theta, \psi) e^{im\varphi} d_{mn}^l(\cos \theta) e^{in\psi} \sin \theta d\varphi d\theta d\psi$$

Les coefficients A_{mn}^l peuvent être mis sous forme matricielle $\mathbf{A}^l$, avec pour chaque degré $l \geq 0$ une matrice de dimension $(2l+1) \times (2l+1)$. On peut également utiliser la notation matricielle pour les éléments de la base de $SO(3)$, *i.e.* $\mathbf{U}^l$. La décomposition de Fourier devient :

$$f(g) = \sum_{l \geq 0} (2l+1) Tr \left(\mathbf{A}^l (\mathbf{U}^l(g))^\dagger \right)$$

avec $Tr(\cdot)$ l'opérateur trace.

4. On note $\varphi \in [0, 2\pi[$, $\theta \in [0, \pi[$ et $\psi \in [0, 2\pi[$ les trois angles d'Euler.

5. Le groupe spécial unitaire $SU(\mathbb{C}^{d_l})$ est l'ensemble des matrices complexes de taille $d_l \times d_l$ qui sont unitaires, $MM^\dagger = I_{d_l}$ et de déterminant égal à 1.

Une application directe du théorème de Peter-Weyl est le théorème de convolution. Soit $f_1, f_2 \in L^2(SO(3), \mathbb{R})$, deux fonctions de $SO(3)$. Leur convolution est également une fonction de $SO(3)$ donnée par :

$$w(g) = (f_1 * f_2)(g) = \int_{SO(3)} f_1(h) f_2(h^{-1}g) dh$$

où l'on voit l'analogie avec la convolution sur $\mathbb{R}$. Une conséquence du théorème de Peter-Weyl est que cette convolution est équivalente à :

$$\mathbf{w}^l = \mathbf{f}_1^l \mathbf{f}_2^l \quad \forall l$$

Ce qui veut dire que les coefficients de Fourier de w sont obtenus par produits matriciels des coefficients de Fourier de f_1 et f_2 . On retrouve bien la transformation d'un produit de convolution en un produit dans le domaine de Fourier, à ceci près qu'ici c'est un produit matriciel pour chaque degré l , avec une taille croissante des coefficients avec l .

Des résultats similaires se retrouvent pour les fonctions sur $\mathcal{S}^2$, via l'action transitive de $SO(3)$ sur $\mathcal{S}^2$. Les fonctions de base sur $\mathcal{S}^2$ sont les harmoniques sphériques $Y_m^l(\varphi, \theta)$ [Dieudonné 1980], les coefficients sont vectoriels, et le produit de convolution entre une fonction de $SO(3)$ et une fonction de $\mathcal{S}^2$ est une fonction de $\mathcal{S}^2$ dont les coefficients de Fourier vectoriels sont obtenus par produit matrice-vecteur pour chaque l . Pour plus de détails, on consultera par exemple [Chirikjian 2000].

L'analyse harmonique non-commutative permet une description "fréquentielle" des fonctions à valeurs sur $SO(3)$ ou $\mathcal{S}^2$. La différence majeure avec le cas classique connu en traitement du signal est que, dans le cas de l'AHNC, les coefficients de Fourier sont matriciels et leur dimension est croissante avec l'ordre (équivalent des hautes fréquences dans le cas commutatif). Notons que dans le cas de $SO(3)$ et $\mathcal{S}^2$, la compacité du support fait que l'analyse harmonique donne une série harmonique infinie. C'est d'ailleurs le cas pour les groupes de Lie matriciels compacts d'une manière générale. Pour les groupes non-compacts comme $SE(3)$ par exemple, l'analyse de Fourier est plus complexe [Chirikjian 2000]. Par la suite, nous utiliserons les résultats d'AHNC principalement pour l'étude des distributions de variables aléatoires sur $SO(3)$ et $\mathcal{S}^2$. En particulier, nous ferons usage des fonctions caractéristiques.

3.2.2 Fonctions caractéristiques

On considère des variables aléatoires à valeurs sur $SO(3)$. Ces variables sont définies sur un espace probabilisé adéquat $(\Omega, \mathcal{A}, \mathbb{P})$. On note $\mathbb{E}$ l'espérance mathématique sur cet espace. Étant donnée une variable aléatoire X sur $SO(3)$, sa fonction caractéristique $\Phi_X(l)$ est :

$$\Phi_X(l) = \mathbb{E}[\mathbf{U}^l(X)] = \int_{SO(3)} p_X(g) \mathbf{U}^l(g) d\mu(g)$$

pour tout l . Parmi les propriétés de ces fonctions caractéristiques (listées dans [Said 2010]), nous en donnons une qui est assez remarquable et utile par la suite. Soit X et Y deux variables aléatoires indépendantes à valeurs sur $SO(3)$. Alors, $Z = XY$ a pour fonction caractéristique :

$$\Phi_Z(l) = \Phi_X(l) \Phi_Y(l) \quad \forall l$$

On voit ici que dans le produit de variables aléatoires sur $SO(3)$, les densités sont convoluées et donc les fonctions caractéristiques multipliées entre elles (avec importance de l'ordre dans le produit). C'est en fait l'équivalent de la somme de variables aléatoires dans le cas scalaire commutatif classique. Dans la suite, nous montrerons comment nous avons

utilisé ces fonctions caractéristiques sur $SO(3)$ pour proposer une méthode d'estimation non-paramétrique des fonctions de phase des milieux aléatoires [Said 2010] (article inclus dans la section 3.6) basée sur la résolution du problème de “decompounding” qui sera abordé dans la section 3.3.

3.3 Processus de Lévy sur $SO(3)$

Nous adoptons ici un modèle de processus de Lévy pour décrire des phénomènes de dépolarisation et de diffusion multiple. Pour ces deux phénomènes, nous utilisons les deux composantes d'un processus de Lévy : le mouvement Brownien et le processus de Poisson composé [Applebaum 2000]. Avant de montrer comment utiliser ces deux processus particuliers dans l'étude de la polarisation, nous rappelons les propriétés d'un processus de Lévy sur $SO(3)$. Une étude complète est disponible dans la littérature [Liao 2004a, Applebaum 2000].

On dit que le processus $\mathbf{S}(t) = \mathbf{S}_t$, $t \geq 0$, défini sur $(\Omega, \mathcal{A}, \mathbb{P})$ et à trajectoire *càdlàg*,⁶ est un processus de Lévy (à gauche⁷) sur $SO(3)$ si :

- Il est à accroissements indépendants : pour tout suite strictement croissante $t_0, t_1, t_2, \dots, t_N$, les $\mathbf{S}_{t_1}^{-1}\mathbf{S}_{t_0}$, $\mathbf{S}_{t_2}^{-1}\mathbf{S}_{t_1}$, $\dots$, $\mathbf{S}_{t_N}^{-1}\mathbf{S}_{t_{N-1}}$ sont indépendants.
- Ses accroissements sont stationnaires : $\mathbf{S}_{t+\tau}^{-1}\mathbf{S}_t = \mathbf{S}_\tau$, $\forall t$.
- Il vérifie la propriété de continuité stochastique : $\forall \varepsilon, \lim_{\tau \rightarrow 0} \mathbb{P}(\|\mathbf{S}_{t+\tau}^{-1}\mathbf{S}_t\| \geq \varepsilon) = 0$.

On rajoute en général la condition que $\mathbf{S}_0 = \mathbf{I}_3$, avec $\mathbf{I}_3$ la matrice identité dans $\mathbb{R}^{3 \times 3}$. Pour plus de détails sur les processus de Lévy on consultera [Cont 2003], et plus particulièrement [Liao 2004b] pour les processus de Lévy sur les groupes de Lie.

Dans la suite, nous allons nous intéresser à deux cas particuliers de processus de Lévy : les processus de diffusion purs (Brownien) et les processus de sauts purs (Poisson composé). Ces deux processus vont être utilisés pour modéliser le comportement d'ondes polarisées. Il est du coup intéressant de comprendre ce que veulent dire physiquement les conditions que nous venons de lister. Utiliser les processus de Lévy induit donc des contraintes de modélisation physiques du milieu dans lequel se propage les ondes polarisées :

- **Accroissements indépendants** : deux parties du milieu de propagation qui ne se chevauchent pas sont indépendantes.
- **Accroissements stationnaires** : le milieu est statistiquement homogène.
- **Continuité stochastique** : une partie du milieu de petite taille ne peut pas engendrer un “grand” changement.

Avec ce choix de modèle, nous allons maintenant voir comment il est possible d'utiliser des processus de Lévy pour étudier le comportement d'ondes polarisées, ou même non-polarisées, dans des milieux “aléatoires”.

3.3.1 Dépolarisation

Si l'on revient à l'étude de la propagation d'une onde polarisée dans une fibre optique, on a l'état de polarisation à une distance z dans la fibre $\mathbf{S}(z) = \mathbf{S}_z$ qui est donné par :

$$\mathbf{S}_z = \mathbf{R}_z \mathbf{S}_0$$

L'indépendance des accroissements et les résultats sur la convolution présentés plus haut permettent de donner une expression de la densité de $\mathbf{S}_z$ (voir [Said 2008a]), *i.e.* de prédire le comportement de la polarisation en fonction de la distance parcourue. De plus, comme

6. *càdlàg* : continu à droite et limité à gauche

7. On définit le processus de Lévy à droite de la même manière, simplement en définissant les incréments à droite. Cette distinction est possible du fait de la non-commutativité du produit.

nous l'avons démontré dans [Said 2008a], le modèle de processus de Lévy permet de prédire toute quantité moyenne de $\mathbf{S}_z$, *i.e.* les quantités de la forme $\mathbb{E}[f(\mathbf{S}_z)]$, avec f une fonction de carré intégrable sur $\mathcal{S}^2$. Ainsi, on a une expression de la forme :

$$\mathbb{E}[f(\mathbf{S}_z)] = \sum_{l \geq 0} (2l+1) \hat{\mathbf{f}}^l e^{\hat{\mathbf{t}}^l z} \xi^l(0)$$

avec $\hat{\mathbf{f}}^l$ les coefficients de Fourier de f , $\xi^l(0)$ les coefficients de Fourier de $\mathbf{S}_0$ et $\hat{\mathbf{t}}^l$ des coefficients (appelés “générateurs du processus $\mathbf{R}_z$ ”) directement déterminés par le canal de propagation (voir [Said 2008a] pour les détails). On peut alors considérer toute sorte de fonction f , et en particulier celles définissant l'entropie de l'état de polarisation ou des quantités liées aux paramètres de l'ellipse de polarisation. Cette approche basée sur les processus de Lévy donne donc un cadre théorique simple pour l'étude du phénomène de dépolarisation.

De plus, nous avons défini dans [Said 2008a] une généralisation du *degré de polarisation*. Classiquement, cette quantité est basé sur des moments d'ordre 2 [Brosseau 1998] des observables. À l'aide de l'AHNC, nous avons proposé la définition de degrés de polarisation d'ordre supérieur basée sur les coefficients de Fourier de la distribution de polarisation sur la sphère de Poincaré. Un tel degré de polarisation peut être utile dans le cas de distributions non-gaussiennes sur la sphère de Poincaré (voir les références données dans [Said 2008a]). Ce degré de polarisation, noté P^l , est défini à tous les ordres l . Il est obtenu par la norme des coefficients de la décomposition de Fourier :

$$P^l = \|\xi^l\| = \|\mathbb{E}[Y^l(\mathbf{S})]\|$$

avec $\mathbf{S}$ la variable à valeurs sur $\mathcal{S}^2$ et Y^l les harmoniques sphériques (voir [Said 2008a] encore pour les détails). Par exemple, un état de polarisation sera pur quand $P^l = 1 \forall l$. Avec le modèle de processus de Lévy que nous avons introduit, on peut également prédire l'évolution de la polarisation lors d'une propagation dans un milieu aléatoire caractérisé par $\mathbf{R}_z$. On a directement la relation :

$$P^l(z) = \|e^{\hat{\mathbf{t}}^l z} \xi^l(0)\|$$

qui nous donne l'évolution des degrés de polarisations à tous les ordres en fonction de la distance parcourue z . Sachant que $\lim_{z \rightarrow 0} e^{\hat{\mathbf{t}}^l z} = 0$ ⁸, on voit que l'effet aléatoire de $\mathbf{R}_z$ tend inexorablement à dépolariser l'onde, et ce à tous les ordres l .

L'utilisation du modèle de processus de Lévy sur $SO(3)$ combinée à de l'AHNC sur la sphère $\mathcal{S}^2$ permet, dans un formalisme original, de définir un degré de polarisation d'ordre supérieur et de décrire l'évolution de ce degré de polarisation lors d'une propagation dans un milieu dont l'action est aléatoire. L'intérêt de notre modèle est qu'il est générique et ne fait pas de supposition sur la distribution des $\mathbf{R}_z$, qui peuvent être Gaussiens ou non. En particulier, ce sont les actions de milieux non-Gaussiens qui devraient tirer avantage de ce formalisme que nous avons proposé pour l'étude de la polarisation en milieu aléatoire.

Nous allons à présent nous intéresser à un type particulier de processus de Lévy pour la description de la propagation en milieu aléatoire, les processus de Poisson composés.

3.3.2 Diffusion multiple et Processus de Poisson composé sur $SO(3)$

À présent, nous nous intéressons à la propagation d'une onde dans un milieu diffusant. On ne considère plus une onde polarisée ici. Nous retrouverons la polarisation et la diffusion

8. Comme expliqué dans [Said 2008a], cette limite vaut 0 du fait que les valeurs propres de $\hat{\mathbf{t}}^l$ ont une partie réelle négative.

multiple dans l'étude de la phase géométrique dans la section 3.4. La diffusion multiple est ici modélisée par un processus de sauts pur sur $SO(3)$.

Nous avons montré dans [Le Bihan 2009, Said 2010] que l'on peut considérer le processus de diffusion des ondes dans un milieu aléatoire comme un processus de Poisson composé sur le groupe des rotations. Le milieu de propagation est composé d'une matrice homogène dans laquelle des inclusions sphériques de propriétés physiques différentes (densité, vitesse de propagation) sont présentes à des positions inconnues. On considère une répartition homogène de ces inclusions. Le modèle que nous avons proposé consiste à étudier le comportement aléatoire de la direction de propagation d'une onde lors de son cheminement dans un tel milieu. Les grandeurs caractéristiques qui nous intéressent sont la *direction de propagation* $\mu(t) \in \mathcal{S}^2$ (un vecteur unitaire de $\mathbb{R}^3$), le *libre parcours moyen* ℓ et la *fonction de phase* des diffuseurs. On ne s'intéresse pas ici à l'amplitude des ondes et les chocs que subit l'onde avec les diffuseurs sont supposés parfaitement élastiques. Le libre parcours moyen est la distance moyenne entre deux événements de diffusion et la fonction de phase décrit comment un diffuseur modifie (aléatoirement) la direction de propagation. La figure 3.3 schématise le modèle que nous considérons.

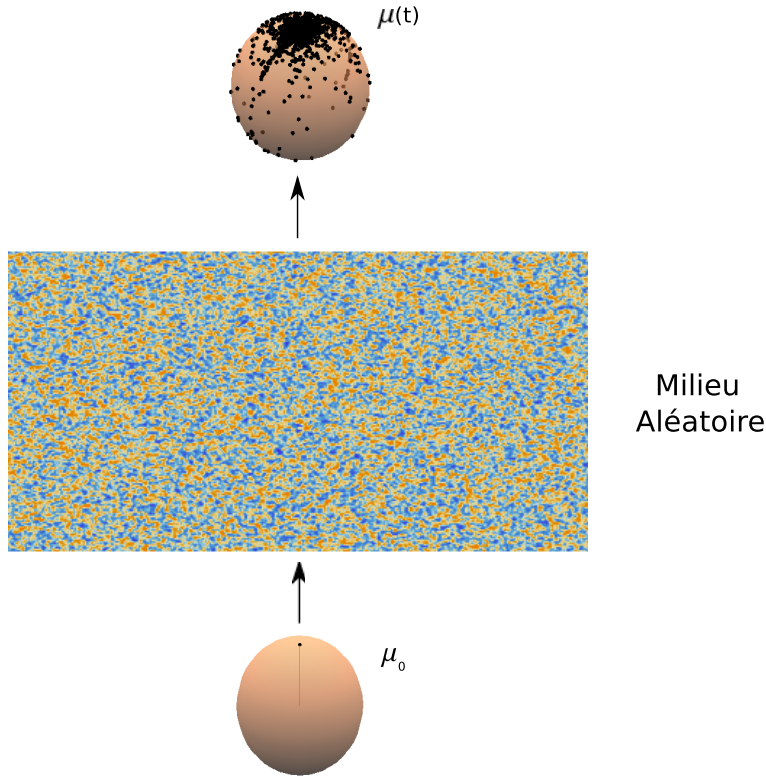


FIGURE 3.3 – Modélisation de la propagation en milieu aléatoire. La direction incidente de l'onde est μ_0 . Après propagation dans le milieu aléatoire, la distribution de la direction de propagation $\mu(t)$ est distribuée aléatoirement sur $\mathcal{S}^2$.

Supposant l'entrée d'une onde plane dans un milieu aléatoire bidimensionnel (cette restriction est obtenue par hypothèse de symétrie de la diffusion, valable pour le cas des diffuseurs parfaitement sphériques et isotropes), nous avons proposé le modèle suivant pour la direction de propagation. Après un temps de propagation t , et en supposant que la direction de propagation en entrée du milieu aléatoire est μ_0 , la direction de propagation $\mu(t)$ peut s'écrire :

$$\mu(t) = \mathbf{R}(t)\mu_0 = \prod_{n=0}^{N(t)} \mathbf{r}_n \mu_0 = \mathbf{r}_{N(t)} \dots \mathbf{r}_1 \mathbf{r}_0 \mu_0$$

avec $\mu(t), \mu_0 \in \mathcal{S}^2$. Les $\mathbf{r}_n$ sont des rotations aléatoires dues aux diffuseurs. Ce sont donc des variables aléatoires à valeurs sur $SO(3)$, *i.e.* $\mathbf{r}_n \in SO(3)$. Elles sont caractérisées par leur densité, également appelée *fonction de phase*, $p_{\mathbf{r}}$. $N(t)$ est un processus de Poisson, de paramètre $\lambda = \tau^{-1}$, où τ est lié au libre parcours moyen : $\tau = \ell/c$, avec c la célérité de l'onde. Clairement, $N(t)$ compte le nombre de diffusions qui sont survenues dans le temps t . $\mu(t)$ est donc un processus à valeurs sur la sphère $\mathcal{S}^2$. $\mathbf{R}(t)$ est un processus de Poisson composé sur $SO(3)$, et par l'action transitive de $SO(3)$ sur $\mathcal{S}^2$, il agit sur la direction de propagation. Les $\mathbf{r}_n$ sont supposés *i.i.d.* et indépendants de $N(t)$.

L'utilisation des processus de Poisson composés pour la description de la diffusion multiple a été introduite dans [Ning 1995], mais uniquement pour estimer la distribution de sortie de l'angle de diffusion. Nous avons proposé d'utiliser ce modèle à des fins d'estimation des paramètres du milieu aléatoire. En effet, à l'aide des résultats d'AHNC et sur les fonctions caractéristiques sur $SO(3)$ présentés dans les sections 3.2.1 et 3.2.2, il est possible de relier la fonction caractéristique de $\mu(t)$ et celle des $\mathbf{r}_n$ de la façon suivante :

$$\phi_{\mu(t)}(l) = \phi_{\mathbf{R}(t)}(l) \phi_{\mu_0}(l) = \exp(\lambda t (\phi_{\mathbf{r}}(l) - I_{(2l+1)})) \phi_{\mu_0}(l)$$

avec $\exp$ l'exponentielle de matrices, $I_{(2l+1)}$ la matrice identité de dimension $(2l+1) \times (2l+1)$, et $\phi_{\mathbf{r}}(l)$ la fonction caractéristique commune des $\mathbf{r}_n$, *i.e.* la TF de la fonction de phase $p_{\mathbf{r}_n}$. On voit donc qu'il existe une relation simple entre l'expression de la fonction caractéristique de $\mu(t)$ et celle des $\mathbf{r}_n$. C'est l'intérêt du modèle basé sur le processus de Poisson composé.

Ceci suggère la possible résolution du problème de **decompounding** : connaissant λ , donner une estimée de $p_{\mathbf{r}}$ à l'aide d'observations de $\mu(t)$. Nous avons étudié de problème de decompounding dans [Said 2010], et présenté sa résolution (définition et propriétés des estimateurs, convergence). Connaissant μ_0 et le paramètre de Poisson λ , on peut ainsi, à l'aide d'un échantillon $[\mu_1, \mu_2, \dots, \mu_M]$ d'angles de diffusion en sortie du milieu, inverser l'expression entre les fonctions caractéristiques et estimer $p_{\mathbf{r}_n}$ par "Transformée de Fourier inverse". Les détails se trouvent dans [Said 2010]. C'est en fait une approche non-paramétrique, équivalente à une méthode des moments sur $SO(3)$.

En combinant cette approche avec une paramétrisation de la *fonction de phase*, il est possible d'estimer cette fonction de phase. Un modèle simple et assez couramment utilisé en physique pour la *fonction de phase* de diffuseurs est le modèle d'Henry-Greenstein. Dans ce modèle, $p_{\mathbf{r}}$ est de la forme :

$$p_{\mathbf{r}}(\cos \theta) = \frac{1 - g^2}{(1 + g^2 - 2g \cos \theta)^{\frac{3}{2}}} = \sum_{l=0}^{+\infty} (2l+1) \hat{p}^l P^l(\cos \theta)$$

avec g le paramètre d'anisotropie ou asymétrie, et θ l'angle de diffusion. On remarque qu'un seul angle intervient dans cette *fonction de phase*, alors que d'une manière très générale, on pourrait avoir trois angles. Une conséquence directe est que les $U^l(\varphi, \theta, \phi)$ introduits plus haut deviennent ici des polynômes de Legendre, *i.e.* des $P^l(\cos \theta)$, conduisant à l'expression de $p_{\mathbf{r}}$ donnée plus haut. Les coefficients de Fourier $\hat{p}^l$ sont alors à valeurs scalaires. De plus,

la *fonction de phase* Henyey-Greenstein est uniquement paramétrée par g , l'anisotropie. On peut alors aisément donner un estimateur de g , via un échantillon de mesures d'angles de diffusion à un instant T , qui a est simplement :

$$\tilde{g} = \left(\frac{\ell}{cT} \log \left(\frac{1}{M} \sum_{m=1}^M P^l(\mu_m) \right) \right)^{1/l}$$

Il est également possible d'obtenir une estimation simultanée de g et λ avec les deux premiers coefficients de Legendre estimés. L'intérêt du modèle de type Poisson composé est qu'il permet, sous réserve d'une modélisation simple des diffuseurs (modèle Henyey-Greenstein ou autre à faible nombre de paramètres), d'obtenir une estimation des caractéristiques de ces diffuseurs à partir de la distribution de l'angle de diffusion en sortie (*i.e.* la distribution des $\mu(t)$). On a donc une résolution possible du problème inverse. Plus de détails sur la technique de **decompounding** sur $SO(3)$ (et d'une manière plus générale sur les groupes de Lie matriciels compacts) se trouvent dans [Said 2010]. Un exemple d'estimation du paramètre g est donné sur la figure 3.4.

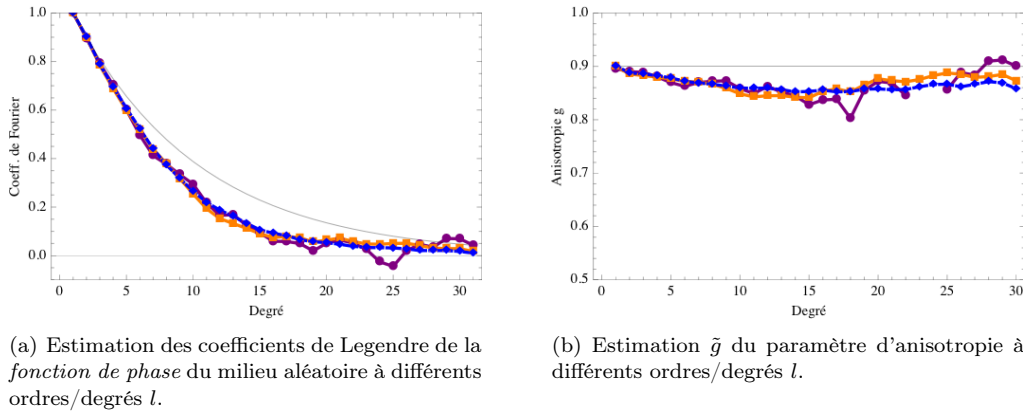


FIGURE 3.4 – Estimation des coefficients de Legendre de la *fonction de phase* et du coefficient d'anisotropie g par la technique de decompounding. Les valeurs théoriques sont en traits pleins gris. Les courbes estimées sont obtenues pour des tailles d'échantillons de : $5 \cdot 10^2$ points (violet), $5 \cdot 10^3$ points (jaune) et $5 \cdot 10^4$ points (bleu)

La technique de decompounding étant basée sur les fonctions caractéristiques, il est possible d'obtenir de manière non-paramétrique une estimée de la fonction caractéristique des diffuseurs (*i.e.* la TF de la *fonction de phase*). Sur la figure 3.4(a), on voit les premiers coefficients estimés de la décomposition de Legendre de la *fonction de phase*. Cette simulation montre un cas où l'on a choisi une distribution de Henyey-Greenstein. D'autres fonctions ont été testées avec ce modèle [Le Bihan 2009].

Nous avons étudié dans [Le Bihan 2009] les limites du modèle Poisson composé pour décrire la diffusion multiple. On peut montrer sa validité dans le régime de diffusion vers l'avant, mais il n'est pas possible de décrire la rétro-diffusion par exemple. Sur les figures 3.5(a) et 3.5(b) (reprises de [Le Bihan 2009]), on peut voir la comparaison entre la distribution des angles de diffusion prédite par le modèle Poisson composé et les simulations Monte Carlo de la diffusion, pour deux types de *fonctions de phase* (Henyey-Greenstein et Gaussienne). Le modèle Poisson composé se comporte bien pour les angles de diffusion proches de 0 (μ proche de 1), mais ne décrit pas correctement la diffusion sur à des angles

importants (μ proche de 0). Il est donc important de noter que ce modèle de processus permet uniquement d'étudier les phénomènes de diffusion vers l'avant.

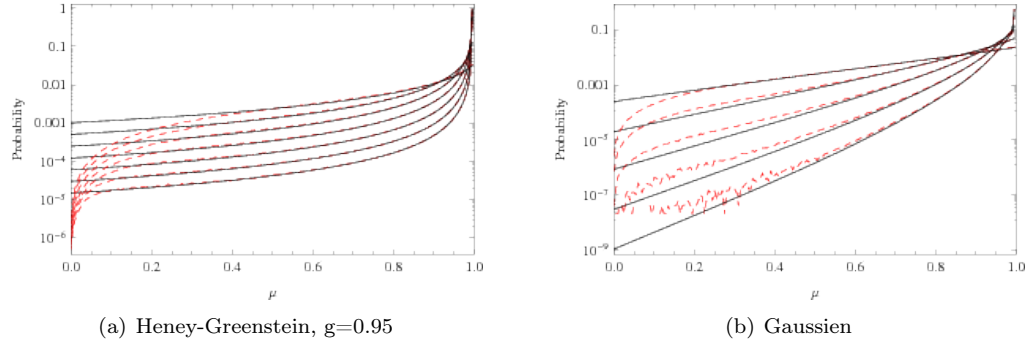


FIGURE 3.5 – Comparaison entre la fonction de phase prédite par le modèle de Poisson composé (trait noir) et la fonction de phase calculée par simulation Monte Carlo de la diffusion multiple (trait rouge), pour deux types de *fonction de phases* (Henyey-Greenstein et Gaussien). Les courbes sont données pour différentes profondeurs de propagation. On note une bonne adéquation entre la fonction de phase prédite et celle obtenue par simulation MC, excepté pour les angles de diffusion proches de $\pi/2$, *i.e.* $\mu \rightarrow 0$.

L'étude des phénomènes de diffusion multiple à l'aide des processus de Poisson est donc intéressante à des fins d'inférence sur le milieu aléatoire. À l'aide d'un modèle relativement simple de processus, il est possible de prédire le comportement d'une onde dans un milieu aléatoire, *i.e.* la distribution de l'angle de diffusion $\mu(t)$. En particulier, ce modèle autorise l'identification de la *fonction de phase* des diffuseurs, sous réserve, dans le cas général, de la connaissance du libre parcours moyen.

On peut imaginer plusieurs extensions de ce modèle pour l'étude de la diffusion multiple des ondes. Les suites de ces travaux en cours sont détaillées dans le chapitre 4. Parmi les extensions déjà explorées, celle incluant la polarisation des ondes dans le modèle a permis d'obtenir déjà de nouveaux résultats, et en particulier la description de l'apparition d'une phase géométrique dans les milieux aléatoires pour les ondes polarisées, présentées dans la Section 3.4.2.

3.4 Phase géométrique des ondes polarisées

La phase géométrique est un concept universel en physique. Depuis sa découverte par Sir M. Berry au début des années 80 [Berry 1984] en mécanique quantique (phase de Berry) et par Hannay en mécanique classique (angle de Hannay), il a été démontré que de nombreux autres systèmes physiques possèdent une phase géométrique [Wilczek 1989]. L'exemple le plus connu en mécanique classique est sans doute le pendule de Foucault [von Bergman 2007]. La phase géométrique traduit la différence de configuration d'un système entre son état initial et son état final lorsque ce système a effectué un chemin fermé dans l'espace des états. Dans le cas du pendule de Foucault, l'espace des états est la position sur terre (relativement au repère fixe des étoiles) et la configuration est la direction d'oscillation du pendule. Il est connu depuis le milieu du XIX^{ème} siècle qu'après un jour sidéral (une rotation complète de la terre, soit 23h57min), le pendule de Foucault n'oscille plus dans la même direction. À la latitude de Paris (48°52' Nord), le pendule fait un angle

de -271° avec sa direction d'oscillation initiale au bout d'un jour. Cet angle entre les deux directions d'oscillation est une *phase géométrique*. Le calcul de cet angle est possible *via* les lois de la mécanique Newtonienne en considérant que le pendule est soumis à l'attraction terrestre et la force de Coriolis, mais également à l'aide d'outils de géométrie différentielle (connexion de Levi-Civita, théorème de Gauss-Bonnet local) [Carmo 1976] dont le concept central est le *transport parallèle* [von Bergman 2007, Oprea 1995, Faure 2011].



FIGURE 3.6 – Directions d'oscillation du pendule de Foucault initiale (rouge) et direction d'oscillation après un jour (vert) dans le plan tangent à la terre. La trajectoire suivie par le plan tangent est représentée en jaune.

L'interprétation du pendule de Foucault par le point de vue géométrique réside dans le fait que le vecteur qui représente la direction d'oscillation du pendule est transporté parallèlement (il suit la connexion de Levi-Civita) dans l'espace tangent à la sphère (terre), *i.e.* c'est un élément de TS^2 .

Ce transport parallèle est induit par le principe d'inertie (façon la moins coûteuse en énergie de passer d'un plan d'oscillation à un autre au cours de la rotation de la terre). L'espace des états est la surface de la terre, et le pendule de Foucault effectue un chemin fermé sur cette sphère au cours d'un jour sidéral (Fig. 3.6). Quand il revient à sa position initiale (par rapport aux étoiles) après un jour sidéral, le vecteur oscillation a subi une rotation de -271° . Cette valeur est donnée par l'holonomie [Carmo 1976] de la trajectoire sur la sphère terrestre⁹. Cette notion de transport parallèle d'un élément de l'espace tangent est à la base de la théorie des phases géométriques [Faure 2011, Wilczek 1989].

Ici, nous allons nous intéresser à la phase géométrique des ondes polarisées. Pour ces ondes, la phase géométrique peut apparaître de deux façons différentes :

- *i)* Lorsqu'une onde dans un état de polarisation pur est amenée à changer d'état de polarisation (changement de position sur la sphère de Poincaré) au moyen de polariseurs. Après un enchaînement de changements de polarisation qui ramènent la

9. L'holonomie est une intégrale de courbure. Plus de détails sont donnés dans [Carmo 1976] ou [Faure 2011] sur ce concept.

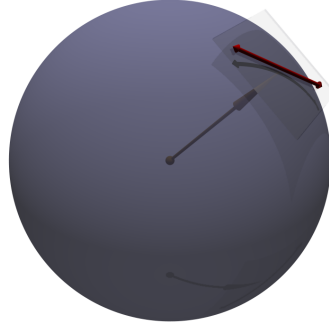


FIGURE 3.7 – Modèle pour l'étude d'une onde polarisée. Le vecteur *direction de propagation* (jaune) est un élément de $\mathcal{S}^2$ et le vecteur polarisation (rouge) est un élément de $T\mathcal{S}^2$

polarisation de l'onde à son état initial, le chemin couvert sur la sphère de Poincaré entoure un angle solide. L'onde polarisée finale est alors déphasée par rapport à l'onde initiale et le déphasage est égal à la moitié de l'angle solide entouré par le chemin. Cette phase géométrique est connue sous le nom de *phase de Pancharatnam* et sa mise en évidence expérimentale date de 1988 [Bhandari 1988, Chyba 1988].

- *ii)* Lorsqu'une onde polarisée se propage suivant une trajectoire tridimensionnelle, la polarisation étant confinée au plan orthogonal à la direction de propagation, elle sera transportée parallèlement lors de la propagation. La direction de propagation étant un vecteur unitaire de $\mathbb{R}^3$, c'est un élément de $\mathcal{S}^2$. La polarisation est contrainte au plan tangent à $\mathcal{S}^2$, *i.e.* elle évolue dans $T\mathcal{S}^2$. Le théorème de Gauss-Bonnet local [Carmo 1976] prédit donc que si la trajectoire suivie par l'onde est "fermée" (point de départ et d'arrivée identiques), la polarisation aura subi une rotation égale à l'aire entourée par la trajectoire.

C'est cette dernière phase qui nous intéresse ici. La phase géométrique pour les ondes polarisées est connue depuis plusieurs années. Elle a été prédite théoriquement par Berry [Berry 1984] (et pour le cas non-adiabatique par Segert [Segert 1987]) et la mise en évidence expérimentale dans le cas adiabatique est due à Tomita et Chiao en 1986 [Tomita 1986]. Nous avons mis en évidence expérimentalement l'existence d'une même phase pour les ondes élastiques, et ce dans un régime non-adiabatique.

3.4.1 Phase géométrique non-adiabatique : mise en évidence expérimentale pour les ondes élastiques

Nous avons mis en place récemment une expérience pour démontrer l'existence de la phase géométrique pour les ondes élastiques [Boulanger 2011] (cet article est inclus dans la section 3.6). En fait, les ondes de cisaillement S sont de même nature que les ondes électromagnétiques : leur polarisation est confinée au plan orthogonal à la direction de propagation. On peut donc appréhender la propagation d'une onde polarisée en observant le trajet du vecteur représentant sa direction de propagation, *i.e.* un élément de $\mathcal{S}^2$ (vecteur jaune sur la figure 3.7), ainsi que la polarisation de l'onde qui évolue dans le plan tangent

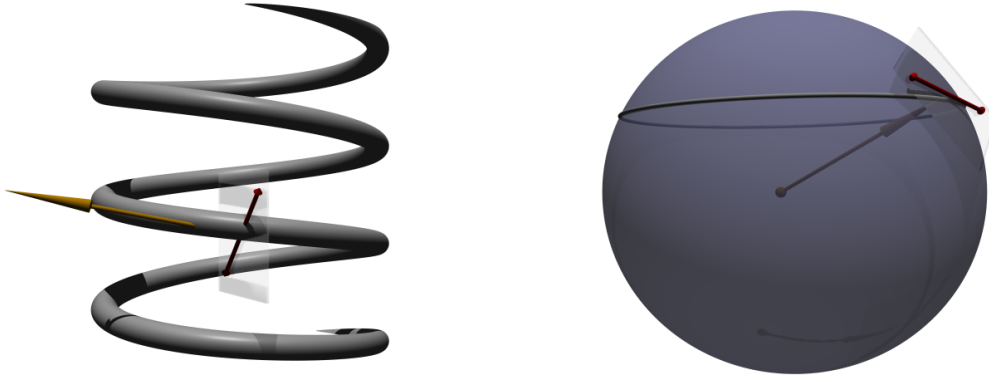


FIGURE 3.8 – Correspondance entre expérience et modèle. À gauche : Lors de la propagation dans le ressort, la direction de propagation (vecteur **jaune**) est tangente à la trajectoire (ressort). La polarisation (vecteur **rouge**) est dans le plan tangent à cette direction de propagation. À droite : Le trajet du vecteur de propagation est en fait le long d'un parallèle de S^2 (courbe grise) situé à une latitude sur S^2 égale à l'angle du ressort. La polarisation, élément de TS^2 , est transportée parallèlement le long de la trajectoire sur la sphère.

TS^2 (vecteur rouge sur la figure 3.7).

Afin de mesurer la phase géométrique pour les ondes S, nous avons fait propager des ondes de flexion dans un guide de section circulaire et à géométrie hélicoïdale : un ressort en métal. La symétrie du ressort est telle que lorsque l'onde le parcourt, cela équivaut pour le vecteur de direction de propagation sur S^2 à effectuer un chemin sur un parallèle de latitude égale à l'angle d'inclinaison de l'hélice/ressort¹⁰. La polarisation étant transportée parallèlement, elle doit donc tourner proportionnellement à la distance de propagation¹¹. Cette relation de proportionnalité avait été observée pour la lumière en faisant propager un faisceau laser dans une fibre optique enroulée autour d'un cylindre de bois [Tomita 1986]. L'expérience que nous avons mis en place a permis de voir le même comportement pour les ondes élastiques, avec une différence majeure tout de même : le caractère **non-adiabatique**.

Les ondes de flexion que nous avons considérées avaient des longueurs d'onde de l'ordre de plusieurs centimètres, et la longueur de propagation était de l'ordre de 1,30m. C'est totalement hors du régime adiabatique¹². Nous avons donc mis en évidence lors de cette expérience la phase géométrique des ondes élastiques hors du régime adiabatique. Cette phase géométrique avait été prédite théoriquement [Aharonov 1987, Anandan 1988, Segert 1987], mais n'avait jamais été observée jusque là.

Les détails de l'expérience se trouvent dans [Boulanger 2011]. Sur les figures 3.9(a)

10. L'angle d'inclinaison de l'hélice est l'angle que fait le vecteur tangent à l'hélice avec la direction horizontale.

11. Les détails du calcul du coefficient de proportionnalité sont donnés dans [Boulanger 2011].

12. En termes classiques, le régime adiabatique revient à avoir des oscillations **rapides** dans le plan tangent TS^2 par rapport à la vitesse de déplacement de ce plan tangent le long de la trajectoire sur S^2 . C'est le cas pour le pendule de Foucault dont les oscillations sont nettement plus rapides que la vitesse de rotation de la terre.

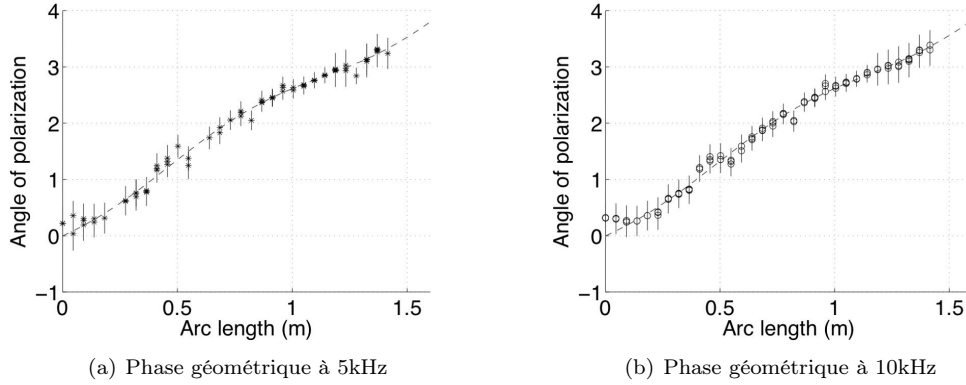


FIGURE 3.9 – Relation entre angle de rotation de la polarisation et distance parcourue par l'onde le long de l'hélice/ressort. La relation est linéaire en première approximation avec une pente indépendante de la fréquence (2.38 rad.m^{-1} à 5kHz et 2.40 rad.m^{-1} à 10kHz) et d'une valeur proche de la valeur prédite théoriquement (2.49 rad.m^{-1}). Voir [Boulanger 2011] pour plus de détails.

et 3.9(b), on présente la phase géométrique (relation de proportionnalité entre distance parcourue et rotation de la direction de polarisation) estimée lors de l'expérience pour des ondes de flexion à 5kHz et 10kHz . L'indépendance en fréquence de la phase est un élément spécifique au caractère purement géométrique de cette phase.

Lors de cette expérience, le chemin suivi par l'onde est tridimensionnel, condition *sine qua none* pour l'apparition d'une phase géométrique, mais déterministe car complètement imposé par la forme du guide. Il peut être intéressant de voir ce qui se passe quand la trajectoire devient aléatoire. C'est ce que nous avons essayé de faire en étudiant l'apparition de la phase géométrique dans le régime de diffusion multiple.

3.4.2 Phase géométrique des ondes élastiques en diffusion multiple

Dans le cas de la lumière, l'apparition d'une phase géométrique lors de la propagation dans un milieu aléatoire a été prédite et observée [Maggs 2001, Rossetto 2002]. Nous avons adopté une approche basée sur les processus aléatoires sur $SO(3)$ pour appréhender cette phase géométrique pour les ondes élastiques [Boulanger 2010].

Il est possible d'adapter le modèle de processus de Poisson sur $SO(3)$ au cas des ondes polarisées. Il faut pour cela considérer la direction de propagation et la polarisation (qui est dans le plan orthogonal à la direction de propagation) simultanément. Pour étudier la diffusion multiple des ondes polarisées, il faut alors considérer le repère $\mathcal{F}$ défini par :

$$\mathcal{F} = [\mathbf{v}, \mathbf{z} \wedge \mathbf{v}, \mathbf{z}]$$

avec $\mathbf{z}$ la direction de propagation et $\mathbf{v}$ le vecteur de polarisation. $\mathcal{F}$ est un triplet de vecteurs¹³ unitaires qui représente l'onde polarisée. Ce modèle est assez logique et intuitif au vu de la figure 3.7. Ainsi, quand l'onde traverse un milieu aléatoire comme décrit dans la section 3.3.2, l'action d'un diffuseur est encore une rotation $\mathbf{r}_{\psi,\theta,\varphi}$. Mais cette rotation est particu-

13. il est nécessaire de considérer $\mathbf{z} \wedge \mathbf{v}$ afin de pouvoir repérer la rotation de la polarisation dans le plan tangent.

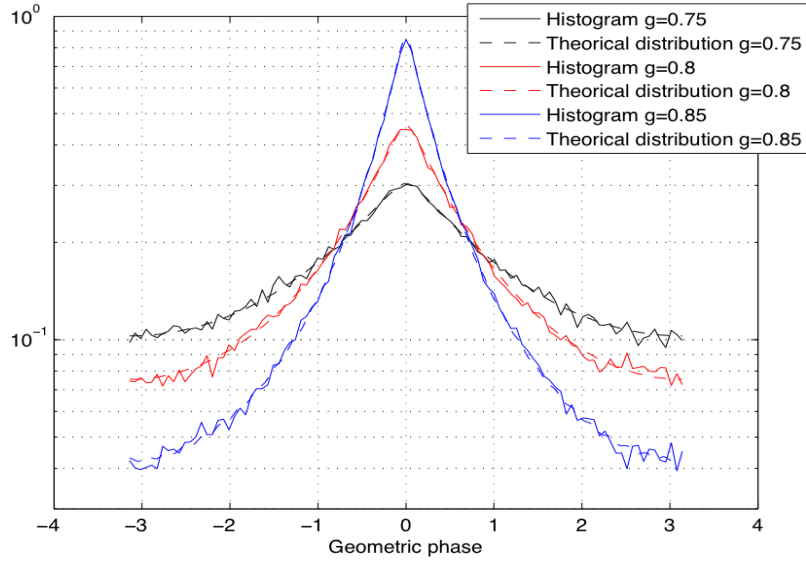


FIGURE 3.10 – Distribution de la phase géométrique obtenue par le modèle processus de Poisson composé avec transport parallèle (traits pointillés) et par simulation Monte Carlo (traits pleins). Les différentes valeurs de g montrent que pour des diffuseurs faiblement anisotropes (g faible) la distribution de phase géométrique s’uniformise rapidement, tandis que pour des diffuseurs fortement anisotropes ($g \rightarrow 1$) la distribution de phase géométrique s’uniformise plus lentement. C’est une marque de l’effet de dépolariation.

lière si la polarisation est transportée parallèlement¹⁴. Nous avons montré [Boulangier 2010] que la contrainte de transport parallèle induit que pour $\mathbf{r}_{\psi, \theta, \varphi}$, on a $\psi = -\varphi$. Cela induit que les actions aléatoires des diffuseurs sont des rotations “contraintes” à deux degrés de liberté (θ et ψ). Cette contrainte entraîne qu’il est impossible d’écrire le processus de diffusion multiple avec transport parallèle comme un processus de Poisson composé (à accroissements indépendants) à gauche. Par contre, le résultat très intéressant que nous avons obtenu est qu’il est possible d’écrire ce processus comme un processus de Poisson composé **à droite** [Boulangier 2010].

Ainsi, après un temps de propagation t , l’onde polarisée $\mathcal{F}_t$ est donnée par :

$$\mathcal{F}_t = \mathcal{F}_0 \prod_{n=0}^{N(t)} \mathbf{r}_{\psi_n, \theta_n} = \mathcal{F}_0 \mathbf{r}_{\psi_1, \theta_1} \mathbf{r}_{\psi_2, \theta_2} \cdots \mathbf{r}_{\psi_{N(t)}, \theta_{N(t)}}$$

qui est donc un processus de Lévy à droite sur $SO(3)$, avec $\mathcal{F}_0$ l’onde polarisée initialement. En utilisant à nouveau le théorème de Peter-Weyl et des techniques d’AHNC, on peut obtenir (voir [Boulangier 2010] pour les détails de calcul) la distribution de la phase géométrique après un temps de propagation t :

$$p(\theta, \phi_{geo}) = 2 \sum_{m \geq 0} \sum_{l \geq m} (2l+1) e^{\lambda(\hat{p}_{\mathbf{r}}^l m, m-1)t} d_{m,m}^l(\theta) \cos(m\phi_{geo})$$

14. Physiquement, le régime dans lequel il y a un transport parallèle de la polarisation est le régime dit de l’optique géométrique, i.e. quand la taille des diffuseurs est grande devant la longueur d’onde [Rossetto 2002]. Nous assumons que nous sommes dans ce régime.

avec ϕ_{geo} la phase géométrique, λ le paramètre de Poisson et $\hat{p}_{\mathbf{r},m,m}^l$ les coefficients de Fourier de la *fonction de phase* des diffuseurs. Comme dans le cas de la diffusion multiple des ondes non polarisées présenté dans la section 3.3.2, l'obtention d'une expression semi-analytique pour la distribution de la phase géométrique permet de prévoir son comportement dans un milieu aléatoire, et permet d'envisager une étape d'estimation (problème inverse) autorisant l'estimation de paramètre du milieu *via* la distribution de la phase géométrique. Cette distribution permet également de prédire la vitesse de dépolarisation des ondes dans un régime de diffusion multiple.

Sur la figure 3.10, on montre la distribution de phase géométrique prédite par le processus de Poisson avec transport parallèle (pointillés) et la distribution obtenue par simulations Monte Carlo de la distribution (traits pleins). Plusieurs distributions sont données pour des valeurs d'anisotropie (g dans le modèle Henyey-Greenstein) des diffuseurs différentes. On remarque une très bonne adéquation de la prédiction du modèle de type processus avec les simulations Monte Carlo.

Ces premiers résultats sur la distribution de la phase géométrique montrent l'intérêt d'utiliser les processus de Poisson sur le groupe des rotations pour étudier le comportement des ondes polarisées. L'étape d'estimation est pour le moment une des perspectives de ces travaux. Ces perspectives sont exposées dans le chapitre 4.

3.5 Conclusion

Nous avons utilisé des outils de théorie des groupes et de géométrie différentielle pour l'étude de signaux à valeurs sur $SO(3)$ et $\mathcal{S}^2$. Nous avons proposé une approche fondée sur les processus de Lévy sur le groupe $SO(3)$ pour modéliser et étudier des phénomènes comme la dépolarisation, la diffusion multiple des ondes (polarisées ou non) et la phase géométrique des ondes élastiques (apparaissant dans les guides d'ondes à géométrie déterministe ou dans les milieux aléatoires). Tous ces modèles ont été validés sur simulations numériques et nous avons également proposé une expérience pour l'observation de la phase géométrique des ondes élastiques dans un régime non-adiabatique. Une telle observation n'avait jamais été faite auparavant. L'intérêt des processus de Poisson composés gauches et droits sur $SO(3)$ est qu'ils permettent une résolution du problème inverse *via* la technique de decompounding. Cette approche pourrait être intéressante dans un cadre expérimental de diffusion multiple afin d'estimer des paramètres du milieu de diffusion par exemple.

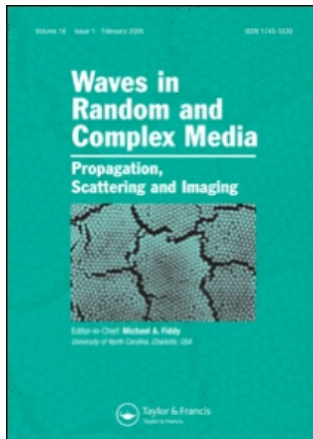
Plus généralement, les travaux présentés dans ce chapitre montrent l'utilité des processus de Lévy pour la description des ondes polarisées, et en même temps ils ouvrent la voie à l'étude de processus plus généraux : les processus non-holonomes et les processus de Markov sur $SO(3)$ et $\mathcal{S}^2$. Les processus non-holonomes qui, après une trajectoire close dans l'espace des états, ne se retrouvent pas dans la même configuration, pourraient être utiles à la compréhension plus fine des phénomènes d'apparition de phases géométriques. Une perspective de ces travaux est donc d'étudier ces processus dans le cadre plus large de l'étude des chaînes de Markov sur $SO(3)$.

3.6 Publications annexées en lien avec ce chapitre

Les articles suivants sont inclus ici :

1. “Higher-order statistics of Stokes parameters in a random birefringent medium”, S. Salem and N. Le Bihan, **Waves in random and complex media**, Vol. 18, No. 2, pp. 275 – 292, 2008.
2. “Decompounding on compact Lie groups”, S. Said, C. Lageman, N. Le Bihan and J.H. Manton, **IEEE Transactions on Information Theory**, Vol. 56, No. 6, pp. 2766-2777, 2010.
3. “Non-adiabatic geometric phase of elastic waves”, J. Boulanger, N. Le Bihan, S. Catheline and V. Rossetto, soumis à **Journal of the Acoustical Society of America**, 2011 (<http://arxiv.org/abs/1103.4506>).

This article was downloaded by:[Australian National University Library]
On: 1 May 2008
Access Details: [subscription number 773444842]
Publisher: Taylor & Francis
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Waves in Random and Complex Media

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t716100762>

Higher-order statistics of Stokes parameters in a random birefringent medium

Salem Said ^a; Nicolas Le Bihan ^a

^a Gipsa-Lab, Département DIS, ENSIEG, Domaine Universitaire, Saint Martin d'Hères Cedex, France

Online Publication Date: 01 May 2008

To cite this Article: Said, Salem and Bihan, Nicolas Le (2008) 'Higher-order statistics of Stokes parameters in a random birefringent medium', Waves in Random and Complex Media, 18:2, 275 — 292

To link to this article: DOI: 10.1080/17455030701762636
URL: <http://dx.doi.org/10.1080/17455030701762636>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Higher-order statistics of Stokes parameters in a random birefringent medium

Salem Said* and Nicolas Le Bihan

*Gipsa-Lab, Département DIS, ENSIEG, Domaine Universitaire,
Saint Martin d'Hères Cedex, France*

(Received 17 July 2007; final version received 20 October 2007)

We present a new model for the propagation of polarized light in a random birefringent medium. This model is based on a decomposition of the higher-order statistics of the reduced Stokes parameters along the irreducible representations of the rotation group. We show how this model allows a detailed description of the propagation, giving analytical expressions for the probability densities of the Mueller matrix and the reduced Stokes vector throughout the propagation. It also allows an exact description of the evolution of averaged quantities, such as the degree of polarization. We also discuss how this model allows a generalization of the concepts of reduced Stokes parameters and degree of polarization to higher-order statistics. We give some notes on how it can be extended to more general random media.

1. Introduction

The interaction of polarized light with a random medium is of interest in many domains. Examples include imagery, telecommunications, medicine and instrumentation. In this paper, we present a new model for the propagation of a polarized lightwave in a random birefringent medium. This problem is relevant to telecommunications. Indeed, it occurs in optical fibres subject to polarization mode dispersion (PMD); see for instance [1–3]. The review [4] gives a comprehensive introduction to PMD. The approach offered in [1, 2] is based on stochastic differential equations satisfied by the Jones (or Mueller) matrix of a certain length of the medium.

Our main aim in this paper is to introduce a new model for the propagation of polarized light in a birefringent random medium. All the main features of this model are generalizable to random media in general. However, we use this simpler situation as a starting point. Unlike the ones given in [1, 2], the model we present is global and not local. This means that instead of using a stochastic differential equation that describes propagation over short ‘infinitesimal’ distances, we concentrate on the linear operators describing propagation over finite distances. The two approaches are complementary. However, the results described here are more far-reaching than the ones obtained using stochastic differential equations.

One of the main features of this model is that it considers the random medium as a whole. In other words, the random medium is considered as a system and a certain type of statistical relation between its input and output is studied. The microscopic characteristics of the medium are thus related to the parameters of the probability laws involved in the input–output relation

*Corresponding author. Email: said@gipsa-lab.inpg.fr

characterizing the medium (system). They can be estimated given samples of the input and output of the medium. In fact, the whole probability density of the Mueller matrix of the medium can, in principle, be estimated given samples of the input and output.¹

The proposed model is based on a decomposition of the higher-order statistics of the reduced Stokes parameters along the irreducible representations of the rotation group. The rotation group is important because a birefringent medium acts on the Poincaré sphere by rotations [5]. The relation between the probability densities, on the Poincaré sphere, of the input and output reduced Stokes vectors will be shown to be given by a spherical convolution. The decomposition along irreducible representations of the rotation group is used to obtain from this relation the laws of transformation of the higher-order statistics of the reduced Stokes vector. On the one hand, these laws of transformation constitute generalizations of the Stokes–Mueller formalism to higher-order statistics. On the other hand, they will be shown to be a powerful tool for describing the propagation of polarized light in a random medium. This whole approach is actually an example of harmonic analysis on the rotation group [6] which is an instance of non-commutative harmonic analysis [6–8].

Classical models for the state of polarization of a lightwave, such as the Jones and Stokes models, only involve statistics of order 1 and 2 [5]. Whenever the fluctuations of the field of the lightwave are non-Gaussian, it is necessary to consider higher-order statistics. Non-Gaussian fields occur in many situations in optics [9]. Different ways of including higher-order statistics have been proposed. For instance, Réfrégier in [10, 11] gives a measure of the degree of polarization based on Kullback relative entropy between the field of the lightwave and a completely isotropic field with the same intensity distribution. In [12], Luis proposes a measure of the degree of polarization based on the mean quadratic distance between the probability density function on the Poincaré sphere, corresponding to the state of polarization of the lightwave, and the probability density function of a uniform distribution on the Poincaré sphere. In [9] Ellis and Dogariu propose using the correlations of the Stokes parameters to discriminate states of polarization with the same second order statistics. In [13], they use the symmetries of the probability distribution of the Stokes parameters on the Poincaré sphere to make this same distinction. All these models emphasize two aspects: that it is important to include higher-order statistics, and that this should be done in accordance with the symmetries of the probability distribution on the Poincaré sphere. Our approach is quite similar to them. It has the additional advantage of formalizing the importance of the spherical symmetry underlying the problem, by using the irreducible representation of the rotation group.

The plan of the paper is the following: In Section 2 we outline the physical situations we wish to consider. In Section 3, we explain the mathematical tools we will be using, namely the irreducible representations of the rotation group. In Section 4 we give the main equations of our model. In Section 5, we discuss how it generalizes the Stokes–Mueller formalism to higher-order statistics. In Section 6 we show how it can be used to describe the propagation of a lightwave in a random birefringent medium. In Section 7, we discuss how our model can be extended to the case of a general random medium and conclude.

2. Physical problem

Let us quickly outline the kind of physical setup we are interested in. We consider the effect of a random birefringent medium on the state of polarization of an incident lightwave. We will be interested in two kinds of relations. The first kind of relations is between the input and output

¹This discussion is illustrated by the example given in Section 6.3.

states of polarization when the medium is taken as a whole. The second kind is the relations governing the changes in the state of polarization of the lightwave during its propagation in the random medium. The second kind of relations can be considered as a more detailed model for the first.

In general, the state of polarization of a lightwave will be modelled as a random variable on the Poincaré sphere [5]. If we are considering the input and output states of polarization to a birefringent random medium taken as a whole, we will denote by S_{in} and S_{out} the random variables on the Poincaré sphere corresponding to the input and output states of polarization. The effect of the random medium on S_{in} is given by its Mueller matrix. Here, it is understood that this Mueller matrix is a matrix-valued random variable [14]. In general, the Mueller matrix acts on the Stokes vector of the incident lightwave and not on the reduced Stokes vector S_{in} . However, in the special case of a birefringent medium the Mueller matrix reduces to a rotation matrix acting on the reduced Stokes vector S_{in} . According to this discussion, the effect of a random birefringent medium is given by a random variable $\mathbf{R}$, with values in the rotation group $SO(3)$, according to the following equation:

$$S_{\text{out}} = \mathbf{R} S_{\text{in}} \quad (1)$$

where S_{in} and S_{out} are random variables on the Poincaré sphere S^2 and $\mathbf{R}$ is a $SO(3)$ -valued random variable.

When considering the propagation of a lightwave in a birefringent random medium we will give its state of polarization by a stochastic process $S(z)$ with values on S^2 . Here $z \geq 0$ is the distance along which the wave has propagated through the random medium. For every z , $S(z)$ is a random variable on S^2 giving the state of polarization of the wave after a distance z of its trajectory in the medium. In relation to Equation (1), we can note $S(0) = S_{\text{in}}$ and $S(Z) = S_{\text{out}}$ for some given distance Z . We make the hypothesis that a length z of any trajectory of the wave can be described using a rotation matrix (as in (1)) denoted $\mathbf{R}(z)$. We have, as a first model for the propagation:

$$S(z) = \mathbf{R}(z)S(0) \quad (2)$$

By associating a matrix $\mathbf{R}(z)$ to every distance z , over which the wave propagates in the medium, we are implicitly making two assumptions. First, that all scattering or beam divergence effects in the medium can be ignored. This amounts to supposing that light remains collimated in the medium [15]. Second, that the properties of the medium, although random, do not vary in time. Under these assumptions, Equation (2) is derived from the wave equation for a spectral component of the effective transverse electrical field of the lightwave in [4, 15].

In the following section we give the mathematical tools we will use to analyse Equations (1) and (2) to obtain the main equations of our model as given in Section 4.

3. Irreducible representations of the rotation group

The main idea of this paper is to decompose the higher-order statistics of the reduced Stokes parameters along the irreducible representations of $SO(3)$. This decomposition is central to the spherical convolution theorem which we present shortly and which allows us to obtain the laws of transformation of the higher-order statistics of the reduced Stokes parameters. These laws of transformation constitute the main equations of our model and are given in Section 4. In fact, we

will be interested in the realization of the irreducible representations of $SO(3)$ in the space of square integrable functions on S^2 . Under this form, the study of the irreducible representations of $SO(3)$ is known as spherical harmonic analysis [6–8].²

Let $f(s) = f(\phi, \theta)$ – where ϕ and θ are the azimuth and polar angle of $s = (s_1, s_2, s_3)$ – be a square integrable function on S^2 . f can be decomposed along the orthogonal basis of spherical harmonics $Y_m^l(s)$, where $l \in \mathbb{N}$ and $-l \leq m \leq l$:

$$f(s) = \sum_{l \in \mathbb{N}} \sum_{m=-l}^{m=l} (2l+1) \hat{f}_m^l Y_m^l(s) \quad (3)$$

where $\hat{f}_m^l$ is the projection:

$$\hat{f}_m^l = \int_{S^2} f(s) \overline{Y_m^l(s)} ds \quad (4)$$

ds is the Haar measure $ds = \frac{1}{4\pi} \sin\theta d\phi d\theta$ and the bar represents complex conjugation. Remember that spherical harmonics are given by the formula:

$$Y_m^l(s) = Y_m^l(\phi, \theta) = \sqrt{\frac{(l-m)!}{(l+m)!}} P_m^l(\cos\theta) e^{im\phi} \quad (5)$$

where the P_m^l are associated Legendre functions, and also by the Cartesian formula:

$$\begin{aligned} Y_m^l(s) &= Y_m^l(s_1, s_2, s_3) \\ &= i^m \frac{\sqrt{(l-m)!(l+m)!}}{2\pi l!} \left(\frac{s_1 + is_2}{\sqrt{s_1^2 + s_2^2}} \right)^m \int_0^{2\pi} (s_3 + i\sqrt{1-s_3^2} \cos t)^l \cos(mt) dt \end{aligned} \quad (6)$$

Under the effect of a rotation $\mathbf{r} \in SO(3)$, the $(2l+1)$ coefficients $\hat{f}_m^l$ (for every $l \in \mathbb{N}$) transform according to an irreducible unitary representation, of dimension $(2l+1)$, of $SO(3)$. In other words, for the rotated function $f_{\mathbf{r}}(s) = f(\mathbf{r}s)$ we have the development:

$$f_{\mathbf{r}}(s) = \sum_{l \in \mathbb{N}} \sum_{m=-l}^{m=l} (2l+1) (\hat{f}_{\mathbf{r}})_m^l Y_m^l(s) \quad (7)$$

where the coefficients $(\hat{f}_{\mathbf{r}})_m^l$ of the development are given (for every $l \in \mathbb{N}$) by the linear transformation:

$$(\hat{f}_{\mathbf{r}})_m^l = \sum_{n=-l}^{n=l} \overline{D_{mn}^l(\mathbf{r})} \hat{f}_n^l \quad (8)$$

The functions on $SO(3)$, $D_{mn}^l(\mathbf{r})$, which give the elements of the matrix of the linear transformation (8), are the matrix elements of the irreducible representation of dimension $(2l+1)$ of

²See these same references for all the results in this section.

$SO(3)$. They can be given explicitly, as functions of the Euler angles (ϕ, θ, ψ) of the rotation $\mathbf{r}$:

$$D_{mn}^l(\mathbf{r}) = D_{mn}^l(\phi, \theta, \psi) = e^{-im\phi} P_{mn}^l(\cos\theta) e^{-in\psi} \quad (9)$$

where the polynomials P_{mn}^l are given by Jacobi polynomials.

The $(2l+1) \times (2l+1)$ matrices $\mathbf{D}^l(\mathbf{r}) = \{D_{mn}^l(\mathbf{r})\}$, $-l \leq m, n \leq l$, are unitary and satisfy the homomorphism property:

$$\mathbf{D}^l(\mathbf{r}_2 \mathbf{r}_1) = \mathbf{D}^l(\mathbf{r}_2) \mathbf{D}^l(\mathbf{r}_1) \quad (10)$$

where $\mathbf{r}_1, \mathbf{r}_2 \in SO(3)$ and l is any natural number.

The functions $D_{mn}^l(\mathbf{r})$, on $SO(3)$, play a similar role to that played by spherical harmonics on S^2 . Indeed, let $g(\mathbf{r})$ be a square integrable function on $SO(3)$, then g can be decomposed along the orthogonal basis of matrix elements $D_{mn}^l(\mathbf{r})$:

$$g(\mathbf{r}) = \sum_{l \in \mathbb{N}} \sum_{m, n=-l}^{m, n=l} (2l+1) \hat{g}_{mn}^l \overline{D_{mn}^l(\mathbf{r})} \quad (11)$$

where $\hat{g}_{mn}^l$ is the projection:

$$\hat{g}_{mn}^l = \int_{SO(3)} g(\mathbf{r}) D_{mn}^l(\mathbf{r}) d\mathbf{r} \quad (12)$$

and $d\mathbf{r}$ is the Haar measure on $SO(3)$, given in terms of Euler angles by $d\mathbf{r} = \frac{1}{8\pi^2} \sin\theta d\phi d\theta d\psi$.

It is interesting, at this point, to rewrite the developments (3) and (11) in matrix form. This will clarify the rest of the article and avoid cumbersome notations. To do this, adopt the following notations: denote by $\hat{\mathbf{f}}^l$, for every $l \in \mathbb{N}$, the $(2l+1)$ -dimensional column vector whose components are the coefficients $\hat{f}_m^l$, $-l \leq m \leq l$, from Equation (3). In the same way, denote by $\mathbf{Y}^l(s)$ the column vector of spherical harmonics $Y_m^l(s)$. Denote by $\hat{\mathbf{g}}^l$ the $(2l+1) \times (2l+1)$ matrix with elements $\hat{g}_{mn}^l$, in Equation (11). The notation $\mathbf{D}^l(\mathbf{r})$ has already been explained. Using these notations, Equation (3) can be rewritten as follows:

$$f(s) = \sum_{l \in \mathbb{N}} (2l+1) (\hat{\mathbf{f}}^l)^T \mathbf{Y}^l(s) \quad (13)$$

where T stands for transposition. Equation (11) can be rewritten as:

$$g(\mathbf{r}) = \sum_{l \in \mathbb{N}} (2l+1) \text{Tr} \left(\hat{\mathbf{g}}^l (\mathbf{D}^l(\mathbf{r}))^\dagger \right) \quad (14)$$

where Tr stands for the trace and $\dagger$ for the Hermitian conjugate.

As explained above, the spherical convolution theorem will be used to obtain the basic equations of our model. This theorem uses the decompositions (13) and (14) to transform a spherical convolution into a family of matrix products. It is an analogue of the classical convolution theorem, which is used to transform a convolution of functions of a real variable into a product of their Fourier transforms.

The convolution of two functions g_2 and g_1 on $SO(3)$ is a function g on $SO(3)$ defined as:

$$g(\mathbf{r}) = (g_2 * g_1)(\mathbf{r}) = \int_{SO(3)} g_2(\mathbf{t}) g_1(\mathbf{t}^{-1}\mathbf{r}) d\mathbf{t} \quad (15)$$

This definition is analogous to that of the convolution of two functions of a real variable. Formally, it can be obtained from it by replacing the usual $\mathbf{r} - \mathbf{t}$ by $\mathbf{t}^{-1}\mathbf{r}$. The convolution of a function g on $SO(3)$ with a function f on S^2 is a function h on S^2 . It has a similar definition:

$$h(s) = (g * f)(s) = \int_{SO(3)} g(\mathbf{t}) f(\mathbf{t}^{-1}s) d\mathbf{t} \quad (16)$$

The convolution theorem states that Equation (15) is equivalent to:

$$\hat{\mathbf{g}}^l = \hat{\mathbf{g}}_2^l \hat{\mathbf{g}}_1^l \text{ for } l \in \mathbb{N} \quad (17)$$

where the product on the right-hand side is a product of $(2l + 1) \times (2l + 1)$ matrices. It also states that Equation (16) is equivalent to:

$$\hat{\mathbf{h}}^l = \hat{\mathbf{g}}^l \hat{\mathbf{f}}^l \text{ for } l \in \mathbb{N} \quad (18)$$

where the product on the right-hand side is a product of the $(2l + 1) \times (2l + 1)$ matrix $\hat{\mathbf{g}}^l$ with the $(2l + 1)$ -dimensional vector $\hat{\mathbf{f}}^l$.

4. Main equations of the model

In this section, we give the main equations of our model. In the following Section 5, they will be interpreted as giving the laws of transformation of the higher-order statistics of the reduced Stokes parameters. In Section 6, they will be used to give a detailed probabilistic description of the propagation of polarized light in a random birefringent medium.

Let us return to the physical situation described by Equation (1) of Section 2. This equation describes a random medium taken as a whole. It gives the relation between the random variables on the Poincaré sphere, $\mathbf{S}_{\text{in}}$ and $\mathbf{S}_{\text{out}}$, describing (respectively) the input and output states of polarization to the medium. In this section, the mathematical tools given in the last Section 3 will be used to analyse Equation (1). First, we will establish the relation, between the probability densities of the random variables $\mathbf{S}_{\text{out}}$, $\mathbf{R}$ and $\mathbf{S}_{\text{in}}$, appearing in this equation, as a spherical convolution. Then we will use the convolution theorem (18) to transform this relation.

Denote by $p_{\text{in}}(s)$ and $p_{\text{out}}(s)$ the probability densities of $\mathbf{S}_{\text{in}}$ and $\mathbf{S}_{\text{out}}$. Denote by $p_{\mathbf{R}}(\mathbf{r})$ the probability density of $\mathbf{R}$. To see that the relation between these three probability densities is given by a spherical convolution (16), apply the law of total probabilities: The probability for $\mathbf{S}_{\text{out}}$ to take a value near $s \in S^2$, conditionally to the event ' $\mathbf{R}$ is near $\mathbf{r} \in SO(3)$ ' is equal to $p_{\text{in}}(\mathbf{r}^{-1}s)$. Indeed, the random variables $\mathbf{S}_{\text{in}}$ and $\mathbf{R}$ are independent (the properties of the medium do not depend on the incident wave). By summing these conditional probabilities, we obtain:

$$p_{\text{out}}(s) = \int_{SO(3)} p_{\mathbf{R}}(\mathbf{r}) p_{\text{in}}(\mathbf{r}^{-1}s) d\mathbf{r} = (p_{\mathbf{R}} * p_{\text{in}})(s) \quad (19)$$

which is a spherical convolution. This relation is analogous to the classical relation stating that the probability density of the sum of two independent real random variables is the convolution of their probability densities [16].

We now apply the convolution theorem – Equation (18) – to this last relation. Denote by ξ_{in}^l and ξ_{out}^l the coefficient vectors in the development (13) of (respectively) p_{in} and p_{out} . Denote by $\mathbf{R}^l$ the coefficient matrices in the development (14) of $p_{\mathbf{R}}$. Then according to (18):

$$\xi_{\text{out}}^l = \mathbf{R}^l \xi_{\text{in}}^l \quad (20)$$

This is the first main equation of our model. First of all, it describes the state of polarization of a lightwave using a family of coefficients, e.g. ξ_{in}^l and ξ_{out}^l , instead of a probability density function on S^2 . It also describes the law of transformation of these coefficients, when the lightwave passes through a random birefringent medium. This description presents some similarities to the model suggested in [12], for the description of quantum polarization effects.

Now consider that S_{in} is affected by the composition of two ‘random elements’. That is, let $S_{\text{out}} = \mathbf{R}_2 \mathbf{R}_1 S_{\text{in}}$, where $\mathbf{R}_2$ and $\mathbf{R}_1$ are, as in (1), random variables in $SO(3)$. Let $\mathbf{R} = \mathbf{R}_2 \mathbf{R}_1$ and denote by $p_{\mathbf{R}}(\mathbf{r})$ the probability density of $\mathbf{R}$ and similarly write $p_{\mathbf{R}_2}$ and $p_{\mathbf{R}_1}$. Applying the same reasoning as that used to obtain Equation (19), we have:

$$p_{\mathbf{R}}(\mathbf{r}) = \int_{SO(3)} p_{\mathbf{R}_2}(\mathbf{t}) p_{\mathbf{R}_1}(\mathbf{t}^{-1} \mathbf{r}) d\mathbf{t} = (p_{\mathbf{R}_2} * p_{\mathbf{R}_1})(\mathbf{r}) \quad (21)$$

In order to apply the convolution theorem (17) to this equation, denote by $\mathbf{R}^l$, $\mathbf{R}_2^l$ and $\mathbf{R}_1^l$ the coefficient matrices in the development (14) of (respectively) $p_{\mathbf{R}}$, $p_{\mathbf{R}_2}$ and $p_{\mathbf{R}_1}$. Then, according to (17):

$$\mathbf{R}^l = \mathbf{R}_2^l \mathbf{R}_1^l \quad (22)$$

This relation can be generalized to the composition of n random elements, $\mathbf{R} = \mathbf{R}_n \mathbf{R}_{n-1} \dots \mathbf{R}_2 \mathbf{R}_1$. Using a similar notation to that in (22), we have:

$$\mathbf{R}^l = \mathbf{R}_n^l \mathbf{R}_{n-1}^l \dots \mathbf{R}_2^l \mathbf{R}_1^l \quad (23)$$

This is the second main equation of our model. It shows how Equation (20) can be applied when the input wave is subjected to the consecutive effect of several random elements. This relation will be used in Section 6 where we study the propagation of polarized light in random birefringent media.

5. Statistical interpretation and generalization of Stokes formalism

In this section, we give a statistical interpretation of the results of the last section, especially Equation (20). We study the relation of these results to the classical Stokes–Mueller formalism [5] as well as to recent works that study the role of higher-order statistics in polarization optics [9,10,12,13]. The main idea is to notice that the coefficient vectors ξ_{in}^l and ξ_{out}^l , of Equation (20), contain combinations of the moments of order l of the random vectors S_{in} and S_{out} , that transform under irreducible representations of $SO(3)$. This observation is used to generalize the notions of reduced Stokes vector and of degree of polarization to higher-order statistics.

5.1. Generalized reduced Stokes vectors

Let us consider a random variable S on the Poincaré sphere, describing the state of polarization of a lightwave. Let $p(s)$ be the probability density of S . If ξ^l are the coefficient vectors of the development (13) of p , then by definition – see Equation (4):

$$\xi^l = \int_{S^2} p(s) \overline{Y^l(s)} ds = \mathbb{E}(\overline{Y^l(S)}) \quad (24)$$

By considering the Cartesian expression (6) for the spherical harmonics $Y^l(s)$, it is possible to see that the coefficient vector ξ^l contains complex combinations of the moments of order l of the vector S . Let us take the example of $l = 1$. Using formula (6) we can see that the vector $Y^1(S)$ is given by³ $Y^1(S) = \left(\frac{S_1 - iS_2}{\sqrt{2}}, S_3, -\frac{S_1 + iS_2}{\sqrt{2}} \right)^T$. It results from this that $\xi^1 = \left(\frac{\mathbb{E}(S_1 + iS_2)}{\sqrt{2}}, \mathbb{E}(S_3), -\frac{\mathbb{E}(S_1 - iS_2)}{\sqrt{2}} \right)^T$. In other words, ξ^1 is related by a complex change of basis to the average reduced Stokes vector $\mathbb{E}(S) = (\mathbb{E}(S_1), \mathbb{E}(S_2), \mathbb{E}(S_3))$.

The degree of polarization is classically defined using the average reduced Stokes vector, that is, using the second order statistics of the field of the lightwave. It is given by [5]:⁴

$$P = \|\mathbb{E}(S)\| = \sqrt{\mathbb{E}^2(S_1) + \mathbb{E}^2(S_2) + \mathbb{E}^2(S_3)} \quad (25)$$

Note that we can also write $P = \|\xi^1\| = \sqrt{|\xi_{-1}^1|^2 + |\xi_0^1|^2 + |\xi_1^1|^2}$.

This example shows that, using the first coefficient vector ξ^1 of the development of the probability density of S , we retrieve the classical average reduced Stokes vector as well as the classical notion of degree of polarization [5]. To generalize this result to higher-order statistics, we construct, for every $l \in \mathbb{N}$, a real version of the $(2l + 1)$ -dimensional coefficient vector ξ^l . This can be done as follows. Define for every l the $(2l + 1)$ -dimensional real vector S^l , by putting:

$$S^l = \begin{cases} S_m^l = \frac{-1}{\sqrt{2}}((-1)^m \xi_{-m}^l + \xi_m^l) = -\sqrt{2}\Re(\xi_m^l) & \text{for } m > 0 \\ S_0^l = \xi_0^l & \text{for } m = 0 \\ S_m^l = \frac{-1}{i\sqrt{2}}((-1)^{-m} \xi_m^l - \xi_{-m}^l) = \sqrt{2}\Im(\xi_{-m}^l) & \text{for } m < 0 \end{cases} \quad (26)$$

We will call the vector S^l the reduced Stokes vector of order l . It contains the moments of order l of the vector S . For example:

$$S^2 = \left(\sqrt{\frac{3}{2}} \mathbb{E}(2S_1 S_2), \sqrt{3} \mathbb{E}(S_3 S_2), \mathbb{E}\left(\frac{3}{2} S_3^2 - \frac{1}{2}\right), \sqrt{3} \mathbb{E}(S_3 S_1), \mathbb{E}\left(\sqrt{\frac{3}{2}} (S_2^2 - S_1^2)\right) \right)^T \quad (27)$$

and, $S^1 = (\mathbb{E}(S_1), \mathbb{E}(S_3), \mathbb{E}(S_2))$, so that S^1 is equivalent to the average reduced Stokes vector $\mathbb{E}(S)$.

³ $Y^l(s)$ has been defined as a column vector, whence the transpose.

⁴For the definition (25) to correspond precisely to the usual definition of the degree of polarization, we must add the hypothesis that the total intensity of the lightwave is independent from the other three components of the Stokes vector. This is not a very restrictive hypothesis when considering birefringent media.

The definition of the vectors $\mathbf{S}^l$ allows the generalization of the notion of degree of polarization to higher-order statistics. Indeed, by analogy with formula (25), we can define:

$$P^l = \|\mathbf{S}^l\| = \|\boldsymbol{\xi}^l\| \quad (28)$$

We will call P^l the degree of polarization of order l . It is possible to prove that $P^l \in [0, 1]$. Indeed, P^l is evidently positive. Note also that:

$$P^l = \|\boldsymbol{\xi}^l\| \leq \sqrt{(2l+1)}|\xi_n^l|$$

where $-l \leq n \leq l$ is such that $\max_{\{-l \leq m \leq l\}} |\xi_m^l| = |\xi_n^l|$. Now [8, 6]:⁵

$$|\xi_n^l|^2 \leq \int_{S^2} p(s) |Y_n^l(s)|^2 ds \leq \int_{S^2} |Y_n^l(s)|^2 ds = \frac{1}{2l+1}$$

So that $P^l \leq 1$. For $l = 1$, the fact that $P^1 = P \in [0, 1]$ is well established [5]. It means that the only physical states of polarization are the ones with $\mathbb{E}(\mathbf{S})$ inside the Poincaré sphere.

5.2. Examples and relation to other work

In this subsection, we give a few examples of how the notions of reduced Stokes vector of order l and degree of polarization of order l , introduced in the last Section 5.1, can be used to distinguish states of polarization which are indistinguishable in the framework of classical models for polarization [5]. We also explain how our model is related to other recent works on higher-order statistics in polarization optics [9,10,12,13].

Remember that – see the introduction – classical models for polarization only use the first and second order statistics of the field of the lightwave [5]. This corresponds to using the average Stokes vector or reduced Stokes vector. This approach is sufficient for Gaussian fields but fails for non-Gaussian fields [9, 13]. Indeed, considering only the average reduced Stokes vector $\mathbb{E}(\mathbf{S})$, or equivalently only $\mathbf{S}^1$, would lead to identifying states of polarization which have the same average $\mathbb{E}(\mathbf{S})$ but might have different higher-order moments of this vector. That is, in the formalism introduced in the last Section 5.1, different $\mathbf{S}^l$ for $l > 1$.

In [13], the three following states of polarization are studied. In the framework of the classical Stokes formalism, they are all considered to be identical states of polarization corresponding to totally depolarized light. However, they all have different higher-order statistics: (i) a state of polarization with reduced Stokes vector $\mathbf{S}$ distributed uniformly on the Poincaré sphere; (ii) a state of polarization with reduced Stokes vector $\mathbf{S}$ distributed uniformly on the equator of the Poincaré sphere (only linearly polarized light); (iii) a state of polarization with reduced Stokes vector $\mathbf{S}$ taking the value $(0, 0, 1)^T$ with probability $1/2$ and the value $(0, 0, -1)^T$ with probability $1/2$ (only left or right circularly polarized light).

In all these three cases $\mathbb{E}(\mathbf{S}) = \mathbf{S}^1 = (0, 0, 0)^T$ and $P = P^1 = 0$, so that, in the classical Stokes formalism, they all correspond to the same state of polarization. Using the vectors $\mathbf{S}^l$, with $l > 1$, we can see how they are different: (i) For this state, the vector $\mathbf{S}^l$ is zero and $P^l = 0$ for all $l \geq 1$. (ii) For this state $\mathbf{S}^1 = (0, 0, 0)^T$ and $P^1 = 0$, however $\mathbf{S}^2 = (0, 0, -1/2, 0, 0)^T$ and $P^2 = 1/2$. (iii) For this state $\mathbf{S}^1 = (0, 0, 0)^T$ and $P^1 = 0$, however $\mathbf{S}^2 = (0, 0, 1, 0, 0)^T$ and $P^2 = 1$.

⁵The first inequality follows from Jensen's inequality.

It appears from these three examples, in addition to the somewhat evident fact that higher-order statistics are necessary when studying non-Gaussian fields, that in order to call a state of polarization totally depolarized, it is not sufficient to have $P = 0$. Indeed, the state of polarization can satisfy $P = 0$ but still refer to a specific type of polarization ellipse: only linear polarization (state (ii)), or only circular polarization (state (iii)). More examples are given in [9, 13]. In [12] a new definition of the degree of polarization is proposed which takes into account higher-order statistics. This definition is based on the following quantity:

$$D = \int_{S^2} [p(s) - 1]^2 ds \quad (29)$$

where $p(s)$ is the probability density of the reduced Stokes vector S . D is actually a quadratic measure of the difference between $p(s)$ and a uniform distribution, whose probability density is equal to 1. The degree of polarization is then defined as [12]:

$$P = \frac{D}{1 + D} \in [0, 1] \quad (30)$$

This definition is closely related to the quantities P^l , which we introduced in the last Section 5.1. In fact [6, 8]:

$$D = \sum_{l \geq 1} (P^l)^2 \quad (31)$$

It is possible to synthesize the results of our model, with the measure of degree of polarization proposed in [12], by defining a totally depolarized state of polarization as one for which $P^l = 0$ for all $l \geq 1$. If we have $P^l = 0$ only for $1 \leq l \leq L$, then we can say that the state of polarization is depolarized to the order L . Classical models only consider depolarization to the order 1. Also, we should consider a state of polarization to be totally polarized only if $P^l = 1$ for all $l \in \mathbb{N}$. This corresponds to a distribution concentrated at one point on the sphere.

Let us make a final observation, without developing it. The measures of degree of polarization, and the criteria for distinguishing states of polarization with the same second order statistics, proposed in this article and in [9, 12, 13], do not take into account the intensity distribution of the lightwave. The definition of degree of polarization proposed by Réfrégier [10, 11], is based on the whole probability distribution of the field of the lightwave, including its intensity distribution. A comparative study of these two general approaches may help clarify the importance of including the intensity distribution in a measure of the degree of polarization.

6. Evolution of the state of polarization during propagation

In this section, we use our main Equations (20) and (23) of Section 4, to study the evolution of the state of polarization of a lightwave, propagating in a random birefringent medium. This problem, as mentioned in the introduction, arises in optical fibre telecommunications [1–3]. By using the model we have introduced in this paper, we will be able to achieve a detailed probabilistic description of the problem at hand. In particular, we will be able to give an exact analytical expression for the probability density on the Poincaré sphere, representing the state of polarization, after any distance of propagation. We will also be able to follow exactly the evolution of different averaged quantities, such as the degree of polarization.

The physical problem we are interested in is the one described by Equation (2) in Section 2. We have a lightwave, propagating in a random birefringent medium. We denote by $z \geq 0$ the distance along which the wave has propagated. To each $z \geq 0$ is associated a random variable $S(z)$ on the Poincaré sphere. It represents the state of polarization after a distance z in the medium. As explained in Section 2, we suppose that there exists, for every z , a rotation $\mathbf{R}(z)$ such that:

$$S(z) = \mathbf{R}(z)S(0) \quad (32)$$

The evolution of the state of polarization during propagation can then be described by a stochastic process $S(z)$ on the sphere, or a stochastic process $\mathbf{R}(z)$ on $SO(3)$. Our description of this evolution is based on the mathematical concept of a *Lévy process* on $SO(3)$ [17], which we introduce in the following Section 6.1.

6.1. Lévy process model

We model $\mathbf{R}(z)$ as a Lévy process on $SO(3)$. This model reflects a set of simple physical properties of the propagation medium. It is based on the following hypotheses:

- **Independent increments:** For $z_1 < z_2$ we have that $\mathbf{R}(z_1)$ and $\mathbf{R}(z_2)\mathbf{R}^{-1}(z_1)$ are independent. Physically, this means that non-overlapping parts of the medium are not coupled.
- **Stationary increments:** For $z_1 < z_2$ we have that $\mathbf{R}(z_2)\mathbf{R}^{-1}(z_1) = \mathbf{R}(z_2 - z_1)$.⁶ Physically, this means that the medium is homogeneous and only locally random. This hypothesis, more generally means that:

$$\mathbf{R}(z) = \mathbf{R}(z - z_n)\mathbf{R}(z_n - z_{n-1}) \dots \mathbf{R}(z_2 - z_1)\mathbf{R}(z_1) \text{ for } z > z_n > \dots > z_2 > z_1 \quad (33)$$

- **Stochastic continuity:** The stochastic process $\mathbf{R}(z)$ is stochastically continuous. This means that the probability for $\mathbf{R}(z_1)$ and $\mathbf{R}(z_2)$ to be different tends to zero as $z_2 - z_1$ goes to zero. Physically, this means that a very short length of the medium cannot induce a big change in the state of polarization.⁷
- We add the simplifying hypothesis that $\mathbf{R}(0) = \mathbf{I}$, where $\mathbf{I}$ is the 3×3 identity matrix.

A stochastic process on $SO(3)$ – or any other Lie group – satisfying these properties is called a left Lévy process, or just a Lévy process [17].

This model can be very effectively reduced, using the main Equations (20) and (23). Denote by $p_{\mathbf{R}(z)}(\mathbf{r})$ the probability density of $\mathbf{R}(z)$. According to equation (33):

$$\mathbf{R}(z) = \mathbf{R}(z - z_1)\mathbf{R}(z_1) \text{ for } z > z_1 \quad (34)$$

Using the results of Section 4 – Equation (21) – we can write:

$$p_{\mathbf{R}(z)}(\mathbf{r}) = (p_{\mathbf{R}(z-z_1)} * p_{\mathbf{R}(z_1)})(\mathbf{r}) \quad (35)$$

⁶This equality is only in distribution, i.e. the two members of the equality have the same probability distribution.

⁷This does not imply that the trajectories on the Poincaré sphere of $S(z)$ are continuous. They can have jump discontinuities.

Using Equation (22), we transform this last equation:

$$\mathbf{R}^l(z) = \mathbf{R}^l(z - z_1)\mathbf{R}^l(z_1) \quad (36)$$

where $\mathbf{R}^l(z)$ are the coefficient matrices in the development (14) of $p_{\mathbf{R}(z)}$.

The stochastic continuity of the process $\mathbf{R}(z)$ implies the continuity in z of the matrices $\mathbf{R}^l(z)$. The only continuous solution of (36) satisfying $\mathbf{R}(0) = I$ is [17]:

$$\begin{aligned} \mathbf{R}^0(z) &= I \\ \mathbf{R}^l(z) &= e^{\mathbf{t}^l z} \text{ for } l \geq 1 \end{aligned} \quad (37)$$

where $\mathbf{t}^l$ (for every $l \geq 1$) is a constant matrix (not a function of z). These constant matrices are called generators of the process $\mathbf{R}(z)$. We have:

$$\mathbf{t}^l = \frac{d}{dz} \mathbf{R}^l(z)|_{z=0} \quad (38)$$

It follows from the decomposition formula (14) that:

$$p_{\mathbf{R}(z)}(\mathbf{r}) = 1 + \sum_{l \geq 1} (2l + 1) \text{Tr}(e^{\mathbf{t}^l z} (\mathbf{D}^l(\mathbf{r}))^\dagger) \quad (39)$$

which gives the probability density of $\mathbf{R}(z)$ for any z .

From this last result, the probability density on S^2 of $\mathbf{S}(z)$ can be derived in a direct way. We have already noted that $\mathbf{S}(z) = \mathbf{R}(z)\mathbf{S}(0)$. If $p_{\mathbf{S}(z)}(s)$ is the probability density of $\mathbf{S}(z)$ and $\xi^l(z)$ are the coefficient vectors in its development (13), then by Equations (19) and (20):

$$\begin{aligned} \xi^0(z) &= I \\ \xi^l(z) &= e^{\mathbf{t}^l z} \xi^l(0) \text{ for } l \geq 1 \end{aligned} \quad (40)$$

Using the decomposition formula (13) it follows that:

$$p_{\mathbf{S}(z)}(s) = 1 + \sum_{l \geq 1} (2l + 1) (e^{\mathbf{t}^l z} \xi^l(0))^T \mathbf{Y}^l(s) \quad (41)$$

which gives the probability density of $\mathbf{S}(z)$ for any z .

By modelling the evolution of the state of polarization, during propagation in a random medium, as a Lévy process, the descriptions (39) and (41) of this evolution have been achieved. Practically, these descriptions give the evolution of the state of polarization in function of the generator matrices $\mathbf{t}^l$, ($l \geq 1$). It is clear from Equation (38) that these matrices characterize the propagation medium locally. That is, they describe propagation over small ‘infinitesimal’ distances. These matrices are not known *a priori*. There are two ways of giving them: The first way is to use a local model for the propagation, in the form of a stochastic differential equation. Such local models can be found in [1, 2]. The second way is to note that these matrices are parameters of the probability density of $\mathbf{S}(z)$. It is possible to estimate them given realizations of $\mathbf{S}(z)$ – see the discussion in the introduction. We will return to these two approaches in Section 6.3.

6.2. Depolarization

The most important effect of a random medium on the state of polarization of a lightwave is to depolarize it. After a long distance of propagation, we can expect the state of polarization of the lightwave to become totally depolarized. Here, we use the mathematical description presented in the last Section 6.1 to give the evolution, during propagation in a random birefringent medium, of the degree of polarization. We see that this evolution tends to a totally depolarized state independently of initial conditions.

It is possible to show, under very general conditions [17], that the (real parts of the) eigenvalues of the generator matrices $\mathbf{t}^l$ in Equation (37) are all negative. Since $\mathbf{R}^l = \mathbf{e}^{t^l z}$, this means that $\mathbf{R}^l \rightarrow 0$ exponentially for large z . By taking the limit of Equation (41) for large z , we find that $p_{S(z)}(s) \rightarrow 1$. In other words the probability distribution of $S(z)$ tends to a uniform distribution on the Poincaré sphere, which is characteristic of a totally depolarized state. Note, from Equation (41), that as the matrices $\mathbf{R}^l$ tend to zero, $p_{S(z)}(s)$ will tend to a uniform distribution independently of the initial distribution $p_{S(0)}(s)$.

We have established that the state of polarization of a lightwave propagating in a random medium tends, with the distance of propagation, to a totally depolarized state represented by a uniform probability distribution on the Poincaré sphere. Let us now examine the dependence on z of the degree of polarization of order l , P^l , for $l \geq 1$. Remember that P^l was defined in Section 5.1, Equation (28), as the norm of the complex vector ξ^l . Using Equation (40), we can write:

$$P^l(z) = \|\xi^l(z)\| = \|\mathbf{e}^{t^l z} \xi^l(0)\| \quad (42)$$

It is clear in this equation that $P^l(z)$ tends to zero independently of initial conditions. This is, in particular, true for the usual degree of polarization $P = P^1$.

We have just given the law of evolution, i.e. the dependence on z , of the degree of polarization of order l , P^l , for all $l \geq 1$. These quantities are an example of what might be called averaged quantities associated to the process $S(z)$. That is, combinations of the averages of functions of $S(z)$. Averaged quantities are, of course, deterministic. Let $f(s)$ be any real square integrable function on the sphere S^2 . An example of an averaged quantity is $\mathbb{E}(f(S(z)))$. The fact that f is real can be used to slightly transform its development (13), in the following way:

$$f(s) = \overline{f(s)} = \sum_{l \in \mathbb{N}} (2l+1) (\hat{f}^l)^\dagger \overline{Y^l(s)} \quad (43)$$

Using the fact that – Equation (24) – $\xi^l(z) = \mathbb{E}(\overline{Y^l(S(z))})$, it is possible to write:

$$\mathbb{E}(f(S(z))) = \sum_{l \in \mathbb{N}} (2l+1) (\hat{f}^l)^\dagger \xi^l(z) \quad (44)$$

or, using the expression (40) for $\xi^l(z)$:

$$\mathbb{E}(f(S(z))) = \sum_{l \in \mathbb{N}} (2l+1) (\hat{f}^l)^\dagger \mathbf{e}^{t^l z} \xi^l(0) \quad (45)$$

The last Equation (45) shows that the evolution of any averaged quantity can be followed exactly if the generator matrices $\mathbf{t}^l$ are known. Averaged quantities include the entropy of the state of polarization, the average parameters of the ellipse of polarization, or any other attribute of the state of polarization that we may wish to study. Equation (45) can for instance be used to establish

that the informational entropy of the state of polarization is strictly increasing during propagation in a random birefringent medium. Thus, depolarization can be associated with an increasing informational entropy.

6.3. Estimation of physical parameters: an example

In Section 6.1, the evolution of $S(z)$ during propagation was modelled as a Lévy process. This model led to an analytical formula (41), for the probability density of $S(z)$, containing the generator matrices $\mathbf{t}^l$ ($l \geq 1$) – see Equations (37) and (38) – as free parameters. As mentioned before, these matrices are related to the local properties of the propagation medium. In order to give them concrete expressions or values, two approaches can be used: The first is to use a local physical model for the evolution of $S(z)$. This model can take the form of a stochastic differential equation [1, 2]. The second way is to estimate them from realizations of $S(z)$, since indeed, these matrices appear as parameters of the probability density of $S(z)$.

A particularly simple case arises when the generator matrices correspond to the stochastic differential equation proposed in [2]. This is a stochastic differential equation describing the evolution of the vector on the Poincaré sphere $S(z)$ during propagation in an optical fibre affected by PMD. The main idea of this model is that $S(z)$ rotates on the Poincaré sphere with an angular velocity which is essentially a white noise vector:

$$\frac{d}{dz} S(z) = \mu \mathbf{W} \times S(z) \quad (46)$$

where μ is a constant, $\mathbf{W}$ is a three-dimensional white noise vector, $\times$ is the vector (cross) product and the equation is to be understood as a Stratonovich stochastic differential equation [2].

This stochastic differential equation is well known in mathematics [6, 17–19]. It describes *Brownian motion* on the sphere S^2 . The generator matrices for this process are given by [6, 17–19]:

$$\mathbf{t}^l = \frac{-\mu^2}{2} l(l+1) \mathbf{I}_l \quad (47)$$

where $\mathbf{I}_l$ is the $(2l+1) \times (2l+1)$ identity matrix. By replacing this result in (41), it follows that:

$$p_{S(z)}(s) = \sum_{l \in \mathbb{N}} (2l+1) e^{\frac{-\mu^2}{2} l(l+1)z} (\boldsymbol{\xi}^l(0))^T \mathbf{Y}^l(s) \quad (48)$$

This situation is particularly simple. By starting from the local model suggested in [2], we arrive at an expression for the probability density of $S(z)$ which depends on only one free parameter, namely μ , instead of having the (infinite) family of matrices $\mathbf{t}^l$ as free parameters.

Note also that, in this case, the degree of polarization of order l takes on a simpler form:

$$P^l(z) = e^{\frac{-\mu^2}{2} l(l+1)z} \|\boldsymbol{\xi}^l(0)\| \quad (49)$$

Now let us see how,⁸ in this simple case which only has one parameter in the probability density of $S(z)$, we can use realizations of $S(z)$ to estimate μ . We need to consider the medium as a whole

⁸For a non-parametric estimation method, which allows the estimation of the whole probability density of S_{out} , independently of any physical model, see [16].

(a closed system). We note $S(0) = S_{\text{in}}$ and consider a length Z of the medium so that we can write $S(Z) = S_{\text{out}}$. The medium is represented by the rotation $\mathbf{R} = \mathbf{R}(Z)$, which uniquely defines its Mueller matrix. In an experimental framework S_{in} should be known to us. Here, it is assumed to be a pure state of polarization such that S_{in} takes the value $(0, 0, 1)^T$ with probability one (left circular polarization). In this case, the probability density of S_{out} is given by the formula⁹ (48):

$$p_{\text{out}}(s) = p_{S_{\text{out}}}(s) = \sum_{l \in \mathbb{N}} (2l + 1) e^{\frac{-\mu^2}{2} l(l+1)Z} Y_0^l(s) \quad (50)$$

where we have replaced the values of $\xi^l(0)$ corresponding to S_{in} , and the degree of polarization is given by – using formula (49):

$$P = P^1 = e^{-\mu^2 Z} \quad (51)$$

Formula (50) gives the probability density function of S_{out} . This probability density function contains μ^2 as a parameter. A standard way of estimating μ^2 is, for instance, maximum-likelihood estimation [20]. Using formula (50), in order to find an analytical expression of the maximum of likelihood estimator of μ^2 , is not a straightforward task. Still, an exact maximum-likelihood estimator of μ^2 can be found numerically in a standard way. When $\mu^2 Z$ is small, the maximum-likelihood estimator of μ^2 can be approximated with the following estimator [21, 22]:

$$\mu^2 \approx \frac{1}{2ZN} \sum_{i=1}^{i=N} \theta_i^2 \quad (52)$$

This estimator is evaluated from realizations S_i – with $i = 1, 2, \dots, N$ – of the random variable S_{out} . Here θ_i is the polar angle of S_i , that is, its angular distance, on the Poincaré sphere, to the initial value $S_{\text{in}} = (0, 0, 1)^T$. The estimator is thus based on the empirical mean of the squared angular distance between the initial state S_{in} and each realization of S_{out} . An estimator similar to this one is used in [15] for the standard deviation of speckle noise. Another approach to the estimation of μ^2 is to estimate the degree of polarization and use Equation (51) to retrieve μ^2 . According to our definition of the degree of polarization, formula (25), P can be estimated as:

$$P \approx \left\| \frac{1}{N} \sum_{i=1}^{i=N} S_i \right\| \quad (53)$$

The degree of polarization can also be estimated from intensity measurements [23].

The example treated in this subsection outlines a certain methodology for the application of our formalism to the characterization of an underlying physical system, using the higher-order statistics of a wavefield that has interacted with it. More concrete examples of such use of higher-order statistics are given in [24–27]. In [9], an experimental setup is described that allows the recording of an ensemble of realizations of a random Stokes vector. In our case, these would be the realizations S_i we need in order to evaluate the estimators (52) or (53). Such an ensemble of realizations allows the practical evaluation of the estimators obtained, for the physical characteristics of the propagation medium, in any application of our model similar to the one described in this section. It also allows the measurement (in fact, the evaluation from

⁹This formula was first given by F. Perrin in 1928 [19].

measurements) of our main statistical descriptors, the generalized Stokes vectors S^l or their complex counterparts ξ^l . In [9], this experimental setup is used to construct the probability density function on the Poincaré sphere, of a random Stokes vector. If the vectors ξ^l have been estimated, the probability density function can alternatively be recovered using Equation (3) – see [16] for a detailed study of this method. It appears that while the higher-order statistics of a non-Gaussian wavefield are not related to the observables of this wavefield, they are often characteristic of physical systems that have interacted with it [9, 24–27]. On the contrary, second order statistics are related to observables but provide little information about the underlying physical system.

7. Conclusions and outlook

This article was aimed at presenting a new model for the propagation of polarized light in random birefringent media. The physical situations to which this model would be applicable are bounded by the hypotheses introduced in Section 2. This model was intended to be adapted to a detailed statistical treatment of the physical problems it describes. This is done by including higher-order statistics of the reduced Stokes vector and by describing the random medium via a certain type of statistical input–output relation. It was argued that this type of relation, as introduced in Section 4, can accommodate a variety of physical models and make it easier to estimate the physical parameters appearing in these models. The model that was presented is based on a decomposition of the higher-order statistics of the reduced Stokes vector along the irreducible representations of the rotation group $SO(3)$, which is the group giving the action of a birefringent medium on the Poincaré sphere. In Section 5.1, this decomposition was used to generalize the notions of reduced Stokes vector and degree of polarization to higher-order statistics. This generalization was discussed in relation to recent work studying the role of higher-order statistics in polarization optics. In addition to these more theoretical results, the decomposition along irreducible representation was used, Section 4, to give the laws of transformation, i.e. the input–output relations, for the higher-order statistics of the reduced Stokes vector of a lightwave propagated through a birefringent random medium. These laws of transformation arise mathematically from the spherical convolution theorem.

In Section 6, the evolution of the state of polarization of a lightwave propagating in a random birefringent medium was studied. It was modelled using the concept of *Lévy processes* on the rotation group, see Section 6.1. The framework of Lévy processes on the rotation group was used to give an analytical expression of the probability density on the Poincaré sphere, representing the state of polarization after propagation over any distance in the medium. In Section 6.2, the Lévy process model was used to study the depolarization of a lightwave by propagation in a random birefringent medium. An analytic law for the evolution of the degree of polarization under the effect of propagation in the medium was given. It was shown that depolarization takes place independently of initial conditions. In Section 6.3, an example of propagation in optical fibres was discussed. In particular, the probability density on the Poincaré sphere, representing the state of polarization of a lightwave propagating in an optical fibre affected by PMD, was given in an analytical form. The estimation of the physical parameters appearing in this probability density, from observations of the reduced Stokes vector, was discussed.

An important question to address is how the model presented in this paper, which is specific to birefringent random media, can be generalized to any random medium. Such a general random medium acts on the Poincaré sphere by nonlinear transformations. This makes it difficult to model using the Poincaré sphere formalism. However, if we use the complete Stokes formalism (i.e. with all four components of the Stokes vector), then the medium acts on the Stokes vector essentially by Lorentz transformations [28, 29]. These are, of course, linear transformations.

In this article, we have deduced the laws of transformation of the higher-order statistics of the reduced Stokes vector by using the irreducible representations of the rotation group. In the case of a general random medium, it is possible to do the same for the higher-order statistics of the complete Stokes vector. The finite-dimensional irreducible representations of the Lorentz group (which contains the rotation group as a subgroup) should then be used. The finite-dimensional irreducible representations of the Lorentz group are known as spinor representations [30]. All the main features of the model presented in this paper can be, in this way, generalized to any random medium. However, some additional technical difficulties would arise, since the Lorentz group, unlike the rotation group, is not compact.

The general idea of the model we have presented is to group as much *a priori* knowledge as possible, about polarized light in random birefringent media, in one consistent statistical framework taking into account the higher-order statistics of the wavefield. In other words, to construct a general signal model for the state of polarization of a lightwave in a random birefringent medium. This signal model would allow the extraction of significant physical information, in a variety of practical problems, using different signal processing techniques such as detection, estimation, filtering, etc. In this paper, we have presented our signal model and given a toy example of how it can be used. We hope to demonstrate the usefulness of this model by applying it to concrete problems in future work.

References

- [1] J. Yang, W.L. Kath, and C.R. Menyuk, *Polarization mode dispersion probability distribution for arbitrary distances*, Optic. Lett. 26 (2001), pp. 1472–1474.
- [2] A. Vanucci and A. Bononi, *Statistical characterization of the Jones matrix of long fibers affected by polarization mode dispersion (PMD)*, J. Lightwave Tech. 20 (2002), pp. 811–821.
- [3] M. Karlsson, *Geometrical interpretation of second order PMD*, J. Lightwave Technol. 24 (2006), pp. 643–651.
- [4] J.P. Gordon and H. Kogelnik, *PMD fundamentals: Polarization mode dispersion in optical fibers*, Proc. Natnl. Acad. Sci. USA 97 (2000), pp. 4541–4550.
- [5] C. Brosseau, *Fundamentals of Polarized Light: A Statistical Approach*, John Wiley, New York, 1998.
- [6] G.S. Chirikjian and A.B. Kyatkin, *Engineering Applications of Non-commutative Harmonic Analysis*, CRC Press, 2000.
- [7] M.E. Taylor, *Noncommutative Harmonic Analysis*, American Mathematical Society, 1986.
- [8] N.Ya. Vilenkin and A.U. Klimyk, *Fonctions spéciales et théorie des représentations des groupes*, (Dunod, Collection Travaux et recherches mathématiques, 1969).
- [9] J. Ellis and A. Dogariu, *Discrimination of globally unpolarized fields through Stokes vector element correlations*, J. Opt. Soc. Am. 22 (2005), pp. 491–496.
- [10] Ph. Réfrégier, *Polarization degree of optical waves with non-Gaussian probability density functions: Kullback relative entropy-based approach*, Optic. Lett. 30 (2005), pp. 1090–1092.
- [11] Ph. Réfrégier and F. Goudail, *Kullback relative entropy and characterization of partially polarized optical waves*, J. Opt. Soc. Am. 23 (2006), pp. 671–678.
- [12] A. Luis, *Degree of polarization in quantum optics*, Phys. Rev. A (2002).
- [13] J. Ellis and A. Dogariu, *Differentiation of globally unpolarized complex random fields*, J. Opt. Soc. Am. 21 (2004), pp. 988–983.
- [14] K. Kim, L. Mandel, and E. Wolf, *Relationship between Jones and Mueller matrices for random media*, J. Opt. Soc. Am. 4 (1987), pp. 433–437.
- [15] J. Park, N.J. Kemp, H.N. Zaatari, H.G. Rylander, and T.E. Milner, *Differential geometry of normalized Stokes vector trajectories in anisotropic media*, J. Opt. Soc. Am. 23 (2006), pp. 679–690.
- [16] P.T. Kim and J.Y. Koo, *Optimal spherical deconvolution*, J. Multivariate Anal. (1999), pp. 21–42.
- [17] M. Liao, *Lévy Processes in Lie Groups*, Cambridge University Press, 2004.
- [18] E.P. Hsu, *Stochastic Analysis on Manifolds*, American Mathematical Society, 2002.
- [19] F. Perrin, *Etude mathématique du mouvement brownien de rotation*, Annales scientifiques de l'ENS, tome 45 (1928).

- [20] S.M. Kay, *Fundamentals of Statistical Signal Processing*, Volume II: Estimation Theory, Prentice Hall, 1998.
- [21] P.T. Fletcher, S. Joshi, C. Lu, and S. Pizer, *Gaussian distributions on Lie groups and their applications to statistical shape analysis*, in *Information Processing in Medical Imaging* (2003), pp. 450–462.
- [22] X. Pennec, *Probabilities and Statistics on Riemannian Manifolds: A Geometric Approach*, INRIA Research Report No. 5093, (2004).
- [23] Ph. Réfrégier, J. Fade, and M Roche, *Estimation precision of the degree of polarization from a single intensity image*, *Optic. Lett.* 32 (2007), pp. 739–741.
- [24] B. Ruffing and J. Fleischer, *Spectral correlation of partially or fully developed speckle patterns generated by rough surfaces*, *J. Opt. Soc. Am.* 2 (1985), pp. 1637–1643.
- [25] J. Ohtsubo and T. Asakura, *Measurement of surface roughness using speckle patterns with non-Gaussian statistics*, *Opt. Commun.* 25 (1978), pp. 1742–1753.
- [26] E. Jakeman, *Polarization characteristics of non-Gaussian scattering by small particles*, *Waves in Random Media* 5 (1995), pp. 427–442.
- [27] E.M. Ortiz, F. Gonzales, and F. Moreno, *Intensity statistics of the light scattered from particulate surfaces: Interacting particles and non-Gaussian effects*, *Optic. Commun.* 181 (2000), pp. 231–238.
- [28] Sudha and A.V. Gopala Rao, *Polarization elements, a group theoretic study*, *J. Opt. Soc. Am.* 18 (2001), pp. 3130–3134.
- [29] D.G.M. Anderson and R. Barakat, *Necessary and sufficient conditions for a Mueller matrix to be derivable from a Jones matrix*, *J. Opt. Soc. Am.* 11 (1994), pp. 2305–2319.
- [30] M.A. Naimark, *Linear Representations of the Lorentz Group*, Pergamon Press, 1964.

Decompounding on Compact Lie Groups

Salem Said, Christian Lageman, Nicolas Le Bihan, and Jonathan H. Manton, *Senior Member, IEEE*

Abstract—Noncommutative harmonic analysis is used to solve a nonparametric estimation problem stated in terms of compound Poisson processes on compact Lie groups. This problem of *decompounding* is a generalization of a similar classical problem. The proposed solution is based on a characteristic function method. The treated problem is important to recent models of the physical inverse problem of multiple scattering.

Index Terms—Compact Lie groups, compound Poisson processes, multiple scattering, noncommutative harmonic analysis, nonparametric estimation.

I. INTRODUCTION

THIS paper studies the following nonparametric estimation problem. Let $(X_n)_{n \geq 1}$ be *i.i.d.* G -valued random variables for some group G , and let e denote the identity element of G . For example, G might be the group of 3×3 orthogonal matrices, in which case each X_n would be a random 3×3 orthogonal matrix and e would be the 3×3 identity matrix. The process

$$Y(t) = \prod_{n=0}^{N(t)} X_n, \quad X_0 = e$$

where $N = (N(t))_{t \geq 0}$ is a Poisson process with parameter $\lambda > 0$, is called a G -valued compound Poisson process. If G is not commutative, the above products are taken to be ordered from left to right, and $Y(t)$ is called a *left* compound Poisson process. It is assumed that the random variables X_n and $N(t)$ are independent of each other, and for simplicity, it is further assumed that the Poisson parameter λ is known. The general problem is to estimate the distribution of the X_n given partial observations of one or more realisations of the compound Poisson process $Y(t)$. Of specific interest, is the case when multiple realisations of $Y(T)$ are available, for some fixed time instant $T > 0$.

Manuscript received July 13, 2009; revised January 06, 2010. Current version published May 19, 2010. Parts of this research were carried out while C. Lageman was at the Australian National University and supported by the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems.

S. Said and N. Le Bihan are with the GIPSA-Lab, Department of Images and Signal, Grenoble, France (e-mail: salem.said@gipsa-lab.grenoble-inp.fr; Bihan@gipsa-lab.grenoble-inp.fr).

C. Lageman is with the Department of Mathematics, University of Würzburg, Am Hubland, 97074 Würzburg, Germany (e-mail: christian.lageman@mathematik.uni-wuerzburg.de).

J. H. Manton is with the University of Melbourne, Department of Electrical and Electronic Engineering, Victoria 3010, Australia (e-mail: jmanton@unimelb.edu.au).

Communicated by H. Bölcskei, Associate Editor for Detection and Estimation.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIT.2010.2046216

The real numbers form a group, with addition being the group operation. Choosing G to be this group results in the ordinary compound Poisson process $y(t) = \sum_{n=0}^{N(t)} x_n$ where $x_0 = 0$ and x_n for $n \geq 1$ are real-valued *i.i.d.* random variables. Estimating the distribution of the x_n is known as decompounding and has been well-studied [1], [2]. In the present paper, decompounding techniques are extended to the case when G is a noncommutative group. This new case can not be obtained trivially and requires ideas from noncommutative harmonic analysis. Although group-valued compound Poisson processes were introduced by Applebaum in [3], the corresponding decompounding problem has not been addressed in generality before.

This paper contributes to the relatively recent trend consisting in the application of noncommutative harmonic analysis (*i.e.*, harmonic analysis on groups) to estimation and inverse problems. It addresses a nonparametric estimation problem stated in terms of compound Poisson processes on compact Lie groups. We refer to this as the problem of *decompounding* on compact Lie groups, since it directly generalizes the classical problem of decompounding for scalar processes. This generalization is mathematically natural and is motivated by the physical inverse problem of multiple scattering. In particular, this paper also contributes to the modelling of multiple scattering using compound Poisson processes.

Compound Poisson processes model the accumulation of rare events. As such, scalar compound Poisson processes are important tools in queuing and traffic problems and in risk theory. The classical problem of decompounding arises in the context of these processes. A functional approach to this problem is given by Buchmann and Grübel [1]. A characteristic function method is studied by van Es *et al.* [2]. The applications of decompounding in queuing problems and risk theory are referenced in [1]. We extend this problem by considering decompounding on compact Lie groups. We approach this new problem by using noncommutative harmonic analysis to generalize the above mentioned method of [2].

The important potential which noncommutative harmonic analysis holds for engineering problems is well illustrated in the book of Chirikjian and Kyatkin [4]. Its importance to nonparametric estimation stems from the fact that it leads to the successful generalization of the highly important concept of characteristic function in probability. In mathematical research, this generalization was pioneered by Grenander [5] and extensively developed by Heyer [6]. It has received special attention in the engineering community. See Yazici [7] and the papers by Kim *et al.* [8]–[11].

The paper is organized as follows. Section II sets down the necessary background in harmonic analysis and characteristic functions on compact Lie groups. Section III introduces compound Poisson processes on compact Lie groups. In Section IV, we state the decompounding problem for these processes and

present our approach based on noncommutative harmonic analysis. In Section V, we propose a model for multiple scattering based on compound Poisson processes on the rotation group $SO(3)$. Within this model, decomposing appears as a physical inverse problem. We apply our approach as described in Section IV to this problem using numerical simulations.

II. CHARACTERISTIC FUNCTIONS ON COMPACT LIE GROUPS

Characteristic functions of scalar and vector-valued random variables are defined using the usual Fourier transform. Their extension to random variables with values on compact Lie groups owes to the tools of harmonic analysis on these groups. Our presentation of characteristic functions is adapted from [5], [12]. Harmonic analysis on compact Lie groups is presented in more detail in recent papers [8], [7]. More thorough classical references thereon include [13], [14].

Let G be a compact connected Lie group with identity e . We denote by μ the biinvariant normalized Haar measure on G . Hilbert spaces of square integrable (with respect to μ) complex and real-valued functions on G are denoted $L^2(G, \mathbb{C})$ and $L^2(G, \mathbb{R})$. A representation of G is a continuous homomorphism $\pi : G \rightarrow GL(V)$ with V a complex Hilbert space and $GL(V)$ the group of invertible bounded linear maps of V . It is called irreducible if any G -invariant subspace of V is trivial i.e., equals $\{0\}$ or V . Two representations $\pi_i : G \rightarrow GL(V_i)$ with $i = 1, 2$ are called equivalent if there exists an invertible bounded linear map $L : V_1 \rightarrow V_2$ such that $L \circ \pi_1(g) = \pi_2(g) \circ L$ for all $g \in G$. Using this relation, the set of irreducible representations of G is partitioned into equivalence classes.

The central result of harmonic analysis on compact groups is the Peter–Weyl theorem. For the current context, it can be stated as follows. Let $\text{Irr}(G)$ be the set of equivalence classes of irreducible representations of G . $\text{Irr}(G)$ is a countable set. If $\delta \in \text{Irr}(G)$ then we have the two following facts. All representations of the class δ have the same finite dimension d_δ . There exists in this class a unitary representation U^δ . Choosing one such representation we can suppose that $U^\delta : G \rightarrow SU(\mathbb{C}^{d_\delta})$ with $SU(\mathbb{C}^{d_\delta})$ the group of special unitary $d_\delta \times d_\delta$ matrices. We distinguish the unit representation $\delta_0 \in \text{Irr}(G)$ where $U^{\delta_0}(g) = 1$ for all $g \in G$. With this choice being fixed, we can state the Peter–Weyl theorem.

Theorem 1 (Peter–Weyl): The functions $d_\delta^{1/2} U_{ij}^\delta$, taken for $\delta \in \text{Irr}(G)$ and $i, j = 1, \dots, d_\delta$ form an orthonormal basis of $L^2(G, \mathbb{C})$.

Note that U_{ij}^δ is the usual notation for the matrix elements of U^δ . For all $f \in L^2(G, \mathbb{C})$ the theorem gives the Fourier pair

$$A_\delta = \int f(g) U^\delta(g)^\dagger d\mu(g) \quad (1)$$

$$f(g) = \sum_{\delta \in \text{Irr}(G)} d_\delta \text{tr}(A_\delta U^\delta(g)) \quad (2)$$

where † denotes the Hermitian conjugate and tr the trace. The Fourier series (2) converges in $L^2(G, \mathbb{C})$.

Consider the example $G = S^1$. It is possible to make the identification $\delta = 0, 1, \dots$. Then $U^\delta(z) = z^\delta$ for $z \in S^1$.

Writing $z = e^{i\theta}$ for some $\theta \in [0, 2\pi]$, this gives the classical Fourier expansion of periodic functions.

We consider random objects and in particular G -valued random variables defined on a suitable probability space $(\Omega, \mathcal{A}, \mathbb{P})$. When referring to the probability density of such a random variable X , we mean a probability density $p_X \in L^2(G, \mathbb{R})$ with respect to μ . The characteristic function of a G -valued random variable is defined as follows. Compare to [5].

Definition 1: Let X be a G -valued random variable. The characteristic function of X is the map ϕ_X given by

$$\delta \mapsto \phi_X(\delta) = \mathbb{E}(U^\delta(X)) \quad \delta \in \text{Irr}(G).$$

Here, $\mathbb{E}$ stands for expectation on the underlying probability space. For all $\delta \in \text{Irr}(G)$, the expectation in the definition is finite since U^δ has unitary values. When X has a probability density p_X its characteristic function gives the Fourier coefficients of p_X as in (1). We have

$$\phi_X(\delta) = \mathbb{E}(U^\delta(X)) = \int p(g) U^\delta(g) d\mu(g) \quad \delta \in \text{Irr}(G).$$

The following proposition 1 recalls the relation between characteristic functions and the concepts of convolution and convergence in distribution. It is a generalization of classical properties for scalar random variables. Remember that a sequence $(X_n)_{n \geq 1}$ of G -valued random variables is said to converge in distribution to a random variable X if for all real-valued continuous function f on G , we have

$$\lim_n \mathbb{E}(f(X_n)) = \mathbb{E}(f(X)).$$

The proof of proposition 1 is straightforward. See [5].

Proposition 1: The following two properties hold.

- 1) Let X and Y be independent G -valued random variables and let $Z = XY$. We have for all $\delta \in \text{Irr}(G)$

$$\phi_Z(\delta) = \phi_X(\delta) \phi_Y(\delta).$$

- 2) A sequence $(X_n)_{n \geq 1}$ of G -valued random variables converges in distribution to a random variable X iff for all $\delta \in \text{Irr}(G)$

$$\lim_n \phi_{X_n}(\delta) = \phi_X(\delta).$$

In order to solve our estimation problem in Section IV we will require random variables to have certain symmetry properties. We deal with these properties here. The following analysis draws on Liao [12], [15].

We will say that a G -valued random variable X is inverse invariant if $X \stackrel{d}{=} X^{-1}$. We will say that it is conjugate invariant if for all $k \in G$ we have that $X \stackrel{d}{=} kXk^{-1}$. As usual $\stackrel{d}{=}$ denotes equality in distribution. The following proposition 2 characterizes these two symmetry properties in terms of characteristic functions. It will be important to remember that for any two G -valued random variables X and Y we have $X \stackrel{d}{=} Y$ iff

$\phi_X = \phi_Y$. This results from the completeness of the basis given by the U^δ as stated in the Peter–Weyl theorem [5].

Proposition 2: The following properties hold.

- 1) X is inverse invariant *iff* for all $\delta \in \text{Irr}(G)$ we have that $\phi_X(\delta)$ is Hermitian.
- 2) Let X be inverse invariant. If $X_1, \dots, X_n$ are independent copies of X then the product $X_1 \dots X_n$ is inverse invariant.
- 3) X is conjugate invariant *iff* for all $\delta \in \text{Irr}(G)$ we have that $\phi_X(\delta) = a_\delta I_{d_\delta}$ where $a_\delta \in \mathbb{C}$ and I_{d_δ} is the $d_\delta \times d_\delta$ identity matrix.
- 4) If X and Y are independent and conjugate invariant then XY is conjugate invariant.
- 5) X is conjugate invariant *iff* for all G -valued random variable Y independent of X we have $XY \stackrel{d}{=} YX$.

Proof:

- 1) Note that for all $\delta \in \text{Irr}(G)$ we have by the homomorphism property of U^δ and the fact that it has unitary values

$$\phi_{X^{-1}}(\delta) = \mathbb{E}(U^\delta(X^{-1})) = \mathbb{E}(U^\delta(X))^\dagger = \phi_X(\delta)^\dagger.$$

- 2) This follows from 1 of proposition 2 and 1 of proposition 1, since the powers of a Hermitian matrix are Hermitian.
- 3) Note that for all $k \in G$ we have that $X \stackrel{d}{=} kXk^{-1}$ *iff* for all $\delta \in \text{Irr}(G)$

$$\mathbb{E}(U^\delta(X)) = \mathbb{E}(U^\delta(kXk^{-1})) = U^\delta(k)\mathbb{E}(U^\delta(X))U^\delta(k)^\dagger$$

identifying ϕ_X on both sides, this becomes

$$\phi_X(\delta) = U^\delta(k)\phi_X(\delta)U^\delta(k)^\dagger.$$

If this relation is verified for all $k \in G$ then $\phi_X(\delta)$ is a multiple of I_{d_δ} . This follows by Schur's lemma [13].

- 4) This follows from 3 of proposition 2 and 1 of proposition 1.
- 5) The *if* part follows by setting $Y = k \in G$ for arbitrary k . The *only if* part follows from 3 of proposition 2 and 1 of proposition 1. ■

Article 1 of proposition 2 motivates a practical recipe for generating inverse invariant random variables from general random variables. Let X and Y be G -valued random variables. Suppose X and Y are independent with $Y \stackrel{d}{=} X^{-1}$. It can be verified by 1 of proposition 2 that $XY \stackrel{d}{=} YX$ and that both these products are inverse invariant. In practice, if we have generated X then we can immediately generate Y as above. In this way, an inverse invariant XY or YX is generated from X .

III. COMPOUND POISSON PROCESSES

Compound Poisson processes on groups naturally generalize scalar compound Poisson processes. They are introduced by Applebaum in [3]. Let us start by recalling the definition of scalar compound Poisson processes. Let $N = (N(t))_{t \geq 0}$ be a Poisson process with parameter $\lambda > 0$. Suppose $(x_n)_{n \geq 1}$ are *i.i.d.* $\mathbb{R}$ -valued random variables. Suppose the family $(x_n)_{n \geq 1}$

is itself independent of N . The following process y is said to be a compound Poisson process

$$y(t) = \sum_{n=0}^{N(t)} x_n$$

G -valued compound Poisson processes are defined by analogy to this formula. We continue with the process N . Let $(X_n)_{n \geq 1}$ be *i.i.d.* G -valued random variables and suppose as before that the family $(X_n)_{n \geq 1}$ is independent of N . The following process Y is said to be a G -valued left compound Poisson process

$$Y(t) = \prod_{n=0}^{N(t)} X_n.$$

We understand that products are ordered from left to right. It is possible to obtain a right compound Poisson process by considering $Y(t)^{-1}$ instead. Thus, the two concepts are equivalent. See [12] and [3].

Before going on, we make the following remark on the above definition of compound Poisson processes. This definition was stated for G a compact connected Lie group. This topological and manifold structure of G is not necessary for the definition, which can be stated in its above form for any group with a measurable space structure. The compact group structure of G allows us to use the Peter–Weyl theorem and characteristic functions. The connected Lie group structure allows the introduction of Brownian noise in Section IV, see [12].

We wish to summarize the symmetry properties of the random variables $Y(t)$ for $t \geq 0$. Note first that for all $t \geq 0$, $Y(t)$ does not have a probability density. Indeed, for all $t \geq 0$ we have $\mathbb{P}(Y(t) = e) \geq \mathbb{P}(N(t) = 0) = e^{-\lambda t}$. It follows that $Y(t)$ has an atom at e . In the absence of a probability density, we study $Y(t)$ for $t \geq 0$ using its characteristic function. This is given in the following Proposition 3 which can be seen to immediately extend the well known formula for scalar compound Poisson processes. This proposition follows [12], [3].

Proposition 3: For all $t \geq 0$ the characteristic function $\phi_{Y(t)}$ of $Y(t)$ is given by

$$\phi_{Y(t)}(\delta) = \exp(\lambda t(\phi_X(\delta) - I_{d_\delta})) \quad (3)$$

for $\delta \in \text{Irr}(G)$, where $\phi_X \equiv \phi_{X_1}$.

Proof: Let $t \geq 0$. $\phi_{Y(t)}$ can be calculated by conditioning over the values of $N(t)$. Using the independence of N and $(X_n)_{n \geq 1}$ we have for $\delta \in \text{Irr}(G)$

$$\phi_{Y(t)}(\delta) = e^{-\lambda t} \sum_{n=0}^{\infty} \frac{(\lambda t)^n}{n!} \mathbb{E} \prod_{m=0}^n U^\delta(X_m).$$

Using the fact that $(X_n)_{n \geq 1}$ are *i.i.d.* it is possible to replace

$$\mathbb{E} \prod_{m=0}^n U^\delta(X_m) = \prod_{m=0}^n \mathbb{E}(U^\delta(X_m)) = \phi_X(\delta)^n$$

the proposition follows by rearranging the sum. ■

Combining Propositions 3 and 2, we have the following proposition. It states that for all $t \geq 0$ the symmetry properties of $Y(t)$ are the same as those of the X_n .

Proposition 4: For all $t \geq 0$ we have:

- 1) if X_1 is inverse invariant, then so is $Y(t)$;
- 2) if X_1 is conjugate invariant, then so is $Y(t)$.

We end this section with Proposition 5. It gives a property of uniformization of the distribution of $Y(t)$ as $t \uparrow \infty$. This is similar to the behavior of the products $X_1 \dots X_n$ for $n \uparrow \infty$, see [5]. For a more general version of Proposition 5 see [12], [15]. We say that a G -valued random variable X is supported by a measurable subset S of G if $\mathbb{P}(X \in S) = 1$. If X and X' are G -valued random variables with $X \stackrel{d}{=} X'$ then X is supported by S iff X' is supported by S . In Proposition 5, U is a G -valued random variable with probability density identically equal to 1. That is, U is uniformly distributed on G .

Proposition 5: If X_1 is not supported by any closed proper subgroup S of G or coset gS , $g \in G$ of such a subgroup then $Y(t)$ converges in distribution to U as $t \uparrow \infty$.

Proof: Under the conditions of the proposition we have for all $\delta \neq \delta_0$ that the eigenvalues of $\phi_X(\delta)$ are all < 1 in modulus [5]. It follows that the eigenvalues of $\phi_X(\delta) - I_{d_\delta}$ all have negative real parts. Thus, when $\delta \neq \delta_0$ we have by (3) that $\phi_{Y(t)}(\delta) \rightarrow 0$ as $t \uparrow \infty$. Moreover, it is immediate that $\phi_{Y(t)}(\delta_0) = 1$ for $t \geq 0$. We conclude using 2 of Proposition 1. Note that [13]

$$\phi_U(\delta) = \int U^\delta(g) d\mu(g) = 0 \quad \delta \neq \delta_0$$

and $\phi_U(\delta_0) = 1$ trivially. ■

IV. DECOMPOUNDING

In existing literature, *decompounding* refers to a set of nonparametric estimation problems involving scalar compound Poisson processes [1], [2]. In this section, we will consider the generalization of these problems to compound Poisson processes on compact Lie groups. The new problems can be stated in the notation of Section III. We refer to them also as decompounding problems. As in the scalar case, they consist in estimation of the common probability density (supposed to exist) of the random variables X_n from observations of the process Y . The unknown common probability density of the X_n will be denoted p . We are unaware of any work on similar problems for vector-valued compound Poisson processes. Our consideration of compact Lie groups is motivated by the applications presented in Section V.

A. Typology of Decompounding Problems

Several decompounding problems can be stated, depending on the nature of the observations made of Y [2]. Decompounding is performed from *high-frequency* observations if an individual trajectory of the process Y is observed over time intervals $[0, T]$ where $T \uparrow \infty$. It is performed from *low-frequency* observations if *i.i.d.* observations are made of the random variable $Y(T)$ for a fixed $T \geq 0$.

Decompounding from high- and low-frequency observations lead to different difficulties. For high-frequency observations,

the problem is greatly simplified if the assumption is made that X_n does not take the value e , for any $n \geq 1$. With probability 1, a trajectory of N has infinitely many jumps over $t \geq 0$. Under the assumption we have made, all these jumps correspond to jumps of Y which we do observe. The jumps of Y then give *i.i.d.* observations of X_1 and the average time between these jumps is $1/\lambda$. In particular, it is important for high-frequency observations to take the limit $T \uparrow \infty$.

Low frequency observations do not give direct access to λ . In scalar decompounding from low-frequency observations, λ is often assumed to be known [1], [2]. In the context of a compact group G , Proposition 5 leads to a difficulty that does not appear in scalar decompounding. Under the conditions of this proposition, if low-frequency observations are made at a sufficiently large time T then these observations will be uniformly distributed on G and will have no memory of the random variables X_n .

A third intermediate type of observations is possible. It is possible to make observations of an individual trajectory of Y at regular time intervals $T, 2T, \dots$. This is in fact equivalent to low-frequency distributions. Remember that N is a Lévy process, *i.e.*, has independent stationary increments. Moreover we have that the $(X_n)_{n \geq 1}$ are *i.i.d.* Using this, it is possible to prove that the G -valued random variables

$$Y(T), Y(T)^{-1}Y(2T), Y(2T)^{-1}Y(3T) \dots$$

are *i.i.d.*. Thus, our observations are *i.i.d.* observations of $Y(T)$. This remark refers to the fact that Y is a left Lévy process in G [12]. We do not develop this here.

B. Noise Model for Low Frequency Observations

We will consider decompounding from low-frequency observations. $T \geq 0$ is fixed and *i.i.d.* observations $(Z_n)_{n \geq 1}$ of a noisy version Z of $Y(T)$ are available. Z is given by Y corrupted by multiplicative noise. We have the noise model

$$Z = MY(T) \tag{4}$$

where M is independent of Y . By 1 of Proposition 1 we have for the characteristic function of Z

$$\phi_Z = \phi_M \phi_{Y(T)}.$$

The noise model is equivalent to having an initial value $Y(0) = M$ with a general distribution. We consider the case of Brownian noise. The characteristic function of M is then given by [12], [8]

$$\phi_M(\delta) = \exp\left(-\lambda_\delta \frac{\sigma^2}{2}\right) I_{d_\delta}$$

where σ^2 is a variance parameter and for $\delta \in \text{Irr}(G)$ the constant λ_δ is the corresponding eigenvalue of the Laplace–Beltrami operator. In particular, $\lambda_{\delta_0} = 0$ and $\lambda_\delta > 0$ for $\delta \neq \delta_0$. It is clear from 3 of Proposition 2 that M is conjugate invariant. It follows by 4 of Proposition 2 that, as far as the distribution of Z is concerned, left and right multiplication of $Y(T)$ by the noise M are indifferent.

It is possible to construct a G -valued process ζ such that $Z \stackrel{d}{=} \zeta(T)$. The corresponding construction is well known in

the theory of group-valued Lévy processes and is referred to as interlacing [3], [12]. Here we only state this construction. Let W be a Brownian motion on G independent of N and with variance parameter $\bar{\sigma}^2$. This is a process with continuous paths and independent stationary increments. Moreover, $W(0) = e$ and for $\delta \in \text{Irr}(G)$

$$\phi_{W(t)}(\delta) = \exp\left(-\lambda_\delta \frac{\bar{\sigma}^2}{2} t\right) I_{d_\delta}.$$

Let $T_0 = 0$ and suppose $(T_n)_{n \geq 1}$ are the jump times of N . The interlaced process ζ is defined as follows. We have $\zeta(0) = e$. For $t > 0$ and $n \geq 1$ we have

$$\zeta(t) = \zeta(T_{n-1})W(T_{n-1})^{-1}W(t) \text{ on } \{T_{n-1} \leq t < T_n\}$$

where the following formula holds at each time T_n (here, $\zeta(T_n-)$ denotes the left limit at T_n):

$$\zeta(T_n) = \zeta(T_n-)X_n.$$

This definition is sufficient, since $T_n \uparrow \infty$ almost surely. The term interlacing comes from the fact that the trajectories of ζ are obtained by introducing the jumps of Y into the trajectories of W as these jumps occur. The trajectories of W are thus interlaced with the jumps of Y .

For $t \geq 0$ the characteristic function of $\zeta(t)$ is given by

$$\phi_{\zeta(t)}(\delta) = \exp\left(t\lambda\phi_X(\delta) - tI_{d_\delta}\left(\lambda + \frac{\lambda_\delta\bar{\sigma}^2}{2}\right)\right) \quad (5)$$

for $\delta \in \text{Irr}(G)$. It follows that we have $Z \stackrel{d}{=} \zeta(T)$ if $T\bar{\sigma}^2 = \sigma^2$.

Although we do not deal with the case of high-frequency observations, we would like to end this subsection with a remark on the role of noise in this case. The trajectories of the interlaced process ζ are noisy versions of the trajectories of Y . However, these trajectories have the same jumps as the trajectories of Y . In this sense, high-frequency observations are unaltered by noise.

C. Characteristic Function Method

We present a characteristic function method for decomposing from low-frequency observations. This method extends a similar one considered in [2]. In carrying out this extension, we are guided by the properties of characteristic functions on G presented in Section II. Our observations $(Z_n)_{n \geq 1}$ and noise model (4) were described in Section IV-B. We aim to estimate the common density p of the X_n . A characteristic function method consists in constructing nonparametric estimates for p from parametric estimates for its Fourier coefficients $\phi_X(\delta)$ given for $\delta \in \text{Irr}(G)$. See [8].

We suppose that λ and σ^2 are known. Equation (5) can be copied as follows:

$$\phi_Z(\delta) = \exp\left(T\lambda\phi_X(\delta) - T\bar{\lambda}I_{d_\delta}\right) \quad \delta \in \text{Irr}(G) \quad (6)$$

where $\bar{\lambda}$ is a constant determined by λ and σ^2 . We refer to this transformation $\phi_X \mapsto \phi_Z$ as the compounding transformation. Decomposing will involve local inversion of the compounding transformation. This is clearly related to inversion

of the matrix exponential in a neighborhood of $\phi_Z(\delta)$ for all $\delta \in \text{Irr}(G)$. Rather than deal with this problem in general, we make the following simplifying assumption.

Assumption: X_1 is inverse invariant.

For all $\delta \in \text{Irr}(G)$ we have by applying 1 of Proposition 2 and (6) to this assumption that $\phi_Z(\delta)$ is Hermitian positive definite. Note Log the unique Hermitian matrix logarithm of a hermitian positive definite matrix. We can now express the inverse of the compounding transformation. From (6), it follows that

$$\phi_X(\delta) = \frac{1}{T\lambda} \text{Log}[\phi_Z(\delta)] + (\bar{\lambda}/\lambda) I_{d_\delta} \quad \delta \in \text{Irr}(G). \quad (7)$$

Let $\delta \in \text{Irr}(G)$. It follows from definition 1 that empirical estimates of $\phi_Z(\delta)$ based on the observations $(Z_n)_{n \geq 1}$ are unbiased and consistent. This is a simple consequence of the strong law of large numbers. See for example [16]. In order to estimate $\phi_X(\delta)$ using (7), it is then important to ensure that the empirical estimates of $\phi_Z(\delta)$ are asymptotically Hermitian positive definite.

We start by defining the empirical estimates $\hat{\phi}_Z^n(\delta)$ for $\delta \in \text{Irr}(G)$ and $n \geq 1$

$$\hat{\phi}_Z^n(\delta) = \frac{1}{2n} \sum_{m=1}^n (U^\delta(Z_m) + U^\delta(Z_m)^\dagger).$$

Hermitian symmetrization of empirical estimates is necessary for the application of (7). Since it is a projection operator, this symmetrization moreover contributes to a faster convergence of the $\hat{\phi}_Z^n(\delta)$ to $\phi_Z(\delta)$.

Continuous dependence of the spectrum of a matrix on its coefficients is a classical result in matrix analysis. Several more or less sophisticated versions of this result exist [17]. For a remarkably straightforward statement, see [18]. For a complex matrix C we will denote $\lambda(C)$ its spectrum. For each $\delta \in \text{Irr}(G)$ and $n \geq 1$ define the event R_δ^n by

$$R_\delta^n = \left\{ \lambda\left(\hat{\phi}_Z^n(\delta)\right) \subset]0, \infty[\right\}.$$

For $\delta \in \text{Irr}(G)$, the sequence $(R_\delta^n)_{n \geq 1}$ controls the convergence of the spectra of the empirical estimates $\hat{\phi}_Z^n(\delta)$. In particular, since $\lim_n \hat{\phi}_Z^n(\delta) = \phi_Z(\delta)$ almost surely

$$\mathbb{P}(\cup_{n \geq 1} \cap_{m \geq n} R_\delta^m) = \lim_n \mathbb{P}(\cap_{m \geq n} R_\delta^m) = 1.$$

Using the events R_δ^n we can write down well defined estimates of ϕ_X . These are denoted $\hat{\phi}_X^n(\delta)$ for $\delta \in \text{Irr}(G)$ and $n \geq 1$

$$\begin{aligned} \hat{\phi}_X^n(\delta) &= 0 \text{ on } \Omega - R_\delta^n \\ \hat{\phi}_X^n(\delta) &= \frac{1}{T\lambda} \text{Log}[\hat{\phi}_Z^n(\delta)] + (\bar{\lambda}/\lambda) I_{d_\delta} \text{ on } R_\delta^n. \end{aligned}$$

This expression gives our parametric estimates for the Fourier coefficients of p . We use them to construct nonparametric estimates based on an expression of the form (2). Let $(\Gamma_l)_{l \geq 1}$ be an increasing sequence of finite subsets $\Gamma_l \subset \text{Irr}(G)$ with the limit $\cup_{l \geq 1} \Gamma_l = \text{Irr}(G) - \{\delta_0\}$. Let $K \geq 0$ and for each $\delta \in \text{Irr}(G)$ define

$$f_\delta = d_\delta e^{-K\lambda_\delta}.$$

For $n \geq 1$ and $l \geq 1$, our nonparametric estimate $\hat{p}_l^n$ is given by

$$\hat{p}_l^n(g) = 1 + \sum_{\delta \in \Gamma_l} f_\delta \text{tr} \left(\hat{\phi}_X^n(\delta) U^\delta(g)^\dagger \right) \quad g \in G. \quad (8)$$

The subscript $l \geq 1$ corresponds to a cutoff or smoothing parameter. Indeed, infinitely many representations are excluded from the sum over Γ_l . A more complete expression of this fact appears in [8]. When $K > 0$ the coefficients f_δ form a convolution mask ensuring that the estimates $\hat{p}_l^n$ can be taken to converge to a smooth probability density. We make this more precise in Section IV-D.

It is usual to rewrite expressions similar to (8) in terms of a group invariant kernel. See [8] and [9]. Such a transformation is not possible here due to the indirect nature of our observations. This is in particular related to the more involved form of the $\hat{\phi}_X^n(\delta)$ as given above.

D. Convergence of Parametric and Nonparametric Estimates

Here, we discuss the convergence of the parametric and nonparametric estimates given in Section IV-C. Our argument is presented in the form of Propositions 6 and 7 below. Proposition 6 gives the consistency of the parametric estimates $\hat{\phi}_X^n(\delta)$. Proposition 7 states a subsequent result for the nonparametric estimates $\hat{p}_l^n$.

For Proposition 6, we will need inequalities (9) and (10). These express stability results for the eigenvalues of Hermitian matrices and for the Hermitian matrix function Log. Let A and B be Hermitian $d \times d$ matrices, for some $d \geq 1$. For $1 \leq i \leq d$, let α_i and β_i be the eigenvalues of A and B , respectively. Suppose they are arranged in nondecreasing order. We have

$$\sum_{i=1}^d (\beta_i - \alpha_i)^2 \leq |B - A|^2 \quad (9)$$

where $|\cdot|$ is the Euclidean matrix norm. This inequality is known as the Wielandt–Hoffman theorem. In [17], it is stated for A and B real symmetric. The general case of Hermitian A and B can be obtained from this statement using a canonical realification isomorphism.

Suppose A and B are positive definite. For our purpose, it is suitable to assume both $\lambda(A)$ and $\lambda(B)$ are contained in an interval $[k, 1]$ for some $k > 0$. Under this assumption we have the following Lipschitz property:

$$|\text{Log}(B) - \text{Log}(A)| \leq \sqrt{d} k^{-2} |B - A|. \quad (10)$$

In order to obtain (10), it is possible to start by expressing $\text{Log}(A)$ as follows:

$$\text{Log}(A) = \int_0^1 (A - I_d)[t(A - I_d) + I_d]^{-1} dt.$$

This expression results from a similar one for the real logarithm applied to each eigenvalue of A . Subtracting the same expression for $\text{Log}(B)$, (10) follows by simple calculations.

Proposition 6: For all $\delta \in \text{Irr}(G)$ we have the limit in probability $\lim_n \hat{\phi}_X^n(\delta) = \phi_X(\delta)$.

Proof: We only need to consider $\delta \neq \delta_0$. Indeed, $\hat{\phi}_X^n(\delta_0) = \phi_X(\delta_0) = 1$ for all $n \geq 1$. Let $\delta \neq \delta_0$, for all $n \geq 1$, we have

$$\left| \hat{\phi}_Z^n(\delta) \right|_{op} \leq \frac{1}{2n} \sum_{m=1}^n |U^\delta(Z_m)|_{op} + |U^\delta(Z_m)^\dagger|_{op} = 1$$

where $|\cdot|_{op}$ is the operator matrix norm. Passing to the limit, we have the same inequality for $\phi_Z(\delta)$. It follows that all eigenvalues of $\hat{\phi}_Z^n(\delta)$ or $\phi_Z(\delta)$ are ≤ 1 . Since $\phi_Z(\delta)$ is positive definite, there exists $k_\delta > 0$ such that $\lambda(\phi_Z(\delta)) \subset [k_\delta, 1]$. For $n \geq 1$, let $\tilde{R}_\delta^n$ be the event

$$\tilde{R}_\delta^n = \left\{ \lambda \left(\hat{\phi}_Z^n(\delta) \right) \subset [k_\delta/2, 1] \right\}.$$

From inequality (9), we have

$$\mathbb{P}(\Omega - \tilde{R}_\delta^n) \leq \mathbb{P} \left(\left| \hat{\phi}_Z^n(\delta) - \phi_Z(\delta) \right| > k_\delta/2 \right).$$

Since $\tilde{R}_\delta^n \subset R_\delta^n$, it follows from inequality (10) that

$$\begin{aligned} \mathbb{P} \left(\left| \hat{\phi}_X^n(\delta) - \phi_X(\delta) \right| > \varepsilon \cap \tilde{R}_\delta^n \right) \\ \leq \mathbb{P} \left(\left| \hat{\phi}_Z^n(\delta) - \phi_Z(\delta) \right| > k_\delta^2 \varepsilon / L \right) \end{aligned}$$

for all $\varepsilon > 0$, where $L = 4\sqrt{d_\delta}/T\lambda$.

The proof can be completed by a usual application of Chebyshev's inequality

$$\mathbb{P} \left(\left| \hat{\phi}_X^n(\delta) - \phi_X(\delta) \right| > \varepsilon \right) \leq \left(\frac{8 + 2L^2/\varepsilon^2}{n} \right) \left(\frac{\sqrt{d_\delta}}{k_\delta^2} \right)^2 \quad (11)$$

for all $\varepsilon > 0$. ■

Proposition 7 relies on Proposition 6 and the Peter–Weyl theorem. It implies the existence of sequences $(\hat{p}_k)_{k \geq 1}$, of nonparametric estimates given by (8), converging to p in probability in $L^2(G, \mathbb{C})$ with any prescribed rate of convergence. Convergence in probability in $L^2(G, \mathbb{C})$ means that the following limit in probability holds:

$$\lim_k \|\hat{p}_k - p\| = 0$$

where $\|\cdot\|$ is the $L^2(G, \mathbb{C})$ norm. It is clear from (8) that for all $k \geq 1$, we have $\hat{p}_k \in L^2(G, \mathbb{C})$. In order to obtain nonparametric estimates in $L^2(G, \mathbb{R})$ and converging to p in the same sense, it is enough to consider the real parts of the $\hat{p}_k$. The following proof of Proposition 7 uses Plancherel's formula as in [8].

Proposition 7: Putting $K = 0$ in (8), we have the limit in probability

$$\lim_l \lim_n \|\hat{p}_l^n - p\| = 0.$$

Proof: For $l \geq 1$, let $p_l \in L^2(G, \mathbb{C})$ be given by

$$p_l(g) = 1 + \sum_{\delta \in \Gamma_l} \text{tr} \left(\phi_X(\delta) U^\delta(g)^\dagger \right)$$

for $g \in G$. By the Peter–Weyl theorem, $\lim_l \|p_l - p\| = 0$. By (8) and Proposition 6, we have $\lim_n \|\hat{p}_l^n - p_l\| = 0$ in probability for all $l \geq 1$. The proposition follows by observing that

$$\|\hat{p}_l^n - p\|^2 = \|\hat{p}_l^n - p_l\|^2 + \|p_l - p\|^2 \quad (12)$$

for all $n, l \geq 1$. ■

Proposition 6 obtained convergence in probability of the parametric estimates $\hat{\phi}_X^n(\delta)$ for all $\delta \in \text{Irr}(G)$. These parametric estimates depend only on the observations. In particular, they can be evaluated without any *a priori* knowledge of p . By introducing such knowledge, it is possible to define parametric estimates $\tilde{\phi}_X^n(\delta)$ converging in the square mean to the same limits $\phi_X(\delta)$. For $\delta \in \text{Irr}(G)$ and $n \geq 1$, the $\tilde{\phi}_X^n(\delta)$ are given by

$$\begin{aligned} \tilde{\phi}_X^n(\delta) &= 0 \quad \text{on } \Omega - \tilde{R}_\delta^n \\ \tilde{\phi}_X^n(\delta) &= \frac{1}{T\lambda} \text{Log} \left[\hat{\phi}_Z^n(\delta) \right] + (\bar{\lambda}/\lambda) I_{d_\delta} \quad \text{on } \tilde{R}_\delta^n \end{aligned}$$

where the events $\tilde{R}_\delta^n$ are as in the proof of Proposition 6, and we assume known *a priori* constants k_δ necessary for their definition. As in (8), we can define nonparametric estimates $\tilde{p}_l^n$ where for $n, l \geq 1$

$$\tilde{p}_l^n(g) = 1 + \sum_{\delta \in \Gamma_l} f_\delta \text{tr} \left(\tilde{\phi}_X^n(\delta) U^\delta(g)^\dagger \right) \quad g \in G.$$

For all $\delta \in \text{Irr}(G)$ and $n \geq 1$, we have

$$\mathbb{E} \left| \tilde{\phi}_X^n(\delta) - \phi_X(\delta) \right|^2 \leq \frac{L'}{n} \left(\frac{d_\delta}{k_\delta^2} \right)^2 \quad (13)$$

where L' is a constant depending on the product $T\lambda$. This follows by a reasoning similar to the proof of Proposition 6. Moreover, for all $n, l \geq 1$, we have after putting $K = 0$

$$\mathbb{E} \|\tilde{p}_l^n - p\|^2 \leq \frac{L'}{n} \sum_{\delta \in \Gamma_l} (d_\delta^3/k_\delta^4) + \|p_l - p\|^2 \quad (14)$$

for the functions p_l defined in the proof of Proposition 7. This follows from Plancherel's formula in (12).

We have characterized the convergence of parametric estimates using (11) and (13) and the convergence of nonparametric estimates using (12) and (14). We make the following remarks on these formulae. Inequalities (11) and (13) only give gross bounds for the rate of convergence of parametric estimates. The quality of these bounds improves when the constants k_δ are greater, *i.e.*, closer to the value 1. This is equivalent to the $L^2(G, \mathbb{R})$ distance between p and the uniform density being greater. This last point can be appreciated in relation to the example of Fig. 3 in Section V-C.

Equations (12) and (14) describe the convergence of nonparametric estimates in a way similar to the one used in [8]. Indeed, the nonparametric estimation error is decomposed into two terms. One is given by the parametric estimation error and the other depends only on p . This second term is given by the convergence of the Fourier series of p . This is determined by the smoothness properties of p . We note the two following differences with [8], both related to the indirect nature of our ob-

servations. First, the first and second terms in (14) can not be identified as the “variance” and “bias” of $\tilde{p}_l^n$. Second, (14) characterizes the nonparametric estimation error as depending on the whole spectrum of p —through the constants k_δ —rather than just its smoothness properties.

We finally return to the role of the parameter K introduced in (8). For simplicity, we have put $K = 0$ for Proposition 7 and inequality (14). Let $K > 0$. The following function $p_K \in L^2(G, \mathbb{R})$ is an infinitely differentiable probability density, compare to [12], [8]

$$p_K(g) = 1 + \sum_{\delta \neq \delta_0} f_\delta \text{tr}(A_\delta U^\delta(g)^\dagger). \quad (15)$$

Using the same K in (8) and proceeding as for proposition 7, it is possible to obtain the limit in probability

$$\lim_l \lim_n \|\hat{p}_l^n - p_K\| = 0.$$

A similar limit also holds for the $\tilde{p}_l^n$. Note that in addition to being smooth, p_K can be chosen arbitrarily close to p in $L^2(G, \mathbb{R})$ for $K > 0$ small enough.

V. DECOMPOUNDING ON $SO(3)$ AND MULTIPLE SCATTERING

This section fulfills two goals. First, it summarizes recent use of compound Poisson processes on the rotation group $SO(3)$ in the modelling of multiple scattering and introduces decomposing on $SO(3)$ as a physical inverse problem. Second, it illustrates the characteristic function method presented in Section IV-C by applying it to a numerical example of decomposing on $SO(3)$. Nonparametric estimation on the rotation group $SO(3)$ has received special attention [11], [9]. It is important to many concrete applications and constitutes a privileged starting point for generalization to compact groups.

A. Compound Poisson Model for Multiple Scattering

Many experimental and applied settings aim to infer the properties of complex, *e.g.*, geophysical or biological, media by considering multiple scattering of mechanical or electromagnetic waves by these media. Inference problems arising in this way are formulated as physical inverse problems within the framework of various approximations of the exact equations of radiative transfer. See [19]–[21].

A compound Poisson model for the direct problem of multiple scattering was considered by Ning *et al.* [22]. It is based on a $\mathbb{R}$ -valued compound Poisson process. Consideration of compound Poisson processes on $SO(3)$ leads to a model of multiple scattering which is sufficiently precise as well as amenable to statistical treatment. This model extends the validity of the small angles approximation of radiative transfer. It also allows the formulation of the physical inverse problem of multiple scattering as a statistical nonparametric estimation problem.

We give an example expanding the above discussion. The development of Section III is converted into the terminology of radiative transfer, see [23]. Certain usual results in harmonic analysis on $SO(3)$ are here referred to freely. They are set down in a precise form in Section V-B.

A scalar plane wave is perpendicularly incident upon a plane parallel multiple scattering layer of thickness H . The velocity of the wave in the layer is normalized so that we have $\tau = \ell$ for the mean free time τ and mean free path ℓ . Coordinates and time origin are chosen so that the wave enters the layer at time 0 with direction of propagation $s(0) = (0, 0, 1)$. After time t in the layer this direction of propagation becomes $s(t) = (s^1(t), s^2(t), s^3(t))$. This is considered to be a random variable with values on the unit sphere $S^2 \subset \mathbb{R}^3$. The distribution of the random variable $s(H)$ is denoted I_H . It is identified with the normalized angular pattern of intensity transmitted by the layer. We return below to the validity of this identification.

The interaction of the wave with the layer takes place in the form of a succession of scattering events. These are understood as interaction of the wave with individual scatterers present at random emplacements throughout the layer. The random number of scattering events up to time $0 \leq t \leq H$ will be denoted $N(t)$. Suppose the n^{th} scattering event takes place at the time $0 \leq T_n \leq H$. This affects the direction of propagation as follows:

$$s(T_n) = s(T_n-)X_n. \quad (16)$$

Here, X_n is a random variable with values in $SO(3)$. It is identified with a random orthogonal matrix. Formula (16) is understood as a matrix equality where $s(T_n)$ and $s(T_n-)$ are line vectors. From (16) and the definition of $N(t)$ we can write for $0 \leq t \leq H$

$$s(t) = s(0) \left(\prod_{n=0}^{N(t)} X_n \right). \quad (17)$$

A certain number of standard physical hypotheses can be incorporated in (17). This will allow for the random product therein to be exhibited as a conjugate invariant compound Poisson process on $SO(3)$.

Under the condition $\ell \ll H$ it is possible to make the hypothesis that the time between successive scattering events has an exponential distribution [21]. This allows us to model $N(t)$ as a Poisson process with parameter $1/\ell$. Moreover, we suppose the scatterers identical and scattering events independent. This amounts to taking the $SO(3)$ -valued random variables X_n to be *i.i.d.*. If the additional assumption is accepted that the number of scattering events is independent of the whole outcome of these events then formula (17) can be rewritten for $0 \leq t \leq H$ as

$$s(t) = s(0)Y(t) \quad (18)$$

where Y is a (left) compound Poisson process on $SO(3)$ with parameter $1/\ell$. It is usual to assume that the random variables X_n have a common probability density p . For homogeneity with Section IV, we mention that p is a square integrable probability density with respect to the Haar measure of $SO(3)$. In the theory of radiative transfer, p is known as the phase function of the layer [23].

In order to simplify the Fourier series of p to a Legendre series (22) we profit from the physical hypothesis of statistical isotropy. This implies that scattering events in the layer as given by (16) are symmetric around the direction of propagation

$s(T_n-)$. Statistical isotropy is a valid assumption in a plurality of concrete situations. It is verified by analytical models such as Gaussian and Henyey–Greenstein phase functions, commonly used to describe scattering in geophysical and biological media [24].

Under the hypothesis of statistical isotropy the phase function p is a zonal function in the sense precised in Section V-B. It admits a Legendre series (22) wherein the coefficients a_δ for $\delta \in \mathbb{N}$ are said to form the associated power spectrum of heterogeneities [23]. If p is the Henyey–Greenstein phase function then the power spectrum of heterogeneities is given by $a_\delta = g^\delta$ for $\delta \in \mathbb{N}$ and p can be expressed in the closed form [24], [25]

$$p(\cos \theta) = \frac{1 - g^2}{(1 + g^2 - 2g \cos \theta)^{\frac{3}{2}}}. \quad (19)$$

In this formula the variable $\theta \in [0, \pi]$ refers to the scattering angle from an individual scatterer. It is given a mathematical definition in formula (22) of Section V-B. The parameter $g \in [0, 1]$ is called the anisotropy or asymmetry parameter. It can be shown to give the average cosine of the scattering angle θ . For the scattering of light waves by water clouds and blood, we have respectively $g = 0.85$ and $g = 0.95$, see [25].

Proposition 3 of Section III can be used to give the angular pattern of transmitted intensity I_H in terms of the power spectrum of heterogeneities. This is expressed in the following (20). This relates the directly observable outcome of multiple scattering in the layer to the constitutive microscopic properties of the layer, typically quite difficult to ascertain directly. Replacing in Proposition 3 the definition of the process Y of (18) and using the Legendre series (22) of p we have

$$\frac{I_H(\theta)}{2\pi} = \sum_{\delta \geq 0} (2\delta + 1) e^{\frac{H}{\ell}(a_\delta - 1)} \int_0^\theta P_\delta(\cos \xi) \sin \xi d\xi \quad (20)$$

where $I_H(\theta)$ is the ratio of intensity transmitted within a pencil of angle 2θ around $s(0)$.

Equation (20) is well known in the small angles approximation of radiative transfer where it is derived under the assumption of strong forward scattering [23]. Mathematically, this translates into a phase function p with a sharp peak around $\theta = 0$. Our probabilistic development of (20) does not explicitly make this assumption. However, the identification of I_H with the angular pattern of transmitted intensity implicitly requires for all the intensity of the wave entering the layer to be transmitted. This precludes an important deviation between $s(0)$ and $s(H)$.

Equation (20) is an interesting starting point for the formulation of the physical inverse problem of multiple scattering. Supposing a situation where this equation holds, being able to invert it implies access to the power spectrum of heterogeneities or alternatively the phase function from direct intensity measurements. This implies inference of physical parameters such as the parameter g of the Henyey–Greenstein phase function or determination of microscopic properties such as the shape of individual scatterers [25].

Our use of compound Poisson processes on $SO(3)$ to model multiple scattering lead to the probabilistic counterpart (18) of (20). In relation to (18), the physical inverse problem inherent

to (20) is reformulated as a statistical estimation problem. This appears as the problem of decompounding on $SO(3)$ or some related parametric estimation problem. A crucial difference between the two approaches is that they proceed from different types of data.

Suppose the distribution of $s(0)$ is known and symmetric around $(0, 0, 1)$ —this is the case in many experimental settings. Instead of carrying out measurements of transmitted intensity, it is possible to make observations of $s(H)$. Under the hypothesis of statistical isotropy these observations of $s(H)$ are equivalent to observations of $Y(H)$. If our objective is to estimate the phase function p then we have to deal with decompounding on $SO(3)$ from low-frequency observations of Y . In many cases, we could be interested in the power spectrum of heterogeneities or some related physical parameters. We then have to deal with a parametric estimation problem.

B. Harmonic Analysis on $SO(3)$

We here make a short digression on harmonic analysis on $SO(3)$ in order to clarify the references made to this subject in Section V-A and to prepare for Section V-C. $SO(3)$ is often used as the archetype compact connected Lie group. Essentially, we will specify the Peter–Weyl theorem as stated in Section II to the case $G = SO(3)$. For the following, see [9] or the more detailed account in [4].

We use the notation of Section II. In particular, μ denotes the Haar measure of $SO(3)$. It is possible to identify $\text{Irr}(SO(3)) = \mathbb{N}$ so that $d_\delta = 2\delta + 1$ for each $\delta \in \text{Irr}(SO(3))$. With this identification, the most current choice of functions $U^\delta : SO(3) \rightarrow SU(d_\delta)$ can be given in analytical form using the parameterization of $SO(3)$ by Euler angles.

The ZYZ Euler angles $\varphi, \psi \in [0, 2\pi]$ and $\theta \in [0, \pi]$ are well defined coordinates only on a subset of $SO(3)$. This is, however, a dense subset in the Euclidean topology of $SO(3)$ and has Haar measure equal to 1. Let $p : SO(3) \rightarrow \mathbb{C}$. If p is continuous or $p \in L^2(SO(3), \mathbb{C})$ it follows that p can be identified with a function of the Euler angles $p \equiv p(\varphi, \theta, \psi)$. The chosen functions U^δ are extended by continuity from the following expression for their matrix elements

$$U_{ab}^\delta(\varphi, \theta, \psi) = e^{-ia\varphi} d_{ab}^\delta(\cos \theta) e^{-ib\psi} \quad (21)$$

for $\delta \in \text{Irr}(SO(3))$ and $-\delta \leq a, b \leq \delta$. The notation d_{ab}^δ is used for the real-valued Wigner d-functions, which can be given in terms of the Jacobi polynomials. For $\delta \in \text{Irr}(SO(3))$ we have $d_{00}^\delta = P_\delta$ the Legendre polynomial of order δ .

The Haar measure μ is expressed in the coordinates (φ, θ, ψ) as follows:

$$d\mu(\varphi, \theta, \psi) = \frac{1}{8\pi^2} \sin \theta d\varphi d\theta d\psi.$$

Suppose a function $p \in L^2(SO(3), \mathbb{C})$ is expressed in the form $p(\varphi, \theta, \psi)$. In order to obtain its Fourier coefficients, it is enough to replace the above expressions for the functions U^δ and μ in formula (1). This formula then reduces to a triple integral. By the Peter–Weyl theorem, the Fourier coefficients of p give rise to a Fourier series approximating p in $L^2(SO(3), \mathbb{C})$.

The class of zonal functions on $SO(3)$ arises in relation to the hypothesis of statistical isotropy mentioned in Section V-A.

We will say that a function $p \in L^2(SO(3), \mathbb{C})$ is zonal if $p \equiv p(\theta)$. That is, if the expression of p in the coordinates (φ, θ, ψ) depends only on θ . Zonal functions form a closed subspace of $p \in L^2(SO(3), \mathbb{C})$. If p is a zonal function then its Fourier series reduces to a Legendre series

$$p(\theta) = \sum_{\delta \geq 0} (2\delta + 1) a_\delta P_\delta(\cos \theta) \quad (22)$$

where for $\delta \geq 0$ the Legendre coefficient a_δ is given by

$$a_\delta = \frac{1}{2} \int_0^\pi p(\theta) P_\delta(\cos \theta) \sin \theta d\theta. \quad (23)$$

Identities (22) and (23) can be found as follows. Let p be a zonal function. For $\delta \in \text{Irr}(SO(3))$ let A_δ be the Fourier coefficients of p obtained by taking $p = f$ in (1). The matrix elements of each A_δ are denoted A_δ^{ab} for $-\delta \leq a, b \leq \delta$. For all δ, a, b as above we have that A_δ^{ab} is given by

$$\frac{1}{8\pi^2} \int_0^{2\pi} \int_0^\pi \int_0^{2\pi} e^{ib\varphi} p(\theta) d_{ba}^\delta(\cos \theta) e^{ia\psi} \sin \theta d\varphi d\theta d\psi$$

which follows using (1). Thus, for all $\delta \in \text{Irr}(SO(3))$ we have that $A_\delta^{ab} \neq 0$ only if $a = b = 0$. In other words, the matrix A_δ contains at most one nonzero element. This is the diagonal element $A_\delta^{00} = a_\delta$ given by identity (23). Identity (22) follows by constructing the Fourier series of p as in (2).

C. Numerical Simulations

Here, we will illustrate the characteristic function method of Section IV-C by applying it to a numerical example of decompounding on $SO(3)$. Within this example we will consider a parametric estimation problem related to a physical inverse problem as described in Section V-A. Our example is of a compound Poisson process Y on $SO(3)$. As in Section V-A, $SO(3)$ -valued random variables are identified with random orthogonal matrices. For $t \geq 0$

$$Y(t) = \prod_{n=0}^{N(t)} X_n$$

where the Poisson process N has parameter $\lambda = 0.3$ and the random variables X_n have a common probability density p given by expression (19). Four values will be considered for the parameter g in this expression: 0.85, 0.9, 0.95, and 0.99. We will put $T = 10$. We simulate a number n of *i.i.d.* observations of $Y(T)$. The following values of n are used: 500, 5000 and 50000. Note that on average the number $N(T)$ of factors involved in the random product $Y(T)$ is equal to 3.

Before going on, we confirm that the method of Section IV-C can be applied for this example. In other words, that the X_n with the proposed density p are inverse invariant. This follows from the development after identities (22) and (23). Indeed, the matrices A_δ obtained for p are diagonal with exactly one nonzero diagonal element $a_\delta = g^\delta$. Since g is real, we have that A_δ is Hermitian for all $\delta \in \text{Irr}(SO(3))$. Inverse invariance follows by 1 of Proposition 2.

We will present three sets of figures. Fig. 1 is concerned with the compounding transformation of p . Fig. 2 illustrates the influence of n on parametric and nonparametric estimation errors.

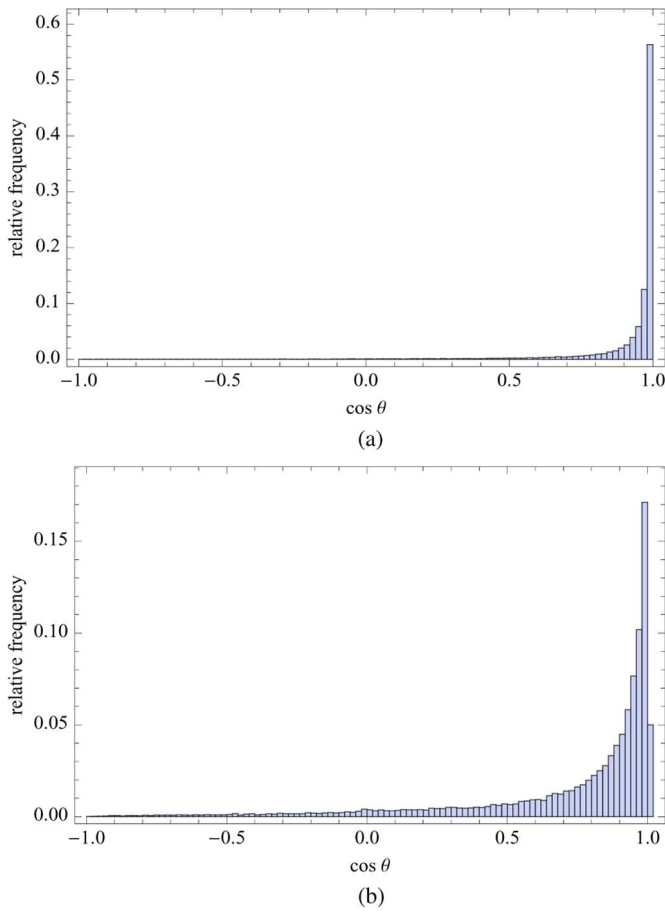


Fig. 1. Compounding transformation of p (histograms). (a) Histogram of $\cos \theta$ under density p . (b) Histogram of $\cos \theta$ under distribution of $Y(T)$.

Fig. 3 studies the influence of g on the nonparametric estimation error for fixed n . For Figs. 1 and 2, we have $g = 0.9$. For Figs. 1 and 3, we have $n = 50000$. We now comment on each of these figures.

Fig. 1 illustrates the relation between the distribution of the X_n as given by the density p and the distribution of $Y(T)$. Both these distributions are studied using histograms. The histogram in Fig. 1(a) is for the cosine of the Euler angle $\theta \in [0, \pi]$ associated with the random variable X_1 . The histogram in Fig. 1(b) is for the cosine of θ associated with $Y(T)$.

Fig. 1 is concerned with the direct compounding transformation rather than the inverse decomposing transformation. It is meant to show the histogram in Fig. 1(b) as function of the one in Fig. 1(a). As expected, the latter histogram appears as a wider version of the former. This corresponds to the content of Proposition 5 of Section III. Note also that the dominant value in Fig. 1(b) has moved away from $\theta = 0$.

For Fig. 2, the observations made of $Y(T)$ are used to carry out the decomposing approach of Section IV-C. Parametric and nonparametric estimation errors are given graphically for different values of n . Fig. 2(a) compares the estimated Legendre coefficients of p to their theoretical values $a_\delta = g^\delta$ for $\delta \geq 0$. In Fig. 2(b), *a priori* knowledge of the analytical form of the a_δ is supposed. This is used to estimate g . A different parametric estimate is obtained from each estimated Legendre coefficient.

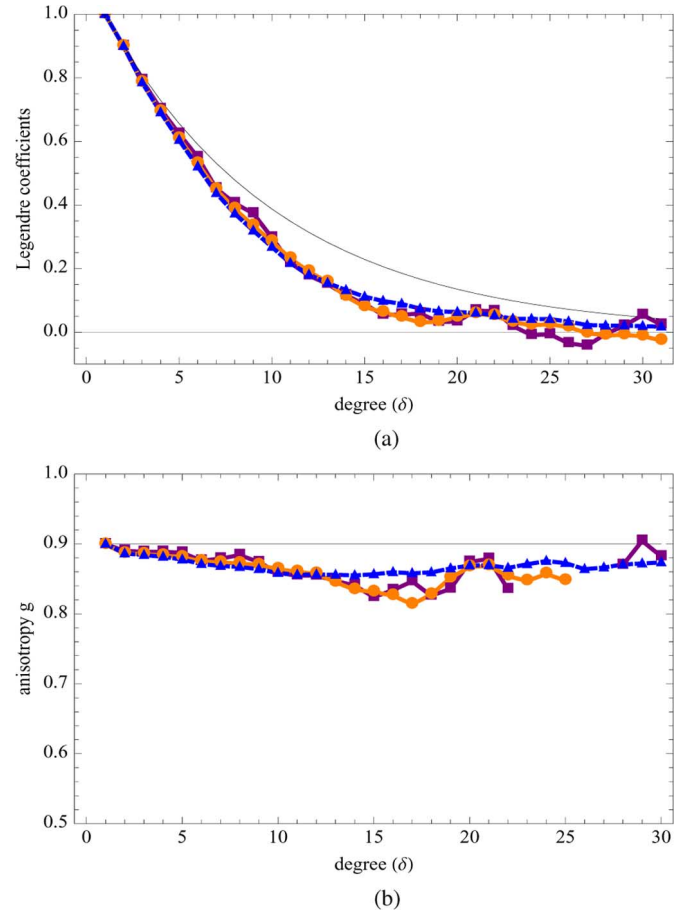


Fig. 2. Influence of n ($\square = 5 \times 10^2$; $\circ = 5 \times 10^3$; $\triangle = 5 \times 10^4$). (a) Estimated Legendre coefficients $\hat{a}_\delta^n$ from decomposing. (b) Corresponding estimates $\hat{g}_\delta^n$ of g (anisotropy parameter).

In Fig. 2(a) and (b), theoretical values are represented by a solid line.

In Fig. 2(a), we have the estimated first $l = 31$ Legendre coefficients for each value of n . Let us call these coefficients $\hat{a}_\delta^n$ for $0 \leq \delta \leq l$ and the corresponding value of n . They can be used to evaluate a nonparametric estimate of p as in formula (8). This is done by replacing them in a truncated Legendre series (22). We have the nonparametric estimate of p which we denote $\hat{p}_l^n$

$$\hat{p}_l^n(\theta) = 1 + \sum_{\delta=1}^{l-1} (2\delta + 1) \hat{a}_\delta^n P_\delta(\cos \theta)$$

where for all values of n we have that $\hat{a}_0^n = a_0 = 1$. Depending on n , the random nonparametric estimation error from $\hat{p}_l^n$ is given by

$$\sum_{\delta < l} (2\delta + 1) (\hat{a}_\delta^n - a_\delta)^2 + \sum_{\delta \geq l} (2\delta + 1) a_\delta^2$$

this is the squared $L^2(SO(3), \mathbb{R})$ distance between $\hat{p}_l^n$ and p . In Fig. 2(a), the sum over $\delta < l$ appears as a weighted quadratic deviation between estimated and theoretical values.

In Fig. 2(b), the estimates $\hat{a}_\delta^n$ are used to give naive estimates $\hat{g}_\delta^n$ of g based on the analytical form of the a_δ . The error in each

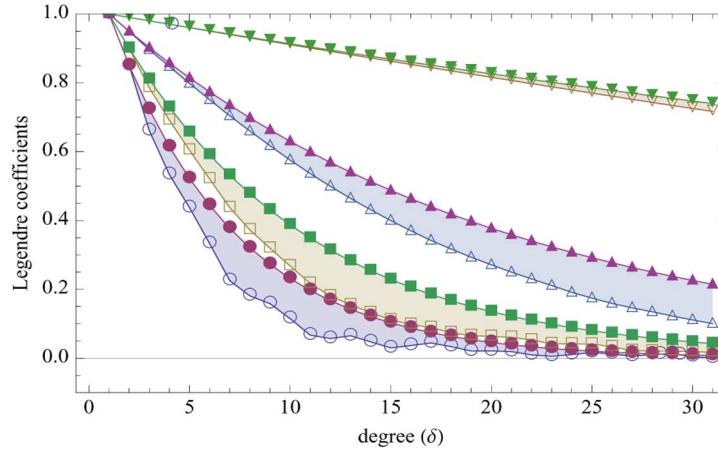


Fig. 3. Influence of g ($\circ = 0.85$; $\square = 0.9$; $\triangle = 0.95$; $\nabla = 0.99$).

of these estimates $\hat{g}_\delta^n$ is directly related to the error in the estimate $\hat{a}_\delta^n$. This latter error is shown for each δ and n in Fig. 2(a). The influence of n is not important for small values of δ . Visually, the $\hat{a}_\delta^n$ in Fig. 2(a) agree independently of n for $0 \leq \delta \leq 5$. For $n = 50000$, the $\hat{a}_\delta^n$ appear to have a regular dependence on δ . For $n = 5000$ and $n = 500$, we have an irregular dependence of the $\hat{a}_\delta^n$ on δ , especially for $\delta \geq 20$. Moreover, for $\delta \geq 25$, we have negative values of $\hat{a}_\delta^n$, clearly inconsistent with the form $a_\delta = g^\delta$. These values do not allow the evaluation of corresponding parametric estimates $\hat{g}_\delta^n$.

Let us remind that g is an important parameter in multiple scattering applications. For multiple scattering media with Henyey–Greenstein phase function (19), g is the main parameter characterizing the scattering process. Its estimation from observations as the ones described in Section V-A is equivalent to a physical inverse problem. This leads to the physical interpretation of the parametric estimation problem represented in Fig. 2(b).

For Fig. 3, we have $n = 50000$. For each value of g , we simulated n observations of $Y(T)$ and calculated estimates of the Legendre coefficients of p as for Fig. 2(a). Estimated and theoretical Legendre coefficients are respectively represented by empty and filled in symbols. It is clear from this figure that the nonparametric estimation error is smaller for larger values of g . Estimation of the Legendre coefficients is virtually exact for $g = 0.99$.

In order to understand this behavior, we note that g in (19) gives the concentration of p near the value $\theta = 0$. Indeed, when $g = 0$ the function p is constant and the random variables X_n are uniformly distributed on $SO(3)$. In the limit $g \uparrow 1$, we have that each random variable X_n is almost surely equal to the identity matrix. Conditionally on the event $\{N(T) > 0\}$, the distribution of $Y(T)$ is a mixture of distributions with Henyey–Greenstein density. More precisely, for all $n > 0$ we have the conditional probability density for the Euler angle θ associated with $Y(T)$

$$p(\theta|N(T) = n) = \frac{1 - g^{2n}}{(1 + g^{2n} - 2g^n \cos \theta)^{\frac{3}{2}}}.$$

In particular, in the limit $g \uparrow 1$ we have that $Y(T)$ is almost surely equal to the identity matrix. Conditionally on $\{N(T) >$

$0\}$, we have in the limit $g \downarrow 0$ that $Y(T)$ is uniformly distributed on $SO(3)$.

Let us note that in our example $\mathbb{P}(N(T) > 0) \simeq 0.96$. Fig. 3 can be understood in light of the above discussion. For greater values of g , observations of $Y(T)$ are concentrated near the identity matrix. This leads to fast convergence of our estimates for the Legendre coefficients of p . For smaller values of g , observations of $Y(T)$ are more dispersed and the convergence of estimates is slower. In the limit $g \downarrow 0$, the observations are close to uniformly distributed on $SO(3)$ and our approach breaks down due to numerical problems.

VI. CONCLUSION

Nonparametric estimation on compact Lie groups, especially using characteristic function methods, is by now a relatively familiar topic in relation to several engineering applications. It has received comprehensive treatment in the case where estimation is carried out directly from some stationary process. That is, from *i.i.d.* observations of a group-valued random variable. This paper has applied a characteristic function method to the problem of decompounding on compact Lie groups. For this problem, nonparametric estimation is required from indirect observations defined in terms of a nonstationary process.

A first approach of decompounding on compact Lie groups was given. It was guided by existing characteristic function methods for the classical problem of decompounding. These methods were transposed directly to the setting of harmonic analysis on compact Lie groups. Under a suitable symmetry assumption, treatment of the indirect nature of observations was simplified. The ensuing nonparametric estimation error was characterized as depending on the whole spectrum of the target density rather than just its smoothness class. In some aspects, our approach of decompounding on compact Lie groups might appear summary. We hope, however, that it will attract attention to various problems of the statistics of nonstationary stochastic processes on groups.

This paper also discussed the importance of decompounding on $SO(3)$ to the physical inverse problem of multiple scattering. Under a probabilistic interpretation of the theory of radiative transfer, models based on compound Poisson processes on $SO(3)$ were found consistent with the results of the

small angles approximation of radiative transfer. The possibility of reformulating physical inverse problems of multiple scattering as parametric or nonparametric statistical estimation problems was discussed. The statistical nature of this new point of view seems desirable given the high complexity of multiple scattering situations. In practice, it might require considerably more elaborate measurements.

ACKNOWLEDGMENT

J. H. Manton would like to thank the Australian Research Council.

REFERENCES

- [1] B. Buchmann and R. Grübel, "Decomposing: An estimation problem for Poisson random sums," *Ann. Statist.*, vol. 31, no. 4, pp. 1054–1074, 2003.
- [2] B. van Es, S. Gugushvili, and P. Spreij, "A Kernel type nonparametric density estimator for decomposing," *Bernoulli*, vol. 13, no. 3, pp. 672–694, 2007.
- [3] D. Applebaum, "Compound Poisson processes and Lévy processes in groups and symmetric spaces," *J. Theoret. Probab.*, vol. 13, no. 2, pp. 383–425, 2000.
- [4] G. Chirikjian and A. Kyatkin, *Engineering Applications of Noncommutative Harmonic Analysis*. Boca Raton, FL: CRC, 2000.
- [5] U. Grenander, *Probabilities on Algebraic Structures*. Hoboken, NJ: Wiley, 1963.
- [6] H. Heyer, *Probability Measures on Locally Compact Groups*. New York: Springer Verlag, 1977.
- [7] B. Yazici, "Stochastic deconvolution over groups," *IEEE Trans. Inf. Theory*, vol. 50, pp. 494–510, 2004.
- [8] J.-Y. Koo and P. Kim, "Asymptotic minimax bounds for stochastic deconvolution over groups," *IEEE Trans. Inf. Theory*, vol. 54, no. 1, pp. 289–298, Jan. 2008.
- [9] P. T. Kim and J.-Y. Koo, "Optimal spherical deconvolution," *J. Multivariate Anal.*, vol. 80, pp. 21–42, 2002.
- [10] P. Kim and D. Richards, "Deconvolution density estimation on compact Lie groups," in *Algebraic Methods in Statistics and Probability*. Providence, RI: AMS, 2001, pp. 155–171.
- [11] P. Kim, "Deconvolution density estimation on $SO(N)$," *The Ann. Statist.*, vol. 26, no. 3, pp. 1083–1102, 1998.
- [12] M. Liao, *Lévy Processes on Lie Groups*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [13] T. Bröcker and T. tom Dieck, *Representations of Compact Lie Groups*. New York: Springer, 1985.
- [14] J. Duistermaat and J. Kolk, *Lie Groups*. New York: Springer Verlag, 2000.
- [15] M. Liao, "Lévy processes and Fourier analysis on compact Lie groups," *Ann. Probab.*, vol. 32, no. 2, pp. 1553–1573, 2004.
- [16] O. Kallenberg, *Foundations of Modern Probability*. New York: Springer Verlag, 2002.
- [17] C. Van Loan and G. Golub, *Matrix Computations*. Baltimore, MD: John Hopkins Univ. Press, 1989.
- [18] J. Uherka and A. Sergott, "On the continuous dependence of the roots of a polynomial on its coefficients," *Amer. Math. Monthly*, vol. 84, no. 5, pp. 368–370, 1977.
- [19] H. Sato and M. Fehler, *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*. New York: Springer, 1998.
- [20] R. Xu, *Particle suspensions: Light scattering methods*. Norwell, MA: Kluwer, 2002.
- [21] P. Sheng, *Wave Scattering, Localization and Mesoscopic Phenomena*. New York: Academic, 1995.
- [22] X. Ning, L. Papiez, and G. Sandinson, "Compound Poisson process method for the multiple scattering of charged particles," *Phys. Rev. E*, vol. 52, no. 5, pp. 5621–5633, 1995.
- [23] A. Ishimaru, *Wave Propagation and Scattering in Random Media*. New York: Academic, 1978, vol. 1,2.
- [24] L. Klimes, "Correlation functions of random media," *Pure Appl. Geophys.*, vol. 159, pp. 1811–1831, 2002.
- [25] A. Kokhanovsky, "Small angle approximations of the radiative transfer theory," *J. Phys. D*, vol. 30, pp. 2837–2840, 1997.

Salem Said recently completed the Ph.D. thesis at the Institut polytechnique de Grenoble, France. The subject of his thesis is "Estimation and filtering of processes in matrix Lie groups".

His general research interests include stochastic filtering and stochastic processes in manifolds. His research activity also includes stochastic modelling in polarization optics and multiple scattering.

Christian Lageman received the Dipl.-Inf. degree in computer science and the Dipl.-Math. and Dr. rer. nat. degrees in mathematics from the University of Würzburg, Germany, in 2001, 2002, and 2007, respectively.

From 2007 to 2008, he worked as a Research Fellow at the Centre for Mathematics and its Applications, The Australian National University, Canberra, Australia. From 2008 to 2009, he was on a postdoctoral fellowship at the Department of Electrical Engineering and Computer Science, University of Liège, Belgium. Since September 2009, he has been with the Department of Mathematics, University of Würzburg, Germany. His current research interests are in nonlinear control theory, optimization, and estimation on manifolds.

Nicolas Le Bihan was born in 1974 in Morlaix, France. He received the B.Sc. degree in physics from the Université de Bretagne Occidentale (UBO, Brest), France, in 1997, and the M.Sc. and Ph.D. degrees in signal processing, respectively in 1998 and 2001, both from the Institut Polytechnique de Grenoble (Grenoble INP), France.

Since 2002, he has been Chargé de Recherche at the Centre National de la Recherche Scientifique (CNRS) and is working with the Department of Images and Signals at the GIPSA-Lab (CNRS UMR 5083) in Grenoble, France.

His research interests include statistical signal processing on groups and manifolds and its applications in polarized wave physics, waves in disordered media, and geophysics.

Jonathan H. Manton (S'95–M'02–SM'03) received the B.S. (mathematics) and B.Eng. (electrical) degrees in 1995 and the Ph.D. degree in 1998, all from the University of Melbourne, Australia.

From 1998 to 2004, he was with the Department of Electrical and Electronic Engineering at the University of Melbourne. During that time, he held a Postdoctoral Research Fellowship then subsequently a Queen Elizabeth II Fellowship, both from the Australian Research Council. In 2005, he became a full Professor in the Department of Information Engineering, Research School of Information Sciences and Engineering (RSISE) at the Australian National University. From July 2006 till May 2008, he was on secondment to the Australian Research Council as Executive Director, Mathematics, Information and Communication Sciences. Currently, he holds a distinguished Chair at the University of Melbourne with the title Future Generation Professor. He is also an Adjunct Professor in the Mathematical Sciences Institute at the Australian National University. His traditional research interests range from pure mathematics (e.g., commutative algebra, algebraic geometry, differential geometry) to engineering (e.g., signal processing, wireless communications). Recently though, led by a desire to participate in the convergence of the life sciences and the mathematical sciences, he has commenced learning neuroscience. He also has extensive experience in software development.

Non-adiabatic geometric phase of elastic waves

Jérémie Boulanger and Nicolas Le Bihan
Gipsa-lab, CNRS, 961, rue de la Houille blanche,

Domaine universitaire BP 46, 38402 Saint-Martin d'Hères CEDEX

Stefan Catheline

Institut des Sciences de la Terre - CNRS & Université J. Fourier, BP 53 38041 Grenoble, FRANCE

Vincent Rossetto*

Université Joseph Fourier, Grenoble - FRANCE and

*Laboratoire de Physique et Modélisation des Milieux Condensés - CNRS, Maison des Magistères - BP 166
25, avenue des Martyrs - 38042 Grenoble CEDEX - FRANCE*

We study the transport of elastic waves in a waveguide with helical shape. Polarization exhibits a geometric phase (or Berry phase): The polarization plane rotates along the helix following a geometric rule called parallel transport. Whereas this experiment is similar to the first experimental evidence of a Berry phase, by Tomita and Chiao (*Phys. Rev. Lett.* **57**, 1986), there is a major difference: The evolution of polarization is not adiabatic. This experiment therefore addresses the universality of the geometric phase beyond the adiabatic regime. We show that properties of the observed geometric phase coincide with the ones predicted by the adiabatic theory. The measured value of the phase is consistent — up to experimental uncertainty — with the theoretical value and no dependency with frequency is observable either.

INTRODUCTION

Polarization is a feature shared by several kind of waves: light and elastic waves, for instance, have two transverse polarization modes. From a more general point of view, quantum particles with spin can also be considered as polarized waves [1]. The polarization degrees of freedom are constrained to lie in the plane orthogonal to the propagation direction. This constraint is responsible, in optics, for the existence of a phase called Berry phase (in quantum mechanics) or geometric phase (in classical physics) [2]. The Berry phase can be different from zero only if the optical path of the polarized beam takes a three-dimensional shape. The Berry phase is a very general physical concept, emerging from the geometry of the phase space. It applies to optical waves and to quantum states as well, it was even discovered in the frame of quantum mechanics [3]. In classical mechanics the geometric phase appears, for instance, in Foucault's pendulum [4].

The fundamental origin of Berry and geometric phases can be found in the geometric description of the phase space. Being of pure geometric origin, such a phase should thus exist for elastic waves [5]. The geometric phase is usually defined and studied for a system, where the oscillations (classical or quantum) occur at a frequency much larger than the rate of direction change. This adiabaticity condition is fulfilled in optics by several orders of magnitude. Foucault's pendulum fulfills this condition as well, because the oscillation period of the pendulum is of the order of the second while the Earth

rotates with a period of one day. Aharonov and Anandan reported, shortly after Berry's discovery, that the adiabaticity is not required for a phase to appear [6]. Samuel and Bhandari later even pointed out that the time evolution can also be non-unitary and non-cyclic [7].

In this article, we report on an experiment showing evidence for the Berry phase for elastic waves and we detail the computation of its value. In this experiment, the adiabaticity condition is not fulfilled. We show that the existence of the phase and its value are nonetheless preserved.

In Section I, we introduce the concept of geometric phase using a purely geometrical description. The experiment is described and its results are discussed in Section II. Section III presents a computation of the geometric phase in the framework of the experiment. We finally conclude by comparing the experimental values to the geometric phase obtained from the theory and discuss the adiabaticity condition.

I. THE GEOMETRIC PHASE FOR POLARIZED WAVES

Polarization is relevantly described by degrees of freedom lying in the transverse plane to the direction of propagation of a wave. We consider in the present work only classical waves (equivalent to quantum waves with spin 1), for which polarization is described by two degrees of freedom in the transverse plane. When the two transverse components oscillate in phase, (or in phase opposition), the polarization state is said to be linear and is characterized by a direction orthogonal to the propagation direction. Any other state can be represented as a combination of two linearly polarized states, so we

* vincent.rossetto@grenoble.cnrs.fr

will only consider linear polarization here, without loss of generality. We will represent polarization as a vector in the plane orthogonal to the direction of propagation.

If the wave travels along a straight path, polarization remains constant along the trajectory; if the direction of propagation is not constant, polarization cannot be kept constant because it is ascribed to remain in the orthogonal plane (see Figure 1). There is therefore a path-dependent transformation transporting polarization along the trajectory. The obvious properties of this transformation are that it is linear, reversible and continuous. As discussed by Segert [5], only one transformation satisfies these requirements: parallel transport.

Considering the continuous path followed by the wave, the direction of propagation can be represented as a unit vector in $\mathbf{R}^3$, *i.e.* an element of the unit sphere $\mathcal{S}^2$. At each point on $\mathcal{S}^2$, the polarization is orthogonal to the propagation vector, so it can be represented in the tangent plane to the sphere, as shown in Figure 1.

Let us consider a geodesic on the unit sphere, *i.e.* a great circle. If polarization is orthogonal to the great circle, it remains constant along the path, as shown in Figure 2(a). Similarly, if polarization is oriented along the great circle, the colinearity holds along the path to preserve symmetry. A general polarization is a linear combination of these two particular linear polarizations. The linearity of parallel transport implies, for linear polarizations, that if one rotates the initial polarization in the initial tangent plane, the parallel transported polarization rotates identically in all the other tangent planes (see Fig. 2(c)). Parallel transport can be extended to any smooth and piecewise differentiable trajectory of the tangent vector on the unit sphere by discretizing the path into elementary arcs of great circles.

A vector submitted to parallel transport along a closed trajectory does not necessarily have the same orientation in the tangent plane at the starting point and after one stride along the trajectory. The angle between the initial and final vectors is algebraically equal to the area enclosed by the trajectory on $\mathcal{S}^2$ (see Fig 3). In differential geometry, this result is known as the holonomy of the trajectory on the tangent bundle of the sphere [8], and is computed using the theory of connections. We will use this important property to interpret the results of the experiment presented in Section II.

II. OBSERVATION OF THE GEOMETRIC PHASE FOR ELASTIC WAVES

This section presents an experiment designed to observe the geometric phase of elastic waves and the results obtained. It is an adaptation of the experiment of Tomita and Chiao [2], performed to observe Berry phase for light.

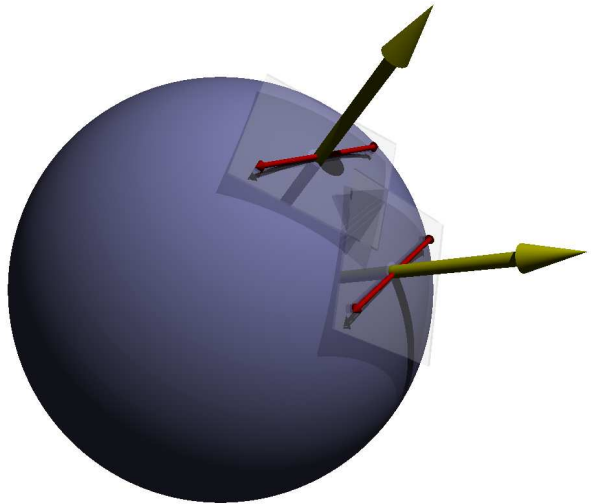


FIG. 1. The sphere of directions is the phase space of the propagating wave. In the Figure, we show two directions as large arrows going out radially. The planes tangent to the sphere at the extremity of these arrows are the orthogonal planes containing the polarization. Polarization is represented as a double arrow lying in the tangent planes. When the direction of transport changes, polarization evolves in such a way that it remains in the tangent plane.

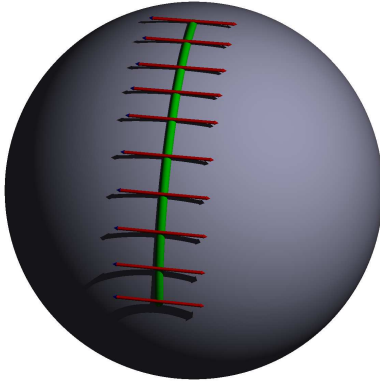
A. Experimental setup

We use a metallic spring as a waveguide for elastic waves. The spring was taken from a car's rear damper. It has a circular section of 13.5 mm making five coils of radius $R = 75 \pm 1$ mm and with a pitch of $P = 91.5 \pm 1$ mm. The helix is suspended to two strings to isolate the system.

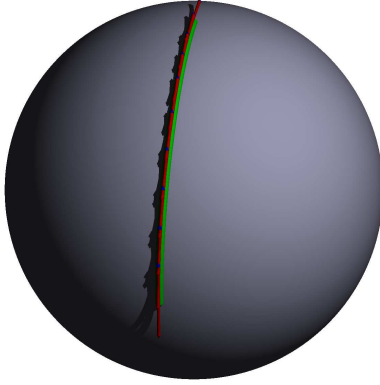
We use two accelerometers located at one end of the spring and record vibrations in two orthogonal directions (see Figure 4). We use piezoelectric accelerometers from Brüel & Kjaer, type 4518-003. Their transfer power is stable up to 10 % in the frequency range from 0.5 Hz to 20 kHz. Pre-processing amplification (Brüel & Kjaer amplifier, type 2694) is performed before signal processing. The sampling frequency is set to 50 kHz.

We record the waves at different distances from the source by moving the source along the helix onto 32 equally spaced positions, which distance from the accelerometers ranges from a few centimeters to 1.45 m. The distances are measured along the helix and correspond, after correction, to the wave propagation length.

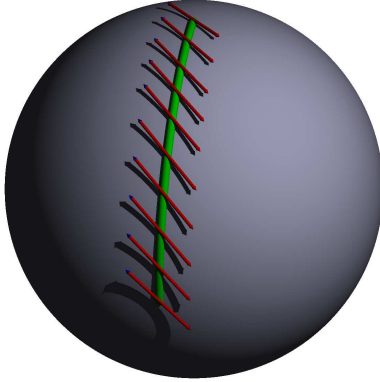
Making weak impacts on the metal spring creates bending waves that are linearly polarized. Impacts are made radially with respect to the helix. The source signal is generated manually, by gently hitting the waveguide with a hammer at the different source positions. Polarization depending only on the amplitude ratio measured by the accelerometers, it is not sensitive to the energy of the source. The hits are made in such a way that they



(a) Parallel transport of a vector orthogonal to displacement

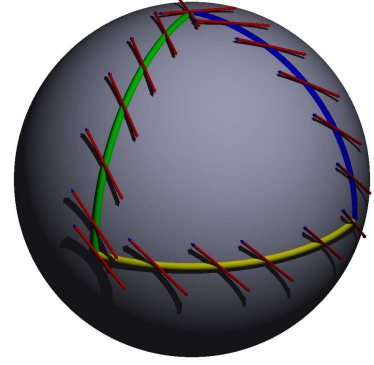


(b) Parallel transport of a vector colinear with displacement

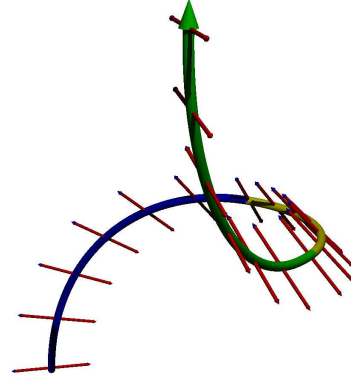


(c) Parallel transport of an arbitrary vector

FIG. 2. Parallel transport of a polarization (double arrows) along a trajectory on the direction sphere. Figures (a) and (b) show two orthogonal linear polarizations forming a basis of the polarization space. On Figure (a), we show that the polarization component orthogonal to the geodesic on the direction sphere remains constant while Figure (b) shows that the component tangent to the geodesic turns and remains tangent to the geodesic. Combining Figures (a) and (b), one deduces that all polarization vectors turn in the same way in all tangent planes along the geodesic, as illustrated in the example of Figure (c). The scalar product is therefore preserved by parallel transport.



(a) Parallel transport of a vector along a closed trajectory on the tangent vector sphere.



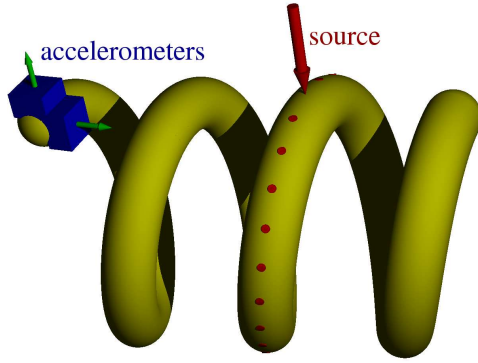
(b) Trajectory in real space, with polarization traced along it.

FIG. 3. A trajectory represented on the tangent vector sphere (a) and in real space (b). At the start and at the end of the trajectory, the tangent vector points upwards. One observes that in a non-flat space, parallel transport is not trivial. The red vector, although it is parallel transported all along the trajectory back to its initial position, rotates between the start and the end of the displacement. The difference angle is equal to the algebraic area enclosed by the trajectory on the sphere (a).

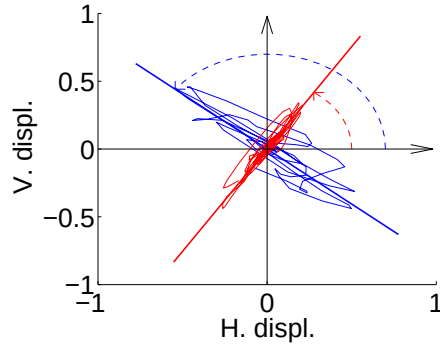
do not saturate the sensors but provide a good signal to noise ratio.

B. Measurements

We make several measurements at each position of the source in order to estimate the systematical errors and reduce their influence on the final result. We show in Figure 5 an example of two signals recorded by the two orthogonal accelerometers. Because of polarization loss, we use only the first periods of oscillation of the signal.



(a) Schematic representation of the experiment.



(b) Parametric plot of the acceleration components.

FIG. 4. (a) Setup of the experiment. Accelerometers are represented as blue cubes, they record acceleration in two orthogonal directions (green arrows). The large arrow symbolizes the source, a radial impact on the helix. The dots indicate some of other source positions. (b): Parametric plots of the two acceleration components recorded by the accelerometers. The direction of polarization is defined by the angle between the main oscillation and the horizontal axis. These signals were recorded 40 cm (red) and 110 cm (blue) away from the accelerometers, in arc length distance.

The power spectrum of the signal recorded on one of the accelerometers is presented in figure 5 (c). The sampling frequency being 50 kHz, the information in the signal is available up to 25 kHz. However, the maximum frequency used for geometric phase measurement is 10 kHz. We did not consider frequencies above 10 kHz.

Bending waves in the waveguide are dispersive; they have different velocities at the two frequencies we focus on: 5 kHz and 10 kHz (see Table I).

Polarization information of the early signal is displayed as parametric plot in 5 (d). We consider only the first oscillations in the waveforms because coupling to other modes depolarizes the bending waves (see Section IV).

In this experiment, we study how the orientation of the linear polarization changes as a function of the arc length along the helix. The recorded signals are not per-

fectly linearly polarized (see Section II C). The principal axis of polarization is obtained from the records using a principal component analysis [9]. This technique consists in obtaining the eigenvectors of the cross-correlation between the two orthogonal signals. The eigenvector with the largest eigenvalue gives the polarization direction. In Figure 4, the right plot represents the polarization of early waves, after a filtering process around 5 kHz for two different locations of the source. Obviously, the axis of polarization rotates when the source location is displaced. We estimate the angle between the horizontal direction and the polarization direction, that we call the *direction of polarization*. The geometric phase is the difference between the angle of polarization measured at a location and that of a reference position, which we choose at the accelerometers (see Figure 4).

Before giving the results of our experiments, we give some details about the experimental errors that have been considered and included in the data analysis.

C. Experimental errors

In this section, we review the sources of experimental error and estimate their values.

- Accelerometers: The measurement of the direction of polarization relies on the orthogonality of the accelerometers. We therefore must take into account the inaccuracy of the orthogonality. We call ϵ the discrepancy between the actual angle and $\pi/2$. Namely, the angle between the direction of the measured accelerations is $\pi/2 + \epsilon$. Furthermore, the coupling between the vibration of the waveguide and the accelerometers cannot be controlled. We

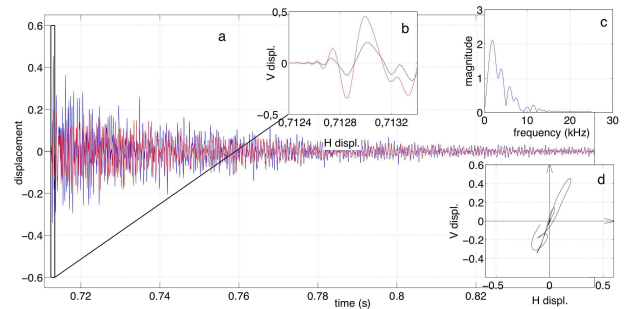


FIG. 5. Example of unfiltered vertical (red) and horizontal (blue) signals recorded by the two orthogonal accelerometers after the emission of one signal from the source. The origin of the time axis is arbitrary. Inset b: zoom from the main graph onto the first oscillations of the first periods of oscillation. Only this part of the signals is used to compute the polarization direction. Inset c: Power spectrum of the displayed time window signal. Inset d: Parametric plot of the first oscillations displayed on the inset b. The signal is visibly linearly polarized. The direction is computed from these data using a principal component analysis (see text).

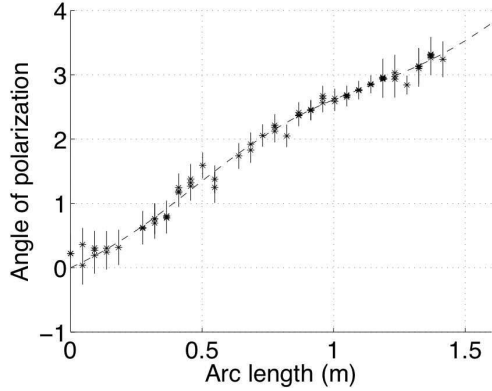


FIG. 6. Geometric phase as a function of arc length along the helix for the 5 kHz mode. The dashed line represents the linear regression corrected with $\epsilon = 0.21$ and $\eta = 0.27$. The slope found is 2.38 rad.m^{-1} .

assume that the coupling of the vertical sensor is $1 + \eta$ times the coupling of the horizontal one.

We take these errors into account in the data analysis to improve the model. The value Φ of the phase is shifted by the amount $\Delta\Phi = \frac{1}{2}(-\epsilon + \epsilon \cos(2\Phi) + \eta \sin(2\Phi))$. Using this expression, we estimate ϵ and η simultaneously with the main slope.

- **Source:** the signal is manually created using a small hammer. The position and the direction of hit are therefore submitted to systematical errors. We estimate that the positional error is around 2 mm and that the directional error is $\pi/20$ rad.
- **External sounds:** although we isolated the waveguide as much as possible from the environment, we can not completely remove ambient sounds from being recorded by the accelerometers.

To minimize the influence of these sources of error, we perform several independent measurements for each source position, this statistically improves the quality of the data. The confidence intervals presented in Figures 6 and 7 include the whole mentioned sources of error.

We obtain the estimates for the slope, ϵ and η fitting the data to a linear relation, corrected by the errors due to accelerometer positioning, and coupling. Both these error contributions result in an oscillation around the linear relation (see Figures 6 and 7).

D. Polarization rotation results

Results are presented separately in figure 6 and 7 for the two frequencies, and in Table I. In these figures one can see that dependency of the geometric phase is in first approximation linear with respect to the arc length. The

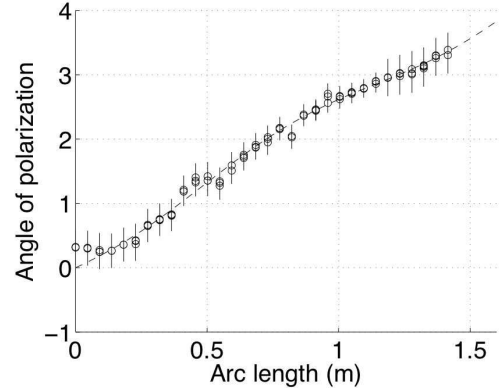


FIG. 7. Geometric phase as a function of arc length along the helix for the 10 kHz mode. The dashed line represents the linear regression corrected with $\epsilon = 0.22$ and $\eta = 0.23$. The slope found is 2.40 rad.m^{-1} .

Frequency	velocity	wavelength	slope (rad.m^{-1})	ϵ	η
5 kHz	2500 m.s^{-1}	0.50 m	2.38 ± 0.1	0.21	0.27
10 kHz	3000 m.s^{-1}	0.24 m	2.40 ± 0.1	0.22	0.23

TABLE I. Physical values measured from the filtered signals. The causes of the experimental errors are discussed in the Section II C

oscillations are due to imperfections related to the accelerometers. We observe a linear behaviour for the two signals filtered at 5 kHz and 10 kHz, with comparable values of the slope. Numerical values are presented in Table I. We do not observe a dependency of the geometric phase with respect to the frequency or equivalently to the wave velocity or to the wavelength. Such a dependency was not expected [7].

In the next section, we present the calculation of the geometric phase for a helix with the same geometry as the waveguide.

III. COMPUTATION OF THE PHASE

The computation of the geometric phase is not as straightforward as explained in the original letter of Tomita and Chiao[2]. A more detailed approach is presented here. We adopt two different points of view : an application of a theorem due to Fuller [10] and an approach based on the explicit computation of the Levi-Civita connection, which defines parallel transport along the helix. The obtained expression for the phase coincides with the phase measured during the experiment described above.

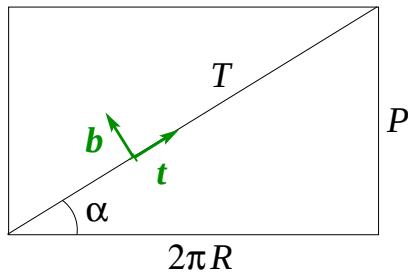


FIG. 8. Description of the geometry of the helix. R is the radius and P is the pitch. The figure shows a flattened cylinder with height equal to one pitch P and width equal to $2\pi R$. Therefore, the trajectory makes one helix turn and is represented by the diagonal. The tangent vector is represented together with the binormal vector $\mathbf{b}$. Note that depending on the helix handedness, the binormal vector orientation could be opposite.

A. Application of Fuller's theorem

As stated in Section I, the geometric phase has originally been defined for a closed trajectory in the phase space. The holonomy of a closed curve on the sphere $\mathcal{S}^2$ is thus useful for the purpose of computing the effect of parallel transport on polarization [11, chap. 4]. However, in the experiment described in Section II, the trajectory followed by the tangent vector is not closed in general. As an example, a trajectory is shown in Figure 9. If the trajectory on the unit sphere is not closed, the holonomy cannot be used as in the closed case. Fortunately, the theorem of Fuller [10] allows to compute a geometric phase by means of a deformation of the trajectory from a reference. In the present case, the choice of the reference is the trajectory $\alpha = 0$ (this is the limit case where the helix becomes a flat wheel) for which it is known that the geometric phase vanishes (essentially because the trajectory is two-dimensional). In Figure 9, we add the reference trajectory (in red) and coloured in yellow the area corresponding to the geometric phase.

If we call R the radius of the helix and P its pitch, then $T = \sqrt{(2\pi R)^2 + P^2}$ is the length of a helix turn. We call s the arc length of the wave trajectory. s/T is the number of turns made by the tangent vector around the unit sphere axis. (See Figure 8)

Fuller's theorem states that the geometric phase difference between the trajectories is equal to the area on the unit sphere spanned during the deformation. In the case of the experiment presented here, the area is proportional to s/T and the proportionality coefficient is $2\pi \sin \alpha$, given by the classic geometry formula for the spherical cap area. We get

$$\Phi = (2\pi \sin \alpha) \times \left(\frac{s}{T}\right) = \frac{2\pi P s}{4\pi^2 R^2 + P^2}. \quad (1)$$

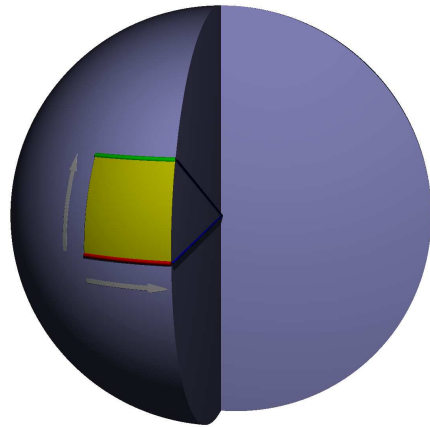


FIG. 9. Geometry of the transformation from the reference trajectory to the helix trajectory. Trajectories followed by the tangent vector are along meridians (vertical arrow on the figure). A displacement of s along the helix corresponds to the angle $2\pi s/T$ on the unit sphere (horizontal arrow). The coloured area is the area spanned by the tangent vector during the deformation. It is equal to the geometric phase difference computed by Fuller's formula.

B. Intrinsic point of view

Another way to compute the geometric phase is to consider the point of view of a walker following the trajectory and parallel transporting a vector $\mathbf{v}$. In the case of elastic waves, vector $\mathbf{v}$ is the polarization vector. The movement is described in the Frénet-Serret frame by the equations (see for example [11])

$$\frac{d\hat{\mathbf{n}}}{ds} = -\kappa \hat{\mathbf{t}} + \tau \hat{\mathbf{b}} \quad \text{and} \quad \frac{d\hat{\mathbf{b}}}{ds} = -\tau \hat{\mathbf{n}}. \quad (2)$$

The curvature κ and the torsion τ of a helix are constants: $\kappa = (2\pi \cos \alpha)/T$ and $\tau = (2\pi \sin \alpha)/T$. Parallel transport of a vector is, by definition, the only way to transport the vector in order to make the covariant derivative vanish along the trajectory. In other words, it corresponds to keep the vector constant from the walker's point of view. The Levi-Civita covariant derivative D projects the derivative in the plane orthogonal to the displacement [8]. We have therefore the definition of the Levi-Civita derivative:

$$D\mathbf{v} = \frac{d\mathbf{v}}{ds} - \left(\frac{d\mathbf{v}}{ds} \cdot \hat{\mathbf{t}}\right) \hat{\mathbf{t}}. \quad (3)$$

Along the trajectory, we must have $D\mathbf{v} = 0$. It follows that

$$D(\mathbf{v} \cdot \hat{\mathbf{n}}) = \tau(\mathbf{v} \cdot \hat{\mathbf{b}}) \quad \text{and} \quad D(\mathbf{v} \cdot \hat{\mathbf{b}}) = -\tau(\mathbf{v} \cdot \hat{\mathbf{n}}). \quad (4)$$

These equations show that a parallel transported vector rotates at rate τ (the Frénet torsion) in the Frénet-Serret frame. The Frénet torsion of a helix being equal

to $(2\pi \sin \alpha)/T$, one obtains

$$\Phi = \frac{2\pi \sin \alpha}{T} \times s, \quad (5)$$

which is identical to the previous expression (1) obtained with the deformation approach and Fuller's theorem.

Using the characteristics of the helix used in the experiment, we find that the slope of the linear dependence between arc length and polarization rotation is theoretically of $2.49 \pm 0.1 \text{ rad.m}^{-1}$. The fit performed in previous section lead to an estimation of $2.40 \pm 0.1 \text{ rad.m}^{-1}$. This is in good agreement with theoretical values. The independence of this value with respect to waves frequency is expected for a geometric phase because it is only due to the geometry of the path in the three dimensional space.

IV. DISCUSSION AND CONCLUSION

The observation of the geometric phase for elastic waves raises different issues than those encountered in optics. This can be seen as a consequence of the fact that the frequency of the waves we measure, around 5 kHz and 10 kHz, is of the same order of magnitude as the rate at which the propagation direction evolves, *i.e.* c/T . These parameters indicate that the time evolution is not adiabatic. Indeed, the wavelength of the bending wave is of the same order of magnitude as the length of a helix coil. The bending wave is even a global deformation of the waveguide.

Apart from bending waves, compression waves and torsion waves can also propagate in the helix. Bending and torsion waves propagate at the same velocity while compression waves have a higher velocity. In a straight rod, these propagation modes are not coupled and bending waves remain polarized at long times. In the helix, the coupling between compression waves and torsion waves increases with curvature and torsion. Bending waves and torsional waves are therefore partially converted into each other, back and forth, during the propagation. This explains why bending waves polarization decreases in our measurements and why we had to consider only the first oscillations of the signals.

We have demonstrated the existence of a geometric phase for elastic waves in a waveguide and in a non-adiabatic regime. This is an experimental evidence for the universality of the geometric phase. Note that the non-adiabatic regime could not be observed in Foucault's pendulum because the trajectory of the pendulum would not remain planar.

In Nature, polarized elastic waves, such as seismic S waves, are observed under certain conditions, and the concept of geometric phase could therefore be used to interpret their polarization under an original angle. For instance, the measure of the degree of polarization in a seismic signal created by a polarized source should contain informations concerning the propagating medium.

-
- [1] Rossetto, V. (2009) *Phys. Rev. E* **80**, 056605.
 - [2] Tomita, A & Chiao, R. Y. (1986) *Phys. Rev. Lett.* **57**, 937–940, 2471.
 - [3] Berry, M. V. (1984) *Proc. R. Soc. Lond. A* **392**, 45–57.
 - [4] von Bergmann, J & von Bergmann, H. (2007) *Am. J. Phys.* **75**, 888–892.
 - [5] Segert, J. (1987) *Phys. Rev. A* **36**, 10–15.
 - [6] Aharonov, Y & Anandan, J. (1987) *Phys. Rev. Lett.* **58**, 1593–1596.
 - [7] Samuel, J & Bhandari, R. (1988) *Phys. Rev. Lett.* **60**, 2339–2341.
 - [8] Nakahara, M. (2003) *Geometry, topology and physics*, Graduate student series in physics. (IOP publishing, Bristol).
 - [9] Vidale, J. (1986) *Bull. of the Seismological Soc. of Am.* **76**, 1393–1405.
 - [10] Fuller, F. B. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 3557–3561.
 - [11] Do Carmo, M. (1976) *Differential Geometry of Curves and Surfaces*. (Prentice Hall).

Travaux en cours et perspectives

Les travaux exposés dans ce manuscrit soulèvent plusieurs questions et ouvrent des pistes de recherches que je compte explorer dans les années à venir. Même si les deux parties des travaux présentés semblent relativement déconnectés, il existe peut être des connexions que je souhaite tenter d'établir.

4.1 Quaternions et signal

L'étude du signal hyperanalytique est un sujet qui me semble intéressant de pousser plus loin. Le lien entre transformation de Fourier et information géométrique et fréquentielle me paraît justifier à lui seul la poursuite de travaux sur les transformations de Fourier sur des structures algébriques non-commutatives. Le signal hyperanalytique étant la représentation "temps-fréquence" la plus simple (comme le signal analytique classique), on peut imaginer développer des outils plus performants pour l'étude des signaux non-stationnaires complexes et non-circulaires. Par exemple, on peut imaginer la distribution de Wigner-Ville quaternionique d'un signal complexe non-circulaire en utilisant la TF Quaternionique au lieu de la TF classique. On imagine aussi qu'un spectrogramme basé sur la TFQ pourra donner des informations sur le comportement du signal. En particulier, l'étude de la phase (en fait, des phases quaternioniques, vu que plusieurs définitions sont possibles) doit apporter un éclairage sur la façon dont le signal évolue dans le plan complexe au cours du temps. On imagine facilement l'intérêt d'une telle représentation sur les signaux polarisés qui sont soumis à des variations de polarisations et ce en fonction de la fréquence et au cours du temps. Un spectrogramme quaternionique pourrait permettre une analyse fine du comportement non-stationnaire polarisé. Un champ d'application pourrait être l'étude de la propagation des ondes polarisées dans les milieux dispersifs par exemple.

Ce qui a été présenté sur le signal hyperanalytique concerne les signaux déterministes. L'extension au cas des signaux aléatoires me semble une idée intéressante également et devrait permettre de relier le signal hyperanalytique à un concept abordé dans ce manuscrit : la circularité. Des travaux en cours avec S.J. Sangwine montrent la $\mathbb{C}$ -circularité¹ du signal hyperanalytique et je souhaite poursuivre cette étude. Nous souhaitons identifier les propriétés de circularité du signal hyperanalytique de processus complexes non-circulaires et éventuellement proposer des tests statistiques d'écart à la circularité quaternionique.

Une autre piste de recherche pour la circularité concerne les statistiques d'ordres supérieurs. Il est connu en traitement du signal que les statistiques d'ordres supérieurs permettent d'identifier des niveaux de circularité discrets [Lacoume 1997]. Il pourrait être intéressant d'étendre ces résultats connus pour les signaux complexes au cas quaternionique. Cela suppose des connaissances sur les statistiques d'ordres supérieurs pour les variables quaternioniques. Or, la non-commutativité des quaternions n'aide pas à l'étude des ordres supérieurs... Quelques résultats récents existent tout de même sur l'extension du théorème de Wick sur $\mathbb{H}$ [Bryc 2009] et j'envisage de travailler sur ce sujet et sur comment utiliser ces nouveaux résultats pour étudier la circularité quaternionique.

1. En fait, une version particulière de la $\mathbb{C}$ -circularité.

Finalement, parmi les travaux en cours sur la TF Quaternionique, j'étudie comment trouver un équivalent pour les séries temporelles à valeurs elles-mêmes quaternioniques. On imagine par exemple que cela ait un intérêt pour l'étude de séquences de quaternions unitaires (séquences temporelles de positions, attitudes, rotations, etc.). Comment définir une TF pour ces séries temporelles, sans passer par la théorie de la représentation et l'isomorphisme entre les quaternions purs et le groupe $SU(2)$? Est-il vraiment impossible de définir une TF octonionique² ? Ou peut-on utiliser la TF biquaternionique que nous avons déjà développée pour analyser des séquences de rotations ? Il pourrait être intéressant d'avoir un signal hyperanalytique à valeurs octonioniques ou biquaternioniques et d'en extraire des informations de vitesses de rotation instantanée par exemple.

L'étude des TF *géométrique* pour l'analyse des séquences temporelles est donc la suite que je souhaite donner à mes travaux sur les quaternions, en y incluant les outils d'analyse développés ces dernières années comme la circularité par exemple. Parmi ces TF *géométrique*, il me semble important également d'inclure les TF basées sur les algèbres de Clifford qui ont connus un certain succès dans la communauté *image* récemment, mais restent peu utilisées dans le domaine du signal. Leur étude fait donc partie des travaux futurs que je compte mener.

4.2 Processus de Poisson : extensions

Dans les travaux que j'ai mené jusqu'à présent, je me suis concentré sur l'utilisation des processus de Lévy, et en particulier sur les processus de Poisson composés sur $SO(3)$. Ces processus permettent de décrire de nombreux phénomènes en lien avec la polarisation. Il existe plusieurs extensions possibles qui peuvent être intéressante à explorer.

4.2.1 Processus sur $SE(3)$, ondes P et S

Dans le modèle de diffusion multiple des ondes scalaires (non polarisées) que nous avons présenté dans le chapitre 3, le processus sur $SO(3)$ est uniquement fonction du temps t . Il ne permet donc pas de modéliser la dépendance spatiale de la distribution des angles de diffusion. Nous avons bien précisé que le processus de Poisson composé sur $SO(3)$ permet de modéliser la diffusion vers l'avant, mais il pourrait être intéressant d'utiliser un processus spatio-temporel afin d'avoir une description spatiale de la diffusion multiple (particulièrement intéressante pour la description de milieu inhomogène par exemple).

Une possibilité pour faire cela est de considérer un processus de Poisson composé sur le groupe spécial Euclidien $SE(3)$. Ce groupe (non-compact) contient en plus de $SO(3)$ les translations dans l'espace 3D, donnant une information de position dans l'espace. Ce groupe est très utilisé en robotique et des outils d'AHNC ont déjà été développé pour analyser les fonctions à valeurs sur ce groupe [Chirikjian 2000]. Cette extension du modèle est relativement directe et facilement réalisable. D'une manière générale, la piste des processus de Poisson spatiaux-temporels doit être intéressante pour l'étude des phénomènes de diffusion multiple. Elle fait partie des perspectives à court terme de mes travaux.

Une autre extension du modèle de processus de Poisson dans le cas scalaire consiste à inclure les conversions entre ondes P et S lors des événements de diffusion. Une onde incidente P dans un milieu peut ainsi se propager en se convertissant en onde S, et éventuellement revenir en P, et ce plusieurs fois. Un modèle simple permettant de considérer ces conversions peut être de considérer une chaîne de Markov cachée à deux états (P et S). On

2. Certains travaux de Y. Tian [Tian 2000] sur les représentations matricielles de certains octonions entretiennent l'espoir de développer la TF octonionique.

peut alors imaginer écrire le processus de diffusion multiple comme l'entrelacement de deux processus de Poisson (un pour chaque type d'onde), avec des probabilités de transition P/S et S/P à chaque événement de diffusion. D'un point de vue "estimation", pouvoir estimer ces coefficients de conversion à partir d'une série temporelle (peu envisageable d'un point de vue expérimental) ou d'une distribution de diffusion est d'un intérêt non négligeable car cette information sur la proportion de conversion apporte une information fine sur le milieu diffusant. D'un point de vue technique, la non-commutativité des variables rend le problème très compliqué : l'ordre des conversions influe sur le résultat final. J'ai commencé à travailler sur ce problème avec une simplification importante qui consiste à paramétrer la fonction de phase des diffuseurs par un seul paramètre d'anisotropie et d'assumer la commutativité des variables. On revient à ce moment-là à un mélange de lois et il est possible de tenter l'estimation des taux de conversion P/S. Réintroduire la non-commutativité dans ce problème fait partie de mes pistes de recherche actuelles.

4.2.2 Taches solaires

Depuis juin 2010, je participe au projet sunspot (projet IXXI Systèmes Complexes) piloté par J.-F. Coeurjolly. Certains des thèmes de ce projet de recherche ont trait aux processus ponctuels sur la sphère. Les données (disponibles jusqu'au milieu du $XIX^{\text{ème}}$ siècle) comportent des informations temporelles (apparition, durée de vie) et spatiales (position, type, taille). Nous avons adopté une approche originale pour appréhender ces données en tentant de modéliser ce phénomène par un processus spatio-temporel sur la sphère, et ce afin de développer des techniques d'estimation statistique, comme par exemple de la prédiction de séries temporelles issues d'observations du processus spatio-temporel de localisation des taches.

Ces travaux s'inscrivent dans la continuité de l'étude des processus sur la sphère que j'ai déjà menée. Mais, dans le cas des taches solaires, le processus que nous regardons est indexé par la sphère. C'est donc une nouvelle piste de recherche que je développe actuellement, toujours tout de même sur un axe alliant géométrie et signal.

4.3 Processus non-holonomes

Les travaux entrepris sur la phase géométrique dans les milieux aléatoires suggèrent l'étude plus large de processus à valeurs dans des espaces particuliers : les fibrés vectoriels. Les processus évoluant dans de tels espaces avec des contraintes de transport parallèle dans la fibre au dessus de l'espace de base doivent exhiber une phase géométrique lorsque leur trajectoire se reboucle. Je souhaite m'intéresser dans le futur à ces processus et d'une manière générale aux signaux possédant une phase géométrique. Ces signaux sont des cas particuliers de signaux à valeurs sur des espaces courbes.

4.3.1 Chaînes de Markov

La contrainte de transport parallèle pour un processus aléatoire sur un espace courbe engendre la perte d'indépendance des incréments. Bien que dans le cas du processus de Poisson sur $SO(3)$ nous ayons pu contourner ce problème pour obtenir un processus de Lévy à droite, il n'est pas garanti que cela soit possible dans tous les cas de figure.

Afin de gérer cette dépendance, j'envisage d'utiliser des processus de pseudo-Poisson (ou processus de Feller), qui sont semblables à des processus de Poisson composés, avec non plus des événements indépendants mais une chaîne de Markov. Ces processus sont connus sur $\mathbb{R}$, mais leur étude sur des espaces courbes n'est pas très répandue. Je souhaite donc

travailler à l'étude de ces processus et souhaite tenter de voir s'ils permettent d'observer et prédire l'apparition de phases géométriques.

4.3.2 Phase géométriques, processus et analyse de Fourier géométrique

Enfin, parmi toutes les perspectives de mes travaux, la dernière que je mentionne ici tente de faire un lien entre mes travaux antérieurs. L'analyse des processus possédant une phase géométrique peut se faire par une AHNC sur l'espace courbe où le processus évolue. Je pense qu'il est possible de décrire spectralement ces signaux avec d'autres outils : les transformations de Fourier géométriques. Ces transformations (dont la TFQ est un cas particulier) sont basées sur les algèbres de Clifford. J'envisage d'étudier leurs propriétés et de voir comment elles pourraient permettre de caractériser les processus non-holonomes. En particulier, l'information géométrique étant disponible *via* la phase de ces transformées, il est possible que les phases géométriques s'expriment naturellement dans le formalisme des transformations de Fourier géométriques. Cette piste de recherche me tient à coeur car elle fait une sorte de connexion entre les travaux présentés dans ce manuscrit.

Conclusion

Les travaux exposés dans ce manuscrit présentent quelques contributions au traitement de signaux dont les échantillons sont à valeurs sur des structures algébriques non-commutatives. Cette propriété de non-commutativité a de multiples conséquences dans le traitement. Nous avons présenté quelques pistes de recherche prenant en compte cette spécificité.

L'utilisation des quaternions en traitement du signal a connu un certain essor ces dernières années, quoique toujours marginale dans la communauté. Pour le moment, beaucoup de travaux se résument à une extension systématique des algorithmes développés depuis plusieurs années pour les signaux à valeurs réelles et complexes. Ce n'est sans doute pas dans cette voie que l'on tirera avantage de l'algèbre des quaternions. Ces derniers sont, comme les complexes dans le plan, le maillon entre algèbre et géométrie dans l'espace Euclidien 3D. Il y a sans doute plus d'avantages à tirer du formalisme des quaternions dans l'analyse de Fourier géométrique ou dans l'analyse des signaux de type rotation ou à valeurs sur la sphère. La position des quaternions au sein des algèbres géométriques (algèbres de Clifford) doit, à mon sens, être exploitée pour les signaux 3D, et ce tant au niveau de l'analyse spectrale que du filtrage. C'est dans cette direction que les extensions des travaux présentés dans les premiers chapitres de ce manuscrit vont se diriger dans un futur proche. Ainsi, les résultats concernant le signal hyperanalytique semblent ouvrir une voie intéressante dans l'analyse géométrique du signal.

Pour les signaux à valeurs sur $SO(3)$ et S^2 , nous avons montré comment ils permettent de modéliser des problèmes physiques et comment ils permettent de prédire l'apparition de certains phénomènes physiques géométriques (phase géométrique, dépolarisation) de manière rigoureuse. L'étude de ces signaux nécessite d'utiliser des outils mathématiques issus de la théorie des groupes (théorie de la représentation) et de la géométrie différentielle (holonomie, transport parallèle). À l'aide de ces outils, il est possible de développer des algorithmes de traitement du signal qui prennent en compte la spécificité de ces signaux. Par exemple, nous avons montré l'intérêt des processus de Poisson composés sur le groupe des rotations pour la modélisation de la diffusion multiple dans le cas des ondes polarisées ou scalaires. Les applications présentées sont fondées sur les processus de Lévy. Nous avons montré quelques résultats d'estimation pour ces processus, se basant systématiquement sur la théorie de la représentation. Ces techniques non-paramétriques ont un fort potentiel dans l'étude des systèmes physiques. Il y a aussi sans doute intérêt à considérer des techniques paramétriques dans la continuité de ces travaux. C'est une des suites possibles parmi d'autres, comme la prise en compte des ondes P et S dans un même modèle de processus de Poisson ou de chaîne de Markov cachée.

Une autre continuation naturelle et intéressante de ces travaux est de considérer les signaux dont les accroissements ne sont plus indépendants. Par exemple, on peut citer le processus de pseudo-Poisson (ou processus de Feller) qui consiste en un processus de Poisson composé dont les éléments de base ne sont plus des variables indépendantes mais une chaîne de Markov. L'étude de tels processus sur les groupes de Lie est assez marginale et principalement menée par des mathématiciens. Mais ces processus ont sans doute un fort potentiel pour la modélisation et l'analyse de phénomènes de couplages entre diffuseurs dans la propagation en milieux aléatoires par exemple. Une contrainte de transport parallèle doit également pouvoir s'exprimer avec de tels signaux. Un autre exemple est la famille des processus non-holonomes : processus qui après une trajectoire fermée dans l'espace des états (espace non plat) ne reviennent pas à leur configuration initiale. Comme nous l'avons montré

pour la phase géométrique, le transport parallèle est une contrainte pour la trajectoire des signaux, et cette contrainte doit pouvoir s'exprimer *via* des processus de pseudo-Poisson. Cette perspective est une des pistes les plus prometteuses à mon avis pour l'étude des signaux possédant une phase géométrique.

Enfin, pour ce qui est des statistiques sur les structures non-commutatives, une autre extension possible est l'étude des processus ponctuels sur des variétés. Cette fois-ci, le processus est à valeurs réelles, mais est indexé par la variété (champ aléatoire). L'étude des taches solaires que nous avons entamée est fondée sur cette approche processus. C'est une piste de recherche nouvelle, mais également un point de rencontre entre plusieurs travaux présentés dans ce manuscrit.

D'une manière générale, je compte continuer à m'intéresser aux signaux et processus dont les échantillons sont à valeurs sur des structures non-commutatives et les espaces courbes, en lien étroit avec la physique. L'apport de la géométrie dans le traitement de ces signaux est encore limité et il reste encore sans doute beaucoup à faire en *traitement du signal géométrique*, avec toujours, pour ma part, un lien privilégié avec la Physique.

Troisième partie

Curriculum Vitae

Le Bihan Nicolas

Né le 22 Mai 1974

Nationalité Française

Marié, trois enfants

Adresse personnelle :

6 Rue de Poisat
38320 POISAT
France

Adresse professionnelle :

GIPSA-Lab,
961 Rue de la Houille Blanche,
Domaine Universitaire, B.P. 46,
38402 SAINT MARTIN D'HERES Cedex
FRANCE

Tél : + 33 4 76 82 64 86

Fax : + 33 4 76 82 63 84

Email : Nicolas.Le-Bihan@gipsa-lab.inpg.fr

<http://www.gipsa-lab.inpg.fr/~nicolas.le-bihan/>

Chargé de recherche CNRS au GIPSA-Lab

1 Positions

2002- Chargé de recherche CNRS au GIPSA-Lab, Grenoble, France.

2001-02 ATER à l'ENSIEG, INPG, Grenoble, France.

2 Cursus

2001 Thèse de doctorat de l'INPG avec les félicitations du jury

Specialité : Traitement du signal

Title : "Linear and multilinear algebra processing of vector signals. Application to multicomponent seismic wave separation."

Jury : M. Basseville, S. Bourennane, P. Comon, P. Duvaut, S.J. Sangwine, J. Meunier, J. Mars, F. Glangaud.

1998 DEA de l'INPG, mention AB.

1992-97 Maîtrise de Physique de l'UBO (Université de Bretagne Occidentale, Brest, France), mention AB.

1992 Baccalauréat C, mention AB.

3 Collaborations internationales

- **Jan. 2003 - Jan. 2005** : "Visiting Fellow" au Dpt. of Computing and Eletronic Systems, University of Essex, Colchester, UK (Collab. avec Dr. S.J. Sangwine).
- **Jun. 2007 - May 2008** : "Visiting Fellow" au the Dpt. of Information Engineering, RSISE, Australian National University, Canberra, Australia (Collab. avec Pr. J.H. Manton).
- **Jun. 2008** : "Visiting Fellow" au the Dpt. of Electrical Engineering, The University of Melbourne, Melbourne, Australia (Collab. avec Pr. J.H. Manton).

4 Projets de recherche

- **Avr. 2007 - Avr. 2009** : Coordinateur du projet de recherche “Analyse statistique non-commutative de signaux sismologiques vectoriels : Application à la mesure d’hétérogénéité de la croûte terrestre”. **Projet PEPS CNRS** du département ST2I. Collaborateurs : Ludovic Margerin (CEREGE, Aix-en-Provence, France), Salem Said (GIPSA-Lab, Grenoble, France) et J.H. Manton (RSISE, ANU, Canberra, Australia).
- **Avr. 2009 - Avr. 2010** : Membre du projet “Phase géométrique des ondes élastiques”. **Projet PEPS CNRS** de l’Institut INP du CNRS. Collaborateurs : Vincent Rossetto (LPMMC, Grenoble, France) and Ludovic Margerin (CEREGE, Aix-en-Provence, France).
- **Juin 2011 - Juin 2013** : Membre du projet “Sunspots : Modelisation et inférence statistique pour le phénomènes des taches solaires”. **Projet de l’Institut Rhône-Alpin des Systèmes complexes (IXXI)**. Collaborateurs : J.-F. Coeurjolly (LJK, Grenoble, France), P.O. Amblard (GIPSA-Lab, Grenoble, France) et J. Liliensten (LPG, Grenoble, France).

5 Encadrement doctoral

J’ai co-encadré 4 thèses déjà soutenues et je co-encadre 1 thèse en cours actuellement.

- ANNABELLE OLLIVIER
Titre : Une nouvelle approche dans l’extraction de paramètres en altimétrie spatiale radar.
Encadrants : Jean-Louis LACOUME, Ouan Zan ZANIFE et Nicolas LE BIHAN
Début : Janvier 2003
Soutenance : 31 Mars 2006
- SEBASTIAN MIRON
Titre : Techniques multilinéaires et hypercomplexes en traitement d’antenne haute résolution.
Encadrants : Jérôme MARS et Nicolas LE BIHAN
Début : Octobre 2002
Soutenance : 4 Octobre 2005
Distinction : Prix de thèse INPG 2006 (specialité : Sciences de la Terre)
- SALEM SAID
Titre : Estimation and filtering of processes in matrix Lie groups
Encadrants : Nicolas LE BIHAN et Stephen J. SANGWINE
Début : Octobre 2006
Soutenance : 17 Décembre 2009
- MANUEL HOBIGER
Titre : 3D characterization of seismological signals using multicomponent signal processing. Application to the study of seismic aléa.
Encadrants : Pierre-Yves BARD, Cécile CORNOU et Nicolas LE BIHAN
Début : Octobre 2007
Soutenance : 19 janvier 2011

- JÉRÉMIE BOULANGER

Titre : Information, géométrie et processus aléatoires sur les variétés : Application à l'étude des milieux aléatoires et des systèmes complexes.

Encadrants : Nicolas LE BIHAN et Jonathan H. MANTON

Spécificité : Co-tutelle entre l'Université de Grenoble et l'Université de Melbourne

Début : Octobre 2010

Soutenance : Prévue en octobre 2013

6 Expertise de thèses (rapporteur)

J'ai été rapporteur pour les thèses suivantes :

- GIOVANNI M. MENANNO

Titre : Seismic multicomponent deconvolution and wavelet estimation by means of quaternions

Université : Université de Pise, Italie

Date : Mars 2010

- CHI LI

Titre : Tracking on manifolds, Quasi-Newton Optimisation algorithms on manifolds

Université : Australian National University (ANU), Canberra, Australie.

Date : Mai 2010

7 Participation à la vie de la recherche

- Relecteur pour les revues : *IEEE Transactions on Signal Processing*, *IEEE Transaction on Information theory*, *Elsevier Signal Processing*, *Geophysics*, *IEEE Journal of Selected Topics in Earth Observations and Remote Sensing*, *IEEE Transactions on Circuits and Systems II*, *Signal, Image and Video Processing*, *Applied Mathematics and Computation*.
- Membre de la commission de spécialistes INPG, Section 61 (2005-2007)
- Membre élu du conseil de laboratoire du Gipsa-Lab (2011 -)
- Co-organisateur du Workshop "Transport of elastic Waves in heterogeneous media", Col de Porte, 14-15 Janvier 2010.
- Co-organisateur de la session Tutorial "Quaternion Signal Processing", Eusipco, Août 2011, Barcelone.

8 Publications et communications

8.1 Participations à des ouvrages [2]

- [1] **N. Le Bihan**, V. Vrabie and J. Mars, *Multidimensional seismic data decomposition by higher order SVD and unimodal-ICA*, in "Signal and Image Processing for remote sensing", edited by C.H. Chen, CRC Press (Taylor and Francis books), Sept. 2006.
- [2] S.J. Sangwine, T.A. Ell and **N. Le Bihan**, *Hypercomplex models and processing for vector images*, in "Multivariate Image Processing", edited by J. Chanussot, C. Collet and K. Chehdi, John Wiley & Sons, 2010.

8.2 Articles dans des revues internationales avec comité de lecture [17]

- [3] **N. Le Bihan** and G. Ginolhac, *Three-Mode Dataset Analysis Using Higher Order Subspace Method : Application to Sonar and Seismo-acoustic Signal Processing*, **Signal Processing**, Vol. 84, Issue 5, pp. 919-942, 2004.
- [4] **N. Le Bihan** and J. Mars, *Singular Value Decomposition of Matrices of Quaternions : A New Tool for Vector-Sensor Signal Processing*, **Signal Processing**, Vol. 84, Issue 7, pp. 1177-1199, 2004.
- [5] S. Miron, **N. Le Bihan** and J. Mars, *Vector-MUSIC for Polarized Seismic Source Localisation*, **EURASIP Journal on Applied Signal Processing**, Vol. 2005, No. 1, pp. 74-84, 2005.
- [6] S. Miron, **N. Le Bihan** and J. Mars, *Quaternion-MUSIC for vector-sensor array processing*, **IEEE Transactions on Signal Processing**, Volume 54, No. 4, 1218 - 1229, 2006.
- [7] V. Vrabie, **N. Le Bihan** and J. Mars, *Multicomponent wave separation using HOSVD - unimodal ICA subspace method*, **Geophysics**, Volume 71, Issue 5, pp. V133-V143, 2006.
- [8] S.J. Sangwine and **N. Le Bihan**, *Quaternion Singular Value Decomposition based on Bidiagonalization to a Real or Complex Matrix using Quaternion Householder Transformations*, **Applied Mathematics and Computation**, Volume 182, Issue 1, pp. 727-738, 2006.
- [9] **N. Le Bihan** and S.J. Sangwine, *Jacobi Method For Quaternion Matrix Singular Value Decomposition*, **Applied Mathematics and Computation**, Vol. 187, Iss. 2, pp. 1265-1271, 2007.
- [10] **N. Le Bihan**, S. Miron and J. Mars, *MUSIC algorithm for vector-sensors array using biquaternions*, **IEEE Transactions on Signal Processing**, Vol. 55, No. 9, pp. 4523-4533, 2007.
- [11] S. Said, **N. Le Bihan** and S.J. Sangwine, *Fast complexified quaternion Fourier transform*, **IEEE Transactions on Signal Processing**, Volume 56, No. 4, pp 1522-1531, 2008.
- [12] S. Said and **N. Le Bihan**, *Higher Order Statistics of Stokes Parameters in a Random Birefringent Medium*, **Waves in Random and Complex Media**, Vol. 18, No. 2, pp. 275-292, 2008.
- [13] S. Buchholz and **N. Le Bihan**, *Polarized Signal Classification by Complex and Quaternionic Multi Layer Perceptrons*, **International Journal of Neural Systems**, Vol. 18, No. 2, pp. 75-85, 2008.
- [14] **N. Le Bihan** and L. Margerin, *Nonparametric estimation of the heterogeneity of a random medium using Compound Poisson Process modeling of wave multiple scattering*, **Physical Review E**, 80, 016601, 2009.

- [15] M. Hobiger, P.Y. Bard, C. Cornou and **N. Le Bihan**, *Single Station Determination of Rayleigh Wave Ellipticity by Using the Random Decrement Technique (RayDec)*, **Geophysical Research Letters**, 36, L14303, 2009.
- [16] S.J. Sangwine and **N. Le Bihan**, *Quaternion Polar Representation with a Complex Modulus and Complex Argument Inspired by the Cayley-Dickson Form*, **Advances in Applied Clifford Algebras**, Vol. 36, No 1, pp. 111-120, 2010.
- [17] S. Said, C. Lageman, **N. Le Bihan** and J.H. Manton, *Decompounding on compact Lie groups*, IEEE Transactions on Information Theory, Vol. 56, Issue 6, pp. 2766-2777, 2010.
- [18] G.M. Menanno and **N. Le Bihan**, *Quaternion polynomial matrix diagonalization for the separation of polarized convolutive mixture*, **Signal Processing**, Vol. 90, Issue 7, pp. 1229-1337, 2010.
- [19] S.J. Sangwine, T.A. Ell and **N. Le Bihan**, *Fundamental representations and algebraic properties of biquaternions or complexified quaternions*, to appear in **Advances in Applied Clifford Algebras**.

8.3 Communications dans des conférences internationales avec comité de lecture [34]

- [20] **N. Le Bihan** and J.-L. Lacoume, *Complex seismic signal and blind wave separation*, 61st Conference of the European Association of Geophysicists and Engineers (EAGE), Helsinki, Finland, Expanded Abstract, 6-22, june 1999.
- [21] **N. Le Bihan** and J. Mars, *Blind wave separation using vector-sensors*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Istanbul, Turkey, june 2000.
- [22] **N. Le Bihan**, J. Mars and H. Pedersen, *Surface wavefield separation on multicomponent sensors*, 62nd Conference of the European Association of Geophysicists and Engineers (EAGE), Glasgow, Scotland, Expanded Abstract, L-03, june 2000.
- [23] **N. Le Bihan** and G. Ginolhac, *Subspace methods for 3D arrays*, SEE Conference on Physics in Image and Signal Processing (PSIP), Marseille, France, pp. 359-364, January 2001.
- [24] **N. Le Bihan** and J. Mars, *New 2D complex and hypercomplex seismic attributes*, 71st Conference of the Society of Exploration Geophysicists (SEG), San Antonio, Texas, USA, September 2001.
- [25] **N. Le Bihan**, S. Lariani and J. Mars, *Seismic cube decomposition before reservoir characterisation*, 71st Conference of the Society of Exploration Geophysicists (SEG), San Antonio, Texas, USA, September 2001.
- [26] **N. Le Bihan**, F. Glangeaud, J. Mars, J.-L. Mari and J. Meunier, *Noise rejection by multicomponent SVD-Wiener joint filters*, 64th conference of the EAGE, Florence, Italy, May 2002.
- [27] **N. Le Bihan** and J. Mars, *Subspace method for vector-sensor wave separation based on quaternion algebra*, XI European Signal Processing Conference (EUSIPCO), Toulouse, France, Sept. 2002.
- [28] V. Vrabie, **N. Le Bihan** and J. Mars, *3D-SVD and partial ICA for 3D array sensor*, 72nd Conference of the Society of Exploration Geophysicists (SEG), Salt Lake City, USA, October 2002.
- [29] S. Miron, M. Guillon, **N. Le Bihan** and J. Mars, *Multidimensional signal processing using quaternions*, SEE Conference on Physics in Image and Signal Processing (PSIP), Grenoble, France, January 2003.
- [30] P. Bas, **N. Le Bihan** and J.-M. Chassery, *Color image watermarking using Quaternion Fourier transform*, IEEE International conference on Acoustics speech and signal processing (ICASSP), Hong-Kong, June 2003.

- [31] **N. Le Bihan** and S.J. Sangwine, *Color Image Decomposition Using Quaternion Singular Value Decomposition*, IEE International conference on Visual Information Engineering (VIE), Guildford, UK, July 2003.
- [32] **N. Le Bihan** and S.J. Sangwine, *Quaternion Principal Component Analysis of Color images*, IEEE International Conference on Image Processing (ICIP), Barcelona, Spain, September 2003.
- [33] S. Miron, **N. Le Bihan** and J. Mars, *Joint estimation of direction of arrival and polarization parameters for multicomponent sensor array*, 66th conference of the EAGE, Paris, France, June 2004.
- [34] S. Miron, **N. Le Bihan** and J. Mars, *Polarized source characterization using Vector-MUSIC*, XII European Signal Processing Conference (EUSIPCO), Vienna, Austria, September 2004.
- [35] P.-O. Amblard and **N. Le Bihan**, *On properness of quaternion valued random variables*, IMA Conference on Mathematics in Signal Processing, Cirencester, UK, December 2004.
- [36] A. Ollivier, **N. Le Bihan**, J.-L. Lacoume and O.Z. Zanife, *Improving speckle filtering with SVD to extract ocean parameters from altimeter radar echoes*, SEE Conference on Physics in Image and Signal Processing (PSIP), Toulouse, France, January 2005.
- [37] S. Miron, **N. Le Bihan** and J. Mars, *High Resolution vector-sensor array processing using quaternions*, IEEE Workshop on Statistical Signal Processing (SSP), Bordeaux, France, 2005.
- [38] S. Miron, **N. Le Bihan** and J. Mars, *High resolution vector-sensor array processing based on biquaternions*, IEEE International Conference on Acoustics speech and signal processing (ICASSP), Toulouse, France, May 2006.
- [39] S. Buchholz and **N. Le Bihan**, *“Optimal separation of polarized signals by quaternionic neural networks”*, XIV European Signal Processing Conference (EUSIPCO), Florence, Italy, September 2006.
- [40] S.J. Sangwine and **N. Le Bihan**, *“Computing the SVD of a quaternion matrix”*, 7th IMA Conference on Mathematics in Signal Processing, Cirencester, UK, December 2006.
- [41] **N. Le Bihan** and S. Buchholz, *“Quaternionic independent component analysis using hypercomplex nonlinearities”*, 7th IMA Conference on Mathematics in Signal Processing, Cirencester, UK, December 2006.
- [42] **N. Le Bihan** and P.-O. Amblard, *“Detection and estimation of Gaussian proper quaternion valued random processes”*, 7th IMA Conference on Mathematics in Signal Processing, Cirencester, UK, December 2006.
- [43] S. Said, **N. Le Bihan** and J.H. Manton, *“Neyman-Pearson detection on $SU(2)$: an application to polarized signals”*, Workshop on Physics in Signal and Image Processing (PSIP), Mulhouse, France, January 2007.
- [44] S. Said, N. Courty, **N. Le Bihan** and S.J. Sangwine, *“Exact principal geodesic analysis for data on $SO(3)$ ”*, XV European Signal Processing Conference (EUSIPCO), Poznan, Poland, September 2007.
- [45] S.J. Sangwine and **N. Le Bihan**, *“Hypercomplex analytic signals : extension of the analytic signal concept to complex signals”*, XV European Signal Processing Conference (EUSIPCO), Poznan, Poland, September 2007.
- [46] **N. Le Bihan** and S.J. Sangwine, *“The H -analytic signal”*, XVI European Signal Processing Conference (EUSIPCO), Lausanne, Switzerland, August 2008.

- [47] S. Said, C. Lageman, **N. Le Bihan** and J.H. Manton, “*Decompounding on compact Lie groups*”, 8th IMA conference on Mathematics in signal processing, Cirencester, UK, December 2008.
- [48] **N. Le Bihan** and S.J. Sangwine, “*About the extension of the 1D analytic signal to improper complex valued signals*”, 8th IMA conference on Mathematics in signal processing, Cirencester, UK, December 2008.
- [49] S. Said and **N. Le Bihan**, *A functional approach to signal processing under nonlinear constraints*, 8th IMA conference on Mathematics in signal processing, Cirencester, UK, December 2008.
- [50] S. Said, C. Lageman, **N. Le Bihan** and J.H. Manton, *Nonparametric estimation for compound Poisson processes on compact Lie groups*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Taipei, Taiwan, April 2009.
- [51] S. Said, **N. Le Bihan** and S.J. Sangwine, *A stability approach to the analysis of rotation time series*, 15th IFAC Symposium on systems identification, Saint-Malo, France, July 2009.
- [52] M. Hobiger, **N. Le Bihan**, C. Cornou and P.-Y. Bard, *Rayleigh wave ellipticity estimation from ambient seismic noise using single and multiple vector-sensor techniques*, XVII European Signal Processing Conference (EUSIPCO), Glasgow, Scotland, August 2009.
- [53] S. Javidi, C. Cheong Took, C. Jahanchahi, **N. Le Bihan** and D. Mandic, *Blind extraction of improper quaternion sources*, IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, May 2011.

8.4 Conférences nationales avec comité de lecture [9]

- [54] **N. Le Bihan**, J.-L. Lacoume, F. Glangaud and S. Gaffet, *Traitement des ondes sismo-acoustiques UBF multicomposantes*, 4^{ème} journées d’acoustique sous-marine, Brest, France, December 1998.
- [55] **N. Le Bihan**, and J. Mars, *Technique d’algèbre multilinéaire pour le traitement sismique multicomposantes*, 5^{ème} journées d’acoustique sous-marine, Brest, France, December 2000.
- [56] **N. Le Bihan**, J.-L. Lacoume et J. Mars, *Séparation d’ondes sur un réseau de capteurs multidimensionnel*, Colloque GRETSI, Toulouse, France, September 2001.
- [57] **N. Le Bihan**, and S.J. Sangwine, *Analyse de signaux vectoriels basée sur le modèle quaternionique*, Colloque GRETSI, Paris, France, September 2003.
- [58] J. Mars, **N. Le Bihan**, and V. Vrabie, *Extensions de la SVD aux données multidimensionnelles, Application à la séparation de sources*, Colloque GRETSI, Paris, France, September 2003.
- [59] P. Bas, **N. Le Bihan**, and J.-M. Chassery, *Utilisation de la Transformé de Fourier Quaternionique en tatouage d’images couleur*, Colloque GRETSI, Paris, France, September 2003.
- [60] S. Miron, **N. Le Bihan**, and J. Mars, *Étude des performances du modèle quaternionique en traitement d’antenne vectorielle*, Colloque GRETSI, Louvain-La-Neuve, Belgium, September 2005.
- [61] **N. Le Bihan**, *Diagonalisation de matrices polynomiales quaternioniques : application à la séparation de mélanges convolutifs d’ondes polarisées*, Colloque GRETSI, Louvain-La-Neuve, Belgium, September 2005.
- [62] **N. Le Bihan**, L. Margerin and S. Salem, *Infrence statistique sur la fonction de phase des milieux alatoires par diffusion multiple des ondes*, Colloque GRETSI, Dijon, France, September 2009.

8.5 Articles soumis [4]

- [63] J. Boulanger, **N. Le Bihan**, S. Catheline and V. Rossetto, *Non-adiabatic geometric phase of elastic waves*, Soumis à **Journal of Acoustical Society of America**, Avril 2011.
- [64] J.F. Coeurjolly and **N. Le Bihan**, *Geodesic Normal distribution on the circle*, Soumis à **Metrika**, Février 2011.
- [65] M. Hobiger, C. Cornou, P.-Y. Bard, **N. Le Bihan** and M. Wathelet, *Inversion of Rayleigh wave ellipticity measurements, Part I : Theoretical aspects*, Soumis à **Geophysical Journal International**, Janvier 2011.
- [66] M. Hobiger, C. Cornou, G. Di Giulio, B. Endrun, F. Renalier, A. Savvaidis, P.-Y. Bard, S. Hailemichael, **N. Le Bihan**, M. Ohrnberger, N. Theodulidis and M. Wathelet, *Inversion of Rayleigh wave ellipticity measurements, Part II : Application to real data measurements*, Soumis à **Geophysical Journal International**, Janvier 2011.

Bibliographie

- [Aharonov 1987] Y. Aharonov et J. Anandan. *Phase change during a cyclic quantum evolution*. Physical review Letters, vol. 58, pages 1593 – 1596, 1987. 112
- [Aki 2002] K. Aki et P.G. Richards. Quantitative seismology. University Science books, 2nd Edition, 2002. 99
- [Alexiadis 2009] D.S. Alexiadis et G.D. Sergiadis. *Estimation of Motions in Color Image Sequences Using Hypercomplex Fourier Transforms*. IEEE Transactions on Image Processing, vol. 18, no. 1, pages 168 –187, 2009. 13
- [Altman 1986] S.L. Altman. Rotations, quaternions and double groups. Dover, 1986. 6, 9, 10
- [Amblard 2004] P.-O. Amblard et N. Le Bihan. *On properness of quaternion valued random variables*. IMA Conference on mathematics in signal processing, 2004. 9, 16, 65
- [Anandan 1988] J. Anandan. *Non-adiabatic non-abelian geometric phase*. Physics Letters A, vol. 133, 1988. 112
- [Applebaum 2000] D. Applebaum. *Compound Poisson processes and Lévy processes in groups and symmetric spaces*. Annals of probability, vol. 13, no. 2, pages 383 – 425, 2000. 104
- [Bas 2003] P. Bas, N. Le Bihan et J.-M. Chassery. *Color image watermarking using quaternion Fourier transform*. IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, pages III – 521–4, 2003. 13, 63
- [Belinfante 1966] J.G. Belinfante, B. Kolman et H.A. Smith. *An introduction to Lie groups and Lie algebra, with applications*. SIAM Review, vol. 8, no. 1, pages 11 – 46, 1966. 8
- [Berry 1984] M.V. Berry. *Quantal phase factors accompanying adiabatic changes*. Proceedings of the Royal Society of London A, vol. 392, pages 45 – 57, 1984. 109, 111
- [Bhandari 1988] R. Bhandari et J. Samuel. *Observation of topological phase by use of a laser interferometer*. Physical Review Letters, vol. 60, pages 1211 – 1213, 1988. 111
- [Bogdanova 205] I. Bogdanova. *Wavelets on Non-Euclidian manifolds*. PhD thesis, École Polytechnique fédérale de Lausanne, 205. 102
- [Boulanger 2010] J. Boulanger. A study of the geometric phase of elastic waves. Rapport de stage de Master, 2010. 113, 114
- [Boulanger 2011] J. Boulanger, N. Le Bihan, S. Catheline et V. Rossetto. *Non-adiabatic geometric phase of elastic waves*. submitted to the Journal of the Acoustical Society of America, 2011. <http://arxiv.org/abs/1103.4506>. 111, 112, 113
- [Brenner 1951] J.L. Brenner. *Matrices of quaternions*. Pacific J. Math., vol. 1, pages 329 – 335, 1951. 11
- [Brosseau 1998] C. Brosseau. Fundamentals of polarized light. a statistical approach. Wiley-Interscience, 1998. 65, 100, 105
- [Bryc 2009] W. Bryc et V. Pierce. *Duality of real and quaternionic random matrices*. Electronic journal of probability, vol. 14, no. 17, pages 452 – 476, 2009. 155
- [Buchholz 2006] S. Buchholz et N. Le Bihan. *Optimal separation of polarized signals by quaternionic neural networks*. XIV European Signal Processing Conference (EU-SIPCO), Florence, Italy, 2006. 63, 65

- [Buchholz 2008] S. Buchholz et N. Le Bihan. *Polarized signal classification by complex and quaternionic multi-layer perceptrons*. International journal of neural systems, vol. 18, no. 2, pages 75 – 85, 2008. 63, 65
- [Bulow 2001] T Bulow et G Sommer. *Hypercomplex signals-a novel extension of the analytic signal to the multidimensional case*. IEEE Transactions on signal processing, vol. 49, no. 11, pages 2844–2852, 2001. 5, 13
- [Bunse-Gerstner 1989] A. Bunse-Gerstner, R. Byers et V. Mehrmann. *A quaternion QR algorithm*. Numerische Mathematik, vol. 55, pages 83 – 95, 1989. 12
- [Carmo 1976] M. Do Carmo. Differential geometry of curves and surfaces. Prentice Hall, 1976. 110, 111
- [Chirikjian 2000] G.S. Chirikjian et A.B. Kyatkin. Engineering applications of noncommutative harmonic analysis. CRC Press, 2000. 102, 103, 156
- [Chyba 1988] T.H. Chyba, L.J. Wang et L. Mandel. *Measurement of the Pancharatnam phase for a light beam*. Optics Letters, vol. 13, no. 7, pages 562 – 564, 1988. 111
- [Cont 2003] R. Cont et P. Tankov. Financial modelling with jump processes. Chapman & Hall / CRC Press, 2003. 104
- [Coxeter 1946] H.S.M Coxeter. *Quaternions and reflections*. The american mathematical monthly, vol. 53, no. 3, pages 136–146, 1946. 9
- [Delsuc 1988] M.A. Delsuc. *Spectrum representation of 2D NMR spectra by hypercomplex numbers*. Journal of magnetic resonance, vol. 77, pages 119 – 124, 1988. 12
- [Dieudonné 1943] J. Dieudonné. *Les déterminants sur un corps non commutatif*. Bulletin de la S.M.F., vol. 71, pages 27–45, 1943. 11
- [Dieudonné 1980] J. Dieudonné. Special functions and linear representations of lie groups. American Mathematical Society, 1980. 102, 103
- [Ell 1992] T.A. Ell. *Hypercomplex spectral transformations*. Ph.D. Dissertation, University of Minnesota, 1992. 12, 13
- [Ell 1993] T.A. Ell. *Quaternion Fourier transforms for analysis of two dimensional linear time invariant partial differential systems*. 32nd Conference on detection and control, pages 1830–1841, 1993. 12, 13
- [Ell 2007] T.A. Ell et S.J. Sangwine. *Hypercomplex Fourier Transforms of Color Images*. IEEE Transactions on Image Processing, vol. 16, no. 1, pages 22 –35, 2007. 13
- [Ernst 1987] R.R. Ernst, G. Bodenhausen et A. Wokaun. Principles of magnetic nuclear resonance in one and two dimensions. Oxford University Press, 1987. 12
- [Faure 2011] F. Faure. Géométrie et topologie pour la physique. note de cours du Master de Physique UJF, 2011. 110
- [Felsberg 2001] M. Felsberg et G. Sommer. *The monogenic signal*. IEEE Transactions on Signal Processing, vol. 49, no. 12, pages 3136 – 3144, 2001. 5
- [Feynman 1963] R. P. Feynman. The feynman lectures on physics, volume I. Addison-Wiley, 1963. 5
- [Gabor 1946] D. Gabor. *Theory of communication*. J. Inst. Elec. Eng., vol. 93, pages 429 – 457, 1946. 5, 70
- [Girard 2004] P.R. Girard. Quaternions, algèbre de clifford et physique relativiste. Presses polytechniques et universitaires romandes, 2004. 6
- [Gordon 2000] J.P. Gordon et H. Kogelnik. *PMD fundamentals : polarization mode dispersion in optical fibers*. Proc. Nat. Acad. Sci. USA, vol. 97, pages 4541 – 4550, 2000. 100

- [Grenander 1963] U. Grenander. Probabilities on algebraic structures. Hoboken, NJ : Wiley, 1963. 102
- [Hahn 1991] S. Hahn. *Multidimensional complex signals with single orthant spectra*. Proceedings of the IEEE, pages 1287 – 1300, 1991. 5
- [Hanson 2006] A.J. Hanson. Visualizing quaternions. Morgan Kaufmann publishers Elsevier, 2006. 6, 10
- [Hisada 1995] Y. Hisada. *An efficient method for computing Green's functions for a layered halfspace with sources and receivers at close depths*. Bulletin of the Seismological Society of America, vol. 84, pages 1456 – 1472, 1995. 66
- [Hobiger 2011] M. Hobiger. *Polarization of Surface waves : characterization, inversion and application to seismic hazard assessment*. PhD thesis, Université Joseph Fourier, 2011. 66, 69
- [Janovská 2003] D. Janovská et G. Opfer. *Givens' transformation applied to quaternion valued vectors*. BIT Numerical mathematics, vol. 43, pages 991 – 1002, 2003. 12
- [Javidi 2011] S. Javidi, C. Cheong took, C. Jahanchahi, N. Le Bihan et D. Mandic. *Blind extraction of improper quaternion sources*. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech republic, 2011. 63
- [Kantor 1989] I.L. Kantor et A.S. Solodovnikov. Hypercomplex numbers : an elementary introduction to algebras. Springer-Verlag, 1989. 6, 7, 8
- [Kostelec 2008] P.J. Kostelec et D.N. Rockmore. *FFTs on the rotation group*. Journal of Fourier analysis and applications, vol. 14, pages 145 – 179, 2008. 102
- [Kuipers 1999] J.B. Kuipers. Quaternions and rotation sequences : a primer with applications to orbits, aerospace and virtual reality. Princeton University Press, 1999. 9, 63, 72
- [Lacoume 1997] J.-L. Lacoume, P.-O. Amblard et P. Comon. Statistiques d'ordre supérieur pour le traitement du signal. Masson, Paris, 1997. 155
- [Le Bihan 2001] N. Le Bihan et J. Mars. *New 2D complex and hypercomplex seismic attributes*. Conference of the Society of Exploration Geophysics (SEG), San Antonio, Texas, USA, 2001. 63
- [Le Bihan 2003] N. Le Bihan et S.J. Sangwine. *Quaternion principal component analysis of Color images*. IEEE International Conference on Image Processing (ICIP), Barcelona, Spain, 2003. 63
- [Le Bihan 2004] N. Le Bihan et J. Mars. *Singular Value Decomposition of Matrices of Quaternions : A New Tool for Vector-Sensor Signal Processing*. Signal Processing, vol. 84, no. 7, pages 1177 – 1199, 2004. 12, 63, 64, 65
- [Le Bihan 2005] N. Le Bihan. *Diagonalisation de matrices polynomiales quaternioniques : application à la séparation de mélanges convolutifs d'ondes polarisées*. Colloque GRETSI, Louvain-La-Neuve, Belgium, September, 2005. 12, 65
- [Le Bihan 2006a] N. Le Bihan et P.-O. Amblard. *Detection and estimation of Gaussian proper quaternion valued random processes*. 7th IMA Conference on Mathematics in Signal Processing, Cirencester, UK, 2006. 16, 63
- [Le Bihan 2006b] N. Le Bihan et S. Buchholz. *Quaternionic independent component analysis using hypercomplex nonlinearities*. 7th IMA Conference on Mathematics in Signal Processing, Cirencester, UK, 2006. 16, 63, 65
- [Le Bihan 2007a] N. Le Bihan, S. Miron et J. Mars. *MUSIC algorithm for vector-sensors array using biquaternions*. IEEE Transactions on Signal Processing, vol. 55, no. 9, pages 4523 – 4533, 2007. 12, 65, 66

- [Le Bihan 2007b] N. Le Bihan et S.J. Sangwine. *Jacobi Method For Quaternion Matrix Singular Value Decomposition*. Applied mathematics and computation, vol. 187, pages 1265 – 1271, 2007. 12
- [Le Bihan 2008] N. Le Bihan et S.J. Sangwine. *The H-analytic signal*. XVI European Signal Processing Conference (EUSIPCO), Lausanne, Switzerland, 2008. 9, 70
- [Le Bihan 2009] N. Le Bihan et L. Margerin. *Nonparametric estimation of the heterogeneity of a random medium using Compound Poisson Process modeling of wave multiple scattering*. Physical Review E, vol. 80, no. 016601, 2009. 106, 108
- [Lee 1949] H.C. Lee. *Eigenvalues and canonical forms of matrices with quaternions coefficients*. Proc. Royal Irish Acad. Sect. A, vol. 52, pages 253 – 260, 1949. 11
- [Liao 2004a] M. Liao. *Lévy processes and Fourier analysis on compact Lie groups*. Annals of probability, vol. 32, no. 2, pages 1553 – 1573, 2004. 104
- [Liao 2004b] M. Liao. *Lévy processes in lie groups*. Cambridge University Press, 2004. 104
- [Lilly 2010] J.M. Lilly et S.C. Olhede. *Bivariate Instantaneous Frequency and Bandwidth*. IEEE Transactions on Signal Processing, vol. 58, no. 2, pages 591–602, 2010. 5
- [Lounesto 2001] P. Lounesto. *Clifford algebra and spinors*. Cambridge University press, 2001. 11
- [Maggs 2001] A.C. Maggs et V. Rossetto. *Whriting photons and Berry phases in polaried multiple scattering*. Physical review Letters, vol. 87, no. 253901, 2001. 113
- [Mandic 2011] D. Mandic. *A Quaternion Gradient Operator and Its Applications*. Signal Processing Letters, IEEE, vol. 18, no. 1, pages 47 – 50, 2011. 16
- [Menanno 2010] G.M. Menanno et N. Le Bihan. *Quaternion polynomial matrix diagonalization for the separation of polarized convolutive mixture*. Signal Processing, vol. 90, pages 1229 – 1337, 2010. 12, 63, 65
- [Miron 2005] S. Miron, N. Le Bihan et J. Mars. *High resolution vector-sensor array processing using quaternions*. IEEE Workshop on statistical signal processing, Bordeaux, France, 2005. 63
- [Miron 2006a] S. Miron, N. Le Bihan et J. Mars. *High resolution vector-sensor array processing based on biquaternions*. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toulouse, France, 2006. 63
- [Miron 2006b] S. Miron, N. Le Bihan et J. Mars. *Quaternion-MUSIC for vector-sensor array processing*. IEEE Transactions on Signal Processing, vol. 54, no. 4, pages 1218 – 1229, 2006. 63, 65
- [Ning 1995] X. Ning, L. Papiez et G. Sandison. *Compound-Poisson-process method for the multiple scattering of charged particles*. Physical Review E, vol. 52, pages 5621 – 5633, 1995. 107
- [Oprea 1995] J. Oprea. *Geometry and the Foucault pendulum*. American Mathematical Monthly, vol. 102, pages 515 – 522, 1995. 110
- [Picinbono 1994] B. Picinbono. *On circularity*. IEEE Transactions on Signal Processing, vol. 42, no. 12, pages 3473 – 3482, 1994. 5, 15
- [Picinbono 1995] B. Picinbono et P. Chevalier. *Widely linear estimation with complex data*. IEEE Transactions on Signal Processing, vol. 43, no. 8, pages 2030 – 2033, 1995. 70
- [QTFM 2005] QTFM. *Quaternion Toolbox for Matlab®*. [Online], 2005. Software library, written by Stephen Sangwine and Nicolas Le Bihan. Available : <http://qtfm.sourceforge.net/>. 16

- [Redfield 2002] S.A. Redfield et Q.Q. Huynh. *Hypercomplex Fourier transforms applied to detection for side-scan sonar*. OCEANS '02 MTS/IEEE, vol. 4, pages 2219 – 2224, 2002. 13
- [Rossetto 2002] V. Rossetto. *Mécanique statistique de systèmes sous contraintes : topologie de l'ADN et simulations électrostatiques*. PhD thesis, Université Paris 6, 2002. 113, 114
- [Said 2008a] S. Said et N. Le Bihan. *Higher-order statistics of Stokes parameters in a random birefringent medium*. Waves in random and complex media, vol. 18, no. 2, pages 275 – 292, 2008. 101, 104, 105
- [Said 2008b] S. Said, N. Le Bihan et S.J. Sangwine. *Fast complexified quaternion Fourier transform*. IEEE Transactions on Signal Processing, vol. 56, no. 4, pages 1522 – 1531, 2008. 15
- [Said 2010] S. Said, C. Lageman, N. Le Bihan et J.H. Manton. *Decompounding on compact Lie groups*. IEEE Transactions on Information theory, vol. 56, no. 6, pages 2766 – 2777, 2010. 103, 104, 106, 107, 108
- [Sangwine 1996] S.J. Sangwine. *Fourier transforms of colour images using quaternion or hypercomplex numbers*. Electronics Letters, vol. 32, no. 21, pages 1979 – 1980, 1996. 6, 13
- [Sangwine 2006] S.J. Sangwine et N. Le Bihan. *Quaternion Singular Value Decomposition based on Bidiagonalization to a Real or Complex Matrix using Quaternion Householder Transformations*. Applied mathematics and computation, vol. 182, pages 727 – 738, 2006. 12
- [Sangwine 2010] S.J. Sangwine et N. Le Bihan. *Quaternion Polar Representation with a Complex Modulus and Complex Argument Inspired by the Cayley-Dickson Form*. Advances in Applied Clifford Algebras, vol. 20, pages 111–120, 2010. 9, 71
- [Sangwine 2011] S.J. Sangwine, T.A. Ell et N. Le Bihan. *Fundamental representations and algebraic properties of biquaternions or complexified quaternions*. to appear in Advances in Applied Clifford Algebras, 2011. 11
- [Schmidt 1986] R.O. Schmidt. *Multiple emitter location and signal parameter estimation*. IEEE Transactions on Antennas and Propagation, vol. 34, pages 276 – 280, 1986. 65, 66
- [Schreier 2010] P.J. Schreier et L.L. Scharf. *Statistical signal processing of complex-valued data : the theory of improper and noncircular signals*. Cambridge University Press, 2010. 5, 9, 16, 70
- [Segert 1987] J. Segert. *Photon Berry's phase as a classical topological effect*. Physical Review A, vol. 36, no. 1, pages 10–15, 1987. 111, 112
- [Serôdio 2001] R. Serôdio, E. Pereira et J. Vitória. *Computing the zeros of quaternion polynomial*. Computers and mathematics with applications, vol. 42, pages 1229 – 1337, 2001. 12
- [Shoemake 1985] K. Shoemake. *Animating rotations with quaternion curves*. SIGGRAPH, vol. 19, no. 3, pages 245–254, 1985. 6, 9
- [Sommer 2001] G. Sommer. *Geometric computing with clifford algebras*. Springer, 2001. 11, 13, 15
- [Taylor 1986] M.E. Taylor. *Noncommutative harmonic analysis*. American Mathematical Society, 1986. 101

- [Tian 2000] Y. Tian. *Matrix representations of octonions and their applications*. Advances in Applied Clifford Algebras, vol. 10, no. 1, pages 61 – 90, 2000. 156
- [Tomita 1986] A. Tomita et R.Y. Chiao. *Observation of Berry's topological phase by use of an optical fiber*. Physical Review Letters, vol. 57, pages 937 – 940, 1986. 111, 112
- [Took 2009] C. Cheong Took. *Study of the quaternion LMS and four-channel LMS algorithms*. Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on, pages 3109 – 3112, 2009. 16
- [Took 2010] C. Cheong Took. *A Quaternion Widely Linear Adaptive Filter*. Signal Processing, IEEE Transactions on, vol. 58, no. 8, pages 4427 – 4431, 2010. 16
- [Ujang 2009] C. Ujang, C. Took et A. Kavcic. *A split quaternion nonlinear adaptive filter*. Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on, pages 1745 – 1748, 2009. 16
- [Vakhania 1998] N. Vakhania. *Random vectors with values in quaternions Hilbert spaces*. Theory of probability and its applications, vol. 43, no. 1, pages 99–115, 1998. 16
- [Via 2010a] J. Via, D. P. Palomar et L. Vielva. *Generalized Likelihood Ratios for Testing the Properness of Quaternion Gaussian Vectors*. IEEE Transactions on Signal Processing, vol. PP, no. 99, page 1, 2010. 16, 65
- [Via 2010b] J. Via, D. Ramirez et I. Santamaria. *Properness and widely linear processing of quaternion random vectors*. IEEE Transactions on Information Theory, vol. 57, no. 7, pages 3502–3514, 2010. 65
- [Via 2010c] J. Via, L. Vielva, I. Santamaria et D.P. Palomar. *Independent component analysis of quaternion Gaussian vectors*. In Sensor Array and Multichannel Signal Processing Workshop (SAM), 2010 IEEE, pages 145 –148, 2010. 16
- [Via 2011] J. Via, D.P. Palomar, L. Vielva et I. Santamaria. *Quaternion ICA from Second-Order Statistics*. to appear in IEEE Transactions on Signal Processing, 2011. 16
- [Ville 1948] J. Ville. *Théorie et applications de la notion de signal analytique*. Cables et Transmission, vol. 2A, pages 61 – 74, 1948. 5, 70
- [von Bergman 2007] J. von Bergman et H. von Bergman. *Foucault pendulum through basic geometry*. American Journal of Physics, vol. 75, no. 10, pages 888 – 892, 2007. 109, 110
- [Ward 1997] J. P. Ward. Quaternions and cayley numbers : Algebra and applications, volume 403 of *Mathematics and Its Applications*. Kluwer, Dordrecht, 1997. 6, 7, 9, 11
- [Wilczek 1989] F. Wilczek et A. Shapere. *Geometric phases in physics*. World scientific Singapore, 1989. 109, 110
- [Wolf 1936] L.A. Wolf. *Similarity of matrices in which elements are real quaternions*. Bull. Amer. Math. Soc., vol. 42, pages 737 – 743, 1936. 11
- [Wu 2008] J. Wu, L. Zou, X. Chen et S. Li. *The estimation of eigenvalues of sum, difference, and tensor product of matrices over quaternion division algebra*. Linear Algebra and its Applications, vol. 428, no. 11-12, pages 3023 – 3033, 2008. 11
- [Yang 2001] J. Yang, W.L. Kath et C.R. Menyuk. *Polarization mode dispersion probability distribution for arbitray distances*. Optics letters, vol. 26, pages 1472 – 1474, 2001. 101
- [Zhang 1997] F. Zhang. *Quaternions and matrices of quaternions*. Linear algebra and its applications, no. 21, pages 21 – 57, 1997. 11, 12