The development of the quaternion normal distribution

by

Mattheüs Theodor Loots

Submitted in fulfillment of the requirements for the degree
Master of Science (Mathematical Statistics)
in the Faculty of Natural and Agricultural Sciences
University of Pretoria, Pretoria

December 2010

Publication data:

Mattheüs Theodor Loots. The development of the quaternion normal distribution. Master's dissertation, University of Pretoria, Department of Statistics, Pretoria, South Africa, December 2010.

Electronic, hyperlinked versions of this dissertation are available online, as Adobe PDF files, at:

http://upetd.up.ac.za/UPeTD.htm

The development of the quaternion normal distribution

by

Mattheüs Theodor Loots
E-mail: theo.loots@gmail.com

Abstract

In this dissertation an overview on the real representation of quaternions in distribution theory is given. The density functions of the p-variate and matrix-variate quaternion normal distributions are derived from first principles, while that of the quaternion Wishart distribution is derived from the real associated Wishart distribution via the characteristic function. Applications of this theory in hypothesis testing is presented, and the density function of Wilks's statistic is derived for quaternion Wishart matrices.

Keywords: characteristic function, Meijer's G-function, quaternion matrix-variate beta type I distribution, quaternion normal distribution, quaternion Wishart distribution, real representation, Wilks's statistic.

Supervisor : Prof. A. Bekker
Department : Department of Statistics
Degree : Master of Science

I, Mattheüs Theodor Loots, declare that the thesis/dissertation, which I hereby submit for the degree, Master of Science (Mathematical Statistics), at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: ....
DATE: 3 December 2010

The financial assistance of the Department of Labour (DoL) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the DoL.

"Tomorrow will be the fifteenth birthday of the quaternions. They started into life, or light, full grown, on the 16th of October, 1843, as I was walking with Lady Hamilton to Dublin, and came up to Brougham Bridge. That is to say, I then and there felt the galvanic circuit of thought closed, and the sparks which fell from it were the fundamental equations between i, j, k; exactly such as I have used them ever since. I pulled out, on the spot, a pocketbook, which still exists, and made an entry, on which, at the very moment, I felt that it might be worth my while to expand the labour of at least ten (or it might be fifteen) years to come. But then it is fair to say that this was because I felt a problem to have been at that moment solved, an intellectual want relieved, which had haunted me for at least fifteen years before."

Kline (1972) [22] as cited in Rautenbach (1983) [28].

To Katryn, with love "The universe rings true wherever you fairly test it." (C.S. Lewis "Surprised by Joy")

Acknowledgements

Without the support of the following list of contributors, this work wouldn't have been possible at all:

- Thank You God, the Creator of all things, for giving me the words to sing Your song. "Not to us, O LORD, not to us but to your name be the glory, because of your love and faithfulness." (PS 115:1, NIV Holy Bible)
- My wife, Katryn, thank you for always keeping an eye out for me. Without your love and continual support this wouldn't have been a reality. (I'll be on dish duty for December!)
- Thank you family and friends for fantastic relationships. In particular, thank you mom (Corrie) for all your typing work.
- Prof. A. Bekker, you are so much more than just a supervisor, it is truly a joy working with you. Prof. J.J.J. Roux, thank you for the cheers from backstage and the initial inspiration. Prof. M. Arashi, thank you Mohammad for your visit, your e-mails and all the invaluable contributions. Prof. K.T. Fang and Prof. D.K. Nagar, thank you for granting access to your research material, as used in this work.
- Dr. L. Fletcher, thank you, without your initiative I wouldn't have had access to many key articles used in this work. Ms. R. Bron, Rhuhanda, thank you for the hours and hours of typing you did, and your willingness to support this venture. To all staff at the University of Pretoria, thank you all for the investment you made in my life.
- The financial assistance of the South African Department of Labour (DoL) and National Research Foundation (NRF) towards this research is hereby acknowledged. To my current employer, TransUnion Analytic and Decision Services, thank you for your financial support and for granting me study leave, especially towards the completion of this dissertation.

Contents

Chapter 1

Introduction

On October 16, 1843, while walking with his wife past the Broome Bridge, Hamilton made a breakthrough in his quest for extending complex numbers with the concept of a system that contained one real and three imaginary parts. In the following excerpt from a letter to his son, Hamilton describes the moment of inspiration .

". . . it is not too much to say that I felt at once the importance. An electric circuit seemed to close; and a spark flashed forth, the herald of many long years to come of definitely directed thought and work, by myself if spared, and at all events on the part of others, if I should even be allowed to live long enough distinctly to communicate the discovery. Nor could I resist the impulse – unphilosophical as it may have been – to cut with a knife on a stone of Brougham Bridge, as we passed it, the fundamental formula with the symbols, i, j, k; namely, $i^2 = j^2 = k^2 = ijk = -1$ *which contains the Solution to the problem . . ."*

Hamilton's excitement at the discovery prompted him to carve the critical equation into a nearby bridge as insurance against the possibility that he might die before he told someone else of his breakthrough. A plaque is now located at Broome Bridge in Dublin to commemorate the event.

(This quotation along with the below photograph is excerpted from the book by Hanson (2006) [16].)

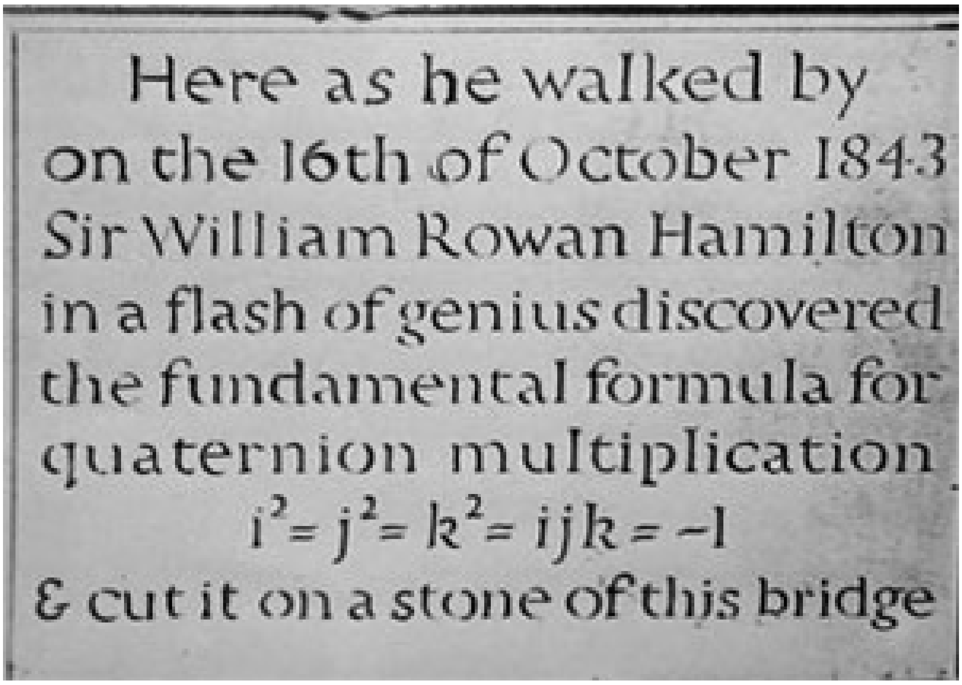

Figure 1.1: Plaque located at Broome Bridge commemorating the discovery of the quaternions by Sir William Rowan Hamilton on October 16, 1843.

1.1 Motivation

Although quaternions had been invented during the first half of the 19^{th} century, they only made their first appearance 132 years later in the statistical literature in an article by Andersson (1975) [1].

Andersson employed an indirect approach in his development of the quaternion normal distribution by imposing conditions on its expected value and covariance. However, both Rautenbach (1983) [28], and Teng and Fang (1997) [31], independently remarked that certain aspects underlying the quaternion distribution theory are lost, or not stated explicitly, when working with these invariant normal models.

Kabe (1976) [17], (1978) [18], (1984) [19], generalised the work done by Goodman (1963) [10], and Khatri (1965) [21], from the complex to the hypercomplex space. For a thorough discussion on the complex distribution theory, see the book by Gupta and Nagar (2009) [14]. Kabe's approach utilised the representation theory, and was further studied by Rautenbach (1983) [28], and more recently by Teng and Fang (1997) [31].

Distribution theory lies at the intersection of probability and statistics and is the foundation from which all statistical theory and application originates. Enhancing our knowledge of distribution theory therefore implies that the general body of knowledge of statistics is improved. Starting from the foundation of the multivariate real normal distribution, it will be endeavoured to show how the representation theory method may be expanded yielding the multivariate quaternion normal, matrix-variate quaternion normal, quaternion chi-squared and quaternion Wishart distributions. In each case, the emphases will be on the relationship between the real and quaternionic space.

1.2 Objectives

- Present a thorough review on the literature on the quaternion distribution theory.
- Specifically focus on the role of the representation theory in quaternion statistics.
- Follow a systematic approach in building up the family of quaternion distributions, starting with the multivariate normal distribution.
- Determine for which normal related quaternion distributions, that already appear in the literature, equivalent derivations may be found, using the representation theory approach.
- Investigate the role of the quaternion normal distribution in hypothesis testing.

1.3 Contributions

- The work of the main contributors to the quaternion distribution theory, utilising the representation theory, is contrasted and presented as a whole.
- The work of Rautenbach (1983) [28] was previously unaccessible to the scholarly community, and of which, some results are now made available in Appendix A.
- For the first time, the matrix-variate quaternion normal and quaternion Wishart distributions are derived from first principles, i.e. from their real counterparts, exposing the relations between their respective density and characteristic functions.

- The role of the quaternion normal distribution in applications is illustrated.

1.4 Dissertation Outline

- In **Chapter 2** a collection of some fundamental mathematical results are given for use in later sections. Probability quantities, such as quaternion probability vectors and matrices, density functions, moments and characteristic functions are also defined.

- **Chapter 3** is first devoted to a review on the derivation of the p-variate quaternion normal distribution using the representation theory, and thereafter, the matrix-variate quaternion normal distribution is derived by generalising this approach.

- The quaternion Wishart distribution, as discussed in **Chapter 4**, is not a new addition to the family of matrix-variate quaternion distributions, however, it will be shown how it relates to its real counterpart. Kabe (1976) [17] derived the quaternion Wishart distribution by extending Sverdrup's lemma to the Q generalized Sverdrup's lemma, while Teng and Fang (1997) [31] validated the results given by Andersson (1975) [1] by deriving the characteristic function of the quaternion Wishart distribution by applying a Fourier transformation. We also note that the quaternion chi-squared distribution reduces to an associated real-valued form.

- **Chapter 5** concludes by showing that a simple quaternion hypothesis may be represented with an associated real hypothesis, and thereafter derive an expression for the density function of Wilks's statistic in the case of quaternion Wishart matrices.

- **Chapter 6** gives some conclusive remarks and a summary on the material covered in this dissertation.

The following Appendices are found towards the end of this dissertation:

- **Appendix A** contains some useful mathematical results regarding the representation theory, complex numbers, quaternions, and algebraic results, including functions and polynomials of quaternions.

- **Appendix B** provides a list of the important acronyms used throughout this work, as well as their associated definitions.

- **Appendix C** lists and defines the notational convensions and mathematical symbols used in this work.

The Index, starting on page 115, contains a list of terms that may be used for reference purposes.

Chapter 2

Quaternion Distribution Theory

A number of useful theorems and other general results that are found in the literature will be discussed in this chapter, which form the basis for further discourse in subsequent chapters.

Some basic mathematical results are presented in Section 2.1 which are required before advancing to a discussion on quaternion probability quantities in Section 2.2. Quaternion probability density functions of quaternion variables, vectors and matrices are the objects under the viewing glass in Section 2.3 for which the respective quaternion moments and characteristic functions are introduced in Section 2.4.

2.1 Mathematical preliminaries

Let $\mathbb{R}$ denote the field of real numbers, and $\mathbb{Q}$ the quaternion (Hamiltonion) division algebra over $\mathbb{R}$, respectively. Hence, every $z \in \mathbb{Q}$ can be expressed as

$$z = x_1 + ix_2 + jx_3 + kx_4,$$

where i, j, and k satisfy the following relations:

$$i^2 = j^2 = k^2 = -1, \qquad ij = -ji = k, \qquad jk = -kj = i, \qquad ki = -ik = j,$$

and where $x_1, x_2, x_3 \in \mathbb{R}$. The conjugate of a quaternion element is defined in a similar fashion to that of a complex number, and is given by:

$$\bar{z} = x_1 - ix_2 - jx_3 - kx_4$$

Now let $M_{n\times p}(\mathbb{R})$ and $M_{n\times p}(\mathbb{Q})$ denote the set of all $n\times p$ matrices over $\mathbb{R}$ and $\mathbb{Q}$, respectively. In the case of square matrices, say $p\times p$, this will be indecated by $M_p(\mathbb{R})$ and $M_p(\mathbb{Q})$ instead. Similar to the scalar form above, any $\mathbf{Z}\in M_{n\times p}(\mathbb{Q})$ may be rewritten as:

$$\mathbf{Z}=[z_{ij}]_{n\times p}=\mathbf{X}_1+i\mathbf{X}_2+j\mathbf{X}_3+k\mathbf{X}_4,$$

where $z_{ij}\in\mathbb{Q}$, and $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{X}_3$, and $\mathbf{X}_4\in M_{n\times p}(\mathbb{R})$. $\mathbf{X}_1$ is the real part of $\mathbf{Z}$, and will be denoted by $\operatorname{Re}\mathbf{Z}$. By setting $n=1$, this reduces to the vector form in an obvious way.

The transpose of a matrix $\underset{n\times p}{\mathbf{Z}}$ will be denoted as $\mathbf{Z}'=\underset{p\times n}{\mathbf{Z}}$. The conjugate transpose of $\mathbf{Z}$ is therefore given by

$$\bar{\mathbf{Z}}'=\left[\bar{z}'_{ij}\right]_{p\times n}=\mathbf{X}'_1-i\mathbf{X}'_2-j\mathbf{X}'_3-k\mathbf{X}'_4.$$

and it is said that $\mathbf{Z}$ is Hermitian if $\bar{\mathbf{Z}}'=\mathbf{Z}$.

The vec operator is frequently used in expressions involving matrices of quaternions, see for instance Li and Xue (2009) [23], and is defined as

$$\text{vec }\mathbf{Z}=\left[\underline{Z}'_1,\ldots,\underline{Z}'_p\right]'\in M_{np\times 1}(\mathbb{Q}),\tag{2.1.1}$$

where $\underline{Z}_\alpha\in M_{n\times 1}(\mathbb{Q}),\alpha=1,\ldots,p$ are the columns of $\mathbf{Z}$.

Throughout this work the representation theory will be used, and although quaternions may be represented by real matrices in various ways, see Teng and Fang (1997) [31], the representation employed by Kabe (1976) [17] and (1984) [19], and Rautenbach (1983) [28] will be preferred. For a detailed exposition on this topic, the reader is referred to Appendix A. Specifically suppose that $z=x_1+ix_2+jx_3+kx_4\in\mathbb{Q}$ may be represented by $\mathbf{z}_0\in M_4(\mathbb{R})$, as

$$\underset{4\times4}{\mathbf{z}_0}=\begin{bmatrix}x_1&-x_2&-x_3&-x_4\\x_2&x_1&-x_4&x_3\\x_3&x_4&x_1&-x_2\\x_4&-x_3&x_2&x_1\end{bmatrix}.$$

Now, if $\underset{n\times p}{\mathbf{Z}}\in M_{n\times p}(\mathbb{Q})$, i.e. the case for matrices with quaternion elements (or vectors by setting $n=1$), it follows that

$$\underset{n\times p}{\mathbf{Z}}=[z_{st}]\,,$$

where $z_{st} = x_{1st} + ix_{2st} + jx_{3st} + kx_{4st} \in \mathbb{Q}, s = 1, \ldots, n,$ and $t = 1, \ldots, p$. By an elementwise generalisation of the representation of the scalar, to the matrix (or vector) form, it follows that

$$\underset{4\times 4}{\mathbf{z}_{0st}} = \begin{bmatrix} x_{1st} & -x_{2st} & -x_{3st} & -x_{4st} \\ x_{2st} & x_{1st} & -x_{4st} & x_{3st} \\ x_{3st} & x_{4st} & x_{1st} & -x_{2st} \\ x_{4st} & -x_{3st} & x_{2st} & x_{1st} \end{bmatrix}.$$

in other words, $\underset{n\times p}{\mathbf{Z}}$ may be represented with the real matrix $\underset{4n\times 4p}{\mathbf{Z}_0}$ as:

$$\underset{4n\times 4p}{\mathbf{Z}_0} = [\mathbf{z}_{0st}].$$

By defining the mapping

$$f\left(\underset{4p\times 4p}{\mathbf{Z}_0}\right) = \underset{p\times p}{\mathbf{Z}} \;\forall \mathbf{Z}_0 \in M_{4p}(\mathbb{R}), \mathbf{Z} \in \mathbb{Q}.$$

it follows from Rautenbach (1983) [28] that f is a faithful representation, as set out in Theorem A.4.7. When $\underset{4p\times 4p}{\mathbf{Z}_0} \in M_{4p}(\mathbb{R})$ and $f(\mathbf{Z}_0) = \mathbf{Z} \in M_p(\mathbb{Q})$ then it will be indicated as $\mathbf{Z}_0 \simeq \mathbf{Z}$.

The trace operator is frequently used in the symplification of expressions, and although the multiplication of quaternions are noncommutative, it may be formulated as follows, see Andersson (1975) [1] and Zhang (1997) [33]. By setting $\mathrm{Retr}(\mathbf{Z}) = \mathrm{tr}(\mathrm{Re}\,\mathbf{Z})$ for $\mathbf{Z} \in M_p(\mathbb{Q})$, it follows that

$$\begin{aligned} \mathrm{Retr}(\mathbf{Z}) &= \frac{1}{2}\,\mathrm{tr}(\mathbf{Z} + \bar{\mathbf{Z}}') \\ \mathrm{Retr}(\mathbf{ZY}) &= \mathrm{Retr}(\mathbf{YZ}) \quad \forall \mathbf{Z}, \mathbf{Y} \in M_p(\mathbb{Q}) \end{aligned} \tag{2.1.2}$$

Moreover, if $\mathbf{Z} = \bar{\mathbf{Z}}' \in M_p(\mathbb{Q})$, i.e. a Hermitian matrix, and using Theorem A.5.11 then this becomes

$$\mathrm{Retr}(\mathbf{Z}) = \mathrm{tr}(\mathbf{Z}) = \sum_{\alpha=1}^{p} \lambda_\alpha,$$

where $\lambda_1, \ldots, \lambda_p$ are the eigenvalues of $\mathbf{Z}$.

The stage is now set for the definition of some concepts specifically pertaining to the development of the quaternion distribution theory.

2.2 Quaternion probability quantities

In this section definitions for a quaternion probability variable, vector and matrix are given, but first some foundational definitions such as random phenomenon, and an event in a sample space, are presented for completeness' sake. For further discussions on these topics, the reader is referred to Rautenbach (1983) [28] and Bain and Engelhardt (1992) [3].

Definition 2.2.1. A random phenomenon is an emperical phenomenon known to yield a different outcome under a fixed set of circumstances, or conditions, whenever it is observed, in such a way as to preserve statistical consistency.

Definition 2.2.2. A sample space, S is a collection of descriptions of all possible observable outcomes of a random phenomenon.

Definition 2.2.3. An event is a subset of the sample space S.

Definition 2.2.4. A quaternion probability variable, $Z(\cdot)$, is a quaternion valued function, which is defined on the elements of the sample space S, in such a way that for every Borel set $B(\mathbb{Q})$ of quaternions, denoted by B, in the range of $Z(\cdot)$, the set $\{s \in S, Z(\cdot) \in B\}$, is an event in S.

It is clear from Definition 2.2.4 that every value z attained by the quaternion probability variable Z is a quaternion, and is therefore of the form

$$z = x_1 + ix_2 + jx_3 + kx_4$$

where x_1 is the real and x_2, x_3 and x_4 are the imaginary components of z respectively.

x_1, x_2, x_3 and x_4 can now be regarded as the observed values of four real probability variables, X_1, X_2, X_3 and X_4, for every possible observable z, respectively. Thus, the quaternion probability variable Z may therefore be thought of as a quaternion linear combination of four real probability variables. These four real probability variables, X_1, X_2, X_3 and X_4 are now inspected more closely, which will lead to the definition of the 4-variate real probability vector $\underset{4\times 1}{\underline{Z}_0} = [X_1, X_2, X_3, X_4]'$.

Definition 2.2.5. Let

$$Z = X_1 + iX_2 + jX_3 + kX_4$$

be a quaternion probability variable.

1. The real probability variable X_1 will be referred to as the real probability component of Z.

2. The real probability variables X_2, X_3 and X_4 will be referred to as the first, second and third imaginary probability components of Z, respectively.

3. The 4-variate real probability vector $\underset{4\times 1}{\underline{Z}_0} = [X_1, X_2, X_3, X_4]'$ will be referred to as the associated real probability vector of Z. This implies that there is a 4-variate real probability vector associated with every quaternion probability variable.

The idea of a quaternion probability variable is now extended to the more general idea of a quaternion probability vector.

Definition 2.2.6. Let $Z_1, \ldots, Z_p$ be p quaternion probability variables. The vector $\underset{p\times 1}{\underline{Z}} = [Z_1, \ldots, Z_p]'$ is called a quaternion probability vector with p variables.

From this definition, it is clear that every possible realisation of the vector $\underset{p\times 1}{\underline{z}}$ is a vector of p quaternions. $\underline{z}$ can now be written in the form

$$\underline{z} = \underline{x}_1 + i\underline{x}_2 + j\underline{x}_3 + k\underline{x}_4$$

where $\underline{x}_1$, $\underline{x}_2$, $\underline{x}_3$ and $\underline{x}_4$ are p-variate quaternion vectors. $\underline{x}_1$, $\underline{x}_2$, $\underline{x}_3$ and $\underline{x}_4$ can once again be viewed as observed values of the real probability vectors $\underline{X}_1$, $\underline{X}_2$, $\underline{X}_3$ and $\underline{X}_4$ respectively. The quaternion probability vector $\underline{Z}$ can therefore be viewed as a quaternion linear combination of four real probability vectors, namely

$$\underline{Z} = \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4.$$

These four real probability vectors $\underline{X}_1$, $\underline{X}_2$, $\underline{X}_3$ and $\underline{X}_4$ are now formally defined, and their resultant $4p$-variate real probability vector

$$\underset{4p\times 1}{\underline{Z}_0} = [\underline{X}_1', \underline{X}_2', \underline{X}_3', \underline{X}_4']'.$$

Definition 2.2.7. Let

$$\underset{p\times 1}{\underline{Z}} = \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4$$

be a quaternion probability vector.

1. The real probability vector $\underset{p\times 1}{\underline{X}_1}$ will be referred to as the real probability component of $\underline{Z}$.

2. The real probability vectors $\underset{p\times 1}{\underline{X}_2}$, $\underset{p\times 1}{\underline{X}_3}$ and $\underset{p\times 1}{\underline{X}_4}$ will be referred to as the first, second and third imaginary probability components of $\underline{Z}$.

3. The $4p$-variate real probability vector

$$\underset{4p\times 1}{\underline{Z}_0} = [\underline{X}_1', \underline{X}_2', \underline{X}_3', \underline{X}_4']'$$

will be referred to as the associated real probability vector of $\underline{Z}$. A $4p$-variate real probability vector is therefore associated with each p-variate quaternion probability vector.

2.3 Quaternion probability density functions

The probability density function (pdf) of a quaternion probability quantity is defined to be algebraically equal to the real pdf of the corresponding associated probability quantity. Probabilities in the case of quaternion probability variables and probability vectors are therefore calculated by integrating the real pdf over the subspaces of $\mathbb{R}^4$ and $\mathbb{R}^{4p}$ respectively. Akin to the complex case, calculations of probabilities are carried out within the realm of multivariate distribution theory.

Definition 2.3.1. Let

$$Z = X_1 + iX_2 + jX_3 + kX_4$$

be a quaternion probability variable with real associated probability vector

$$\underset{4\times 1}{\underline{Z}_0} = [X_1, X_2, X_3, X_4]'.$$

A function $f_Z(z)$ of the quaternion variable $z = x_1 + ix_2 + jx_3 + kx_4$ is called the pdf of Z if

$$f_Z(z) = f_{\underline{Z}_0}(\underline{z}_0)$$

where $f_{\underline{Z}_0}(\underline{z}_0)$ is the 4-variate pdf of $\underline{Z}_0$.

Remark 2.3.2. 1. It is clear that the pdf of a quaternion probability variable a real valued function of a quaternion variable is. The pdf of $f_Z(z)$ therefore yields the probability associated with a quaternion probability variable Z. Suppose that Z varies over the quaternion region B with positive probability and let G_1 be any subset of B. Then it follows that

$$P\left[Z \in G_1\right] = \int\limits_{G_1} f_Z(z)dz.$$

Every probabilistic statement made regarding Z is in actual fact equivalent to that made about $\underline{Z}_0$, where the associated region for G_1 are given by

$$G_{01} \subset \mathbb{R}^4,$$

thus

$$P\left[Z \in G_1\right] = P\left[\underline{Z}_0 \in G_{01}\right]$$

which can be written as

$$\int\limits_{G_1} f_Z(z)dz = \int\limits_{G_{01}} f_{\underline{Z}_0}(\underline{z}_0)d\underline{z}_0. \tag{2.3.1}$$

2. From the definition above it seems that the quaternion pdf has similar properties as that of the real pdf, namely:

(a) $\int\limits_B f_Z(z)dz = 1$, and

(b) $f_Z(z) \geq 0 \forall z$.

The pdf of a quaternion probability vector is now defined in a similar fashion to that of a quaternion probability variable.

Definition 2.3.3. Let

$$\underset{p\times 1}{\underline{Z}} = \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4$$

be a quaternion probability vector with real associated probability vector given by

$$\underset{4p\times 1}{\underline{Z}_0} = [\underline{X}_1', \underline{X}_2', \underline{X}_3', \underline{X}_4']'.$$

A function $f_{\underline{Z}}(\underline{z})$ of the p quaternion variables $z_s = x_{1s} + ix_{2s} + jx_{3s} + kx_{4s},\ s = 1, \ldots, p$ is called the probability density function (pdf) of $\underline{Z}$ if

$$f_{\underline{Z}}(\underline{z}) = f_{\underline{Z}_0}(\underline{z}_0)$$

where $f_{\underline{Z}_0}(\underline{z}_0)$ is the $4p$-variate real pdf of $\underset{4p\times 1}{\underline{Z}_0}$.

All the remarks made in Remark 2.3.2 on $f_Z(z)$ also hold in the case for $f_{\underline{Z}}(\underline{z})$. Furthermore, it is also true that

$$\int_{G_2} f_{\underline{Z}}(\underline{z})d\underline{z} = \int_{G_{02}} f_{\underline{Z}_0}(\underline{z}_0)d\underline{z}_0, \tag{2.3.2}$$

where G_2 is a subspace of some quaternion space and G_{02} is the associated subset of $\mathbb{R}^{4p}$.

Note, here it is required that the associated counterpart of $\underset{p\times 1}{\underline{Z}}$ is given by $\underset{4p\times 1}{\underline{Z}_0}$ and not necessarily by $\underset{4p\times 4}{\mathbf{Z}_0}$ as defined earlier. This is due to the derivation of the quaternion normal distribution, which is discussed in Chapter 3.

The question now arises whether matrix-variate quaternion distributions can be defined. It will be shown in Chapter 3 that the problem of finding the pdf of a quaternion random matrix $\underset{n\times p}{\mathbf{Z}}$ reduces to a multivariate problem by using the vec operator as defined in (2.1.1). Specifically, it will be shown that the pdf of $\mathbf{Z}$ is algebraically equivalent to that of a $4pn$-variate real variable $\underset{4pn\times 1}{\mathrm{vec}\,\mathbf{Z}_0}$, where $\mathbf{Z}_0$ is the real associated probability matrix of $\mathbf{Z}$. For this reason, all definitions in the current chapter dealing with quaternion probability vectors apply equally well to quaternion probability matrices.

2.4 Quaternion moments and characteristic functions

In this section the moments of quaternion probability variables and vectors are defined. Their characteristic functions, for which the forms relating to specific distributions will be derived in the chapters that follow, are also derived.

2.4.1 Moments

Suppose

$$Z = X_1 + iX_2 + jX_3 + kX_4$$

is a quaternion probability variable with pdf $f_Z(z)$ and that Z varies over the quaternion space B with positive probability. Let

$$\underset{4\times 1}{\underline{Z}_0} = [X_1, X_2, X_3, X_4]'$$

be its associated real probability vector with pdf $f_{\underline{Z}_0}(\underline{z}_0)$. Let $g(z)$ be any real or quaternion valued function of the quaternion variable $z = x_1 + ix_2 + jx_3 + kx_4$ and suppose that

$$g(z) = g_1(x_1, x_2, x_3, x_4) + ig_2(x_1, x_2, x_3, x_4) + jg_3(x_1, x_2, x_3, x_4) + kg_4(x_1, x_2, x_3, x_4)$$

where $g_1(x_1, x_2, x_3, x_4)$, $g_2(x_1, x_2, x_3, x_4)$, $g_3(x_1, x_2, x_3, x_4)$ and $g_4(x_1, x_2, x_3, x_4)$ are real functions of the real variables x_1, x_2, x_3 and x_4 respectively. From (2.3.1) it follows that

$$\begin{aligned}\int\limits_B g(z) f_Z(z) dz = &\int\limits_{B_0} g_1(x_1, x_2, x_3, x_4) f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \\ &+ i\int\limits_{B_0} g_2(x_1, x_2, x_3, x_4) f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \\ &+ j\int\limits_{B_0} g_3(x_1, x_2, x_3, x_4) f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \\ &+ k\int\limits_{B_0} g_4(x_1, x_2, x_3, x_4) f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \qquad (2.4.1)\end{aligned}$$

where the integrals on the right are integrated with respect to the variables x_1, x_2, x_3 and x_4 and over the region $B_0 \subset \mathbb{R}^4$.

Similar results are obtained for functions of a quaternion vector $\underset{p\times 1}{\underline{z}} = \underline{x}_1 + i\underline{x}_2 + j\underline{x}_3 + k\underline{x}_4$. If

$$g(\underline{z}) = g_1(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4) + ig_2(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4) + jg_3(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4) + kg_4(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4)$$

be a real or quaternion valued function of $\underset{p\times 1}{\underline{z}} = \underline{x}_1 + i\underline{x}_2 + j\underline{x}_3 + k\underline{x}_4$ with $g_1(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4)$, $g_2(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4)$, $g_3(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4)$ and $g_4(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4)$ real valued functions of $\underline{x}_1$, $\underline{x}_2$, $\underline{x}_3$ and $\underline{x}_4$ respectively. It follows from (2.3.2) that

$$\begin{aligned}\int\limits_B g(\underline{z}) f_{\underline{Z}}(\underline{z}) d\underline{z} = &\int\limits_{B_0} g_1(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4) f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \\ &+ i \int\limits_{B_0} g_2(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4) f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \\ &+ j \int\limits_{B_0} g_3(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4) f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \\ &+ k \int\limits_{B_0} g_4(\underline{x}_1, \underline{x}_2, \underline{x}_3, \underline{x}_4) f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \end{aligned} \tag{2.4.2}$$

where the integrals on the right are integrated with respect to the variables $\underline{x}_1$, $\underline{x}_2$, $\underline{x}_3$ and $\underline{x}_4$ and over the region $B_0 \subset \mathbb{R}^{4p}$.

Definition 2.4.1. Suppose that $Z = X_1 + iX_2 + jX_3 + kX_4$ is a quaternion probability variable, with pdf $f_Z(z)$, which varies over the quaternion region B and let $g(Z)$ be a real or quaternion valued function of Z. The expected value of $g(Z)$ is defined as

$$E\left[g(Z)\right] = \int\limits_B g(z) f_Z(z) dz.$$

From (2.4.1) this expected value can now be calculated.

Remark 2.4.2. 1. From Definition 2.4.1 the following special cases follow:

(a) The expected value of Z is given by:

$$\begin{aligned}\mu \equiv &E\left[Z\right] \\ = &\int\limits_B z f_Z(z) dz \\ = &\int\limits_{B_0} x_1 f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 + i \int\limits_{B_0} x_2 f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 + j \int\limits_{B_0} x_3 f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 + k \int\limits_{B_0} x_4 f_{\underline{Z}_0}(\underline{z}_0) d\underline{z}_0 \\ \therefore \mu \equiv &E\left[Z\right] \\ = &E\left[X_1\right] + iE\left[X_2\right] + jE\left[X_3\right] + kE\left[X_4\right]\end{aligned}$$

(b) Let $g(Z) = [Z - E[Z]] \left[\overline{Z - E[Z]}\right]$ then

$$\begin{aligned}\sigma^2 \equiv& \operatorname{var}(Z)\\ =& E[X_1 - E[X_1]]^2 + E[X_2 - E[X_2]]^2 + E[X_3 - E[X_3]]^2 + E[X_4 - E[X_4]]^2\\ =& \operatorname{var}(X_1) + \operatorname{var}(X_2) + \operatorname{var}(X_3) + \operatorname{var}(X_4). \end{aligned} \tag{2.4.3}$$

2. It is therefore clear that $\sigma^2 \equiv \operatorname{var}(Z)$ is a measure of the joint dispersion of the components X_1, X_2, X_3 and X_4. Note however that

$$\operatorname{var}(Z) = \operatorname{var}(X_1 + \operatorname{var}(X_2) + \operatorname{var}(X_3) + \operatorname{var}(X_4)$$

should not be attributed to independence among the respective components, but rather to the special way in which the variance is defined in (2.4.3). In Chapter 3, however, quaternion normal variables are discussed, which are characterised by independence of the components X_1, X_2, X_3 and X_4, such that

$$\begin{aligned}\operatorname{cov}(X_1, X_2) =&0, \qquad \operatorname{cov}(X_1, X_3) =0, \qquad \operatorname{cov}(X_1, X_4) =0\\ \operatorname{cov}(X_2, X_3) =&0, \qquad \operatorname{cov}(X_2, X_4) =0, \qquad \operatorname{cov}(X_3, X_4) =0.\end{aligned}$$

for these probability variables. Hence, in this case it follows that

$$\begin{aligned}\operatorname{var}(Z) =& \operatorname{var}(X_1 + iX_2 + jX_3 + kX_4)\\ =& \operatorname{var}(X_1) + \operatorname{var}(X_2) + \operatorname{var}(X_3) + \operatorname{var}(X_4)\\ =& \operatorname{var}(X_1 + X_2 + X_3 + X_4).\end{aligned}$$

The reason for this special definition of the variance will be discussed in the following chapter.

The covariance between two quaternion probability variables is now defined.

Definition 2.4.3. Let $Z_1 = X_{11} + iX_{12} + jX_{13} + kX_{14}$ and $Z_2 = X_{21} + iX_{22} + jX_{23} + kX_{24}$ be two quaternion probability variables. The covariance between Z_1 and Z_2 are defined as

$$\operatorname{cov}(Z_1, Z_2) = E[Z_1 - E[Z_1]] \left[\overline{Z_2 - E[Z_2]}\right] \tag{2.4.4}$$

where the expected value is calculated using (2.4.2) for the case $p = 2$.

From Definition 2.4.3 it is clear that the covariance between Z_1 and Z_2 is a quaternion quantity, namely

$$\begin{aligned}
&\operatorname{cov}(Z_1, Z_2)\\
=&\operatorname{cov}(X_{11}, X_{21}) + \operatorname{cov}(X_{12}, X_{22}) + \operatorname{cov}(X_{13}, X_{23}) + \operatorname{cov}(X_{14}, X_{24})\\
&+ i\left\{\operatorname{cov}(X_{12}, X_{21}) - \operatorname{cov}(X_{11}, X_{22}) - \operatorname{cov}(X_{13}, X_{24}) + \operatorname{cov}(X_{14}, X_{23})\right\}\\
&+ j\left\{\operatorname{cov}(X_{13}, X_{21}) - \operatorname{cov}(X_{11}, X_{23}) + \operatorname{cov}(X_{12}, X_{24}) - \operatorname{cov}(X_{14}, X_{22})\right\}\\
&+ k\left\{\operatorname{cov}(X_{14}, X_{21}) - \operatorname{cov}(X_{11}, X_{24}) - \operatorname{cov}(X_{12}, X_{23}) + \operatorname{cov}(X_{13}, X_{22})\right\}. \qquad (2.4.5)
\end{aligned}$$

Furthermore, it follows that

$$\operatorname{cov}(Z_1, Z_2) = \operatorname{cov}(\overline{Z_2, Z_1}). \qquad (2.4.6)$$

Definition 2.4.4. Suppose that

$$\begin{aligned}
\underset{p\times 1}{\underline{Z}} &= \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4\\
&= [Z_1, \ldots, Z_p]'
\end{aligned}$$

is a quaternion probability vector.

1. The expected value of $\underline{Z}$ is the p component quaternion vector

$$\underset{p\times 1}{\underline{\mu}} = \left[E\left[Z_1\right], \ldots, E\left[Z_p\right]\right]' \qquad (2.4.7)$$

with components equal to the expected values of $Z_1, \ldots, Z_p$ respectively.

2. The covariance matrix of $\underline{Z}$ is the quaternion Hermitian matrix

$$\underset{p\times p}{\boldsymbol{\Sigma}} = \begin{bmatrix} \operatorname{var}(Z_1) & \operatorname{cov}(Z_1, Z_2) & \ldots & \operatorname{cov}(Z_1, Z_p)\\ \operatorname{cov}(Z_2, Z_1) & \operatorname{var}(Z_2) & \ldots & \operatorname{cov}(Z_2, Z_p)\\ \vdots & \vdots & \ldots & \vdots\\ \operatorname{cov}(Z_p, Z_1) & \operatorname{cov}(Z_p, Z_2) & \ldots & \operatorname{var}(Z_p) \end{bmatrix}. \qquad (2.4.8)$$

Remark 2.4.5. 1. From (2.4.7) it is clear that

$$\underset{p\times 1}{\underline{\mu}} = \begin{bmatrix} E\left[X_{11} + iX_{21} + jX_{31} + kX_{41}\right] \\ \vdots \\ E\left[X_{1p} + iX_{2p} + jX_{3p} + kX_{4p}\right] \end{bmatrix}$$
$$= \begin{bmatrix} E\left[X_{11}\right] \\ \vdots \\ E\left[X_{1p}\right] \end{bmatrix} + i \begin{bmatrix} E\left[X_{21}\right] \\ \vdots \\ E\left[X_{2p}\right] \end{bmatrix} + j \begin{bmatrix} E\left[X_{31}\right] \\ \vdots \\ E\left[X_{3p}\right] \end{bmatrix} + k \begin{bmatrix} E\left[X_{41}\right] \\ \vdots \\ E\left[X_{4p}\right] \end{bmatrix}$$
$$= \underline{\mu}_{X_1} + i\underline{\mu}_{X_2} + j\underline{\mu}_{X_3} + k\underline{\mu}_{X_4}$$

with $\underset{p\times 1}{\underline{\mu}_{X_1}}$, $\underset{p\times 1}{\underline{\mu}_{X_2}}$, $\underset{p\times 1}{\underline{\mu}_{X_3}}$ and $\underset{p\times 1}{\underline{\mu}_{X_4}}$ the average vectors of $\underline{X}_1$, $\underline{X}_2$, $\underline{X}_3$ and $\underline{X}_4$ respectively.

2. From (2.4.8) it is clear that

$$\underset{p\times p}{\mathbf{\Sigma}} = E\left[\underline{Z} - \underline{\mu}\right]\left[\overline{\underline{Z} - \underline{\mu}}\right]'.$$

2.4.2 Characteristic functions

The characteristic functions of quaternion probability variables, vectors and matrices are the object of the discussion in this subsection.

Definition 2.4.6. Suppose that $Z = X_1 + iX_2 + jX_3 + kX_4$ is a quaternion probability variable. The characteristic function (cf) of Z is defined as

$$\phi_Z(z) = E\left[\exp \frac{1}{2}\iota(\bar{Z}t + \bar{t}Z)\right] \tag{2.4.9}$$

where $t = t_1 + it_2 + jt_3 + kt_4$ is a quaternion number and ι the usual imaginary complex root.

Remark 2.4.7. From (2.4.9) it follows that

$$\phi_Z(t) = E\left[\exp \iota(X_1t_1 + X_2t_2 + X_3t_3 + X_4t_4)\right]$$
$$= \phi_{\underline{Z}_0}(\underline{t}_0)$$

where $\phi_{\underline{Z}_0}(\underline{t}_0)$ the cf of the associated real probability variable, $\underset{4\times 1}{\underline{Z}_0} = [X_1, X_2, X_3, X_4]'$ and and further is $\underset{4\times 1}{\underline{t}_0} = [t_1, t_2, t_3, t_4]' \in M_{4\times 1}(\mathbb{R})$. It is therefore clear that the cf of a quaternion probability variable is equivalent to the cf of a 4-variate real probability vector.

Definition 2.4.8. Let $\underset{p\times 1}{\underline{Z}} = \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4$ be a quaternion probability vector. The cf of $\underline{Z}$ is defined as

$$\phi_{\underline{Z}}(\underline{t}) = E\left[\exp\frac{1}{2}\iota\left(\underline{\bar{Z}}'\underline{t} + \underline{\bar{t}}'\underline{Z}\right)\right], \tag{2.4.10}$$

where $\underset{p\times 1}{\underline{t}} = \underline{t}_1 + i\underline{t}_2 + j\underline{t}_3 + k\underline{t}_4$ is a quaternion vector and ι is the usual imaginary complex root.

Remark 2.4.9. 1. In a similar fashion as in the univariate case, it follows that

$$\phi_{\underline{Z}}(\underline{t}) = E\left[\exp\iota\left(\underline{X}_1'\underline{t}_1 + \underline{X}_2'\underline{t}_2 + \underline{X}_3'\underline{t}_3 + \underline{X}_4'\underline{t}_4\right)\right]$$

such that

$$\phi_{\underline{Z}}(\underline{t}) = \phi_{\underline{Z}_0}(\underline{t}_0).$$

$\phi_{\underline{Z}_0}(\underline{t}_0)$ is the cf of the associated real probability vector, $\underset{4p\times 1}{\underline{Z}_0} = [\underline{X}_1', \underline{X}_2', \underline{X}_3', \underline{X}_4']'$ and further is $\underset{4p\times 1}{\underline{t}_0} = [\underline{t}_1', \underline{t}_2', \underline{t}_3', \underline{t}_4']' \in M_{4p\times 1}(\mathbb{Q})$. It is therefore clear that the cf of a quaternion probability vector is equivalent to the cf of a $4p$-variate real probability vector.

2. It should be noted that the cf of the real probability component of a quaternion probability quantity can be derived by using the particular cf in the quaternion case by setting the imaginary components of $\underline{t} \in M_{p\times 1}(\mathbb{Q})$ to zero. From Remark 2.4.7 it follows that $\phi_Z(t) = \phi_{X_1}(t_1)$ when $t_2 = 0$, $t_3 = 0$ and $t_4 = 0$. Similarly, from 1 above, it follows that $\phi_{\underline{Z}}(\underline{t})$ reduces to $\phi_{\underline{X}_1}(\underline{t}_1)$ when $\underline{t}_2 = \underline{0}$, $\underline{t}_3 = \underline{0}$ and $\underline{t}_4 = \underline{0}$.

Definition 2.4.10. Let $\underset{n\times p}{\mathbf{Z}} = \mathbf{X}_1 + i\mathbf{X}_2 + j\mathbf{X}_3 + k\mathbf{X}_4$ be a quaternion probability matrix. The cf of $\mathbf{Z}$ is defined as

$$\phi_{\mathbf{Z}}(\mathbf{T}) = E\left[\exp\frac{\iota}{2}\operatorname{tr}\left(\bar{\mathbf{Z}}'\mathbf{T} + \bar{\mathbf{T}}'\mathbf{Z}\right)\right], \tag{2.4.11}$$

where $\underset{n\times p}{\mathbf{T}} = \mathbf{T}_1 + i\mathbf{T}_2 + j\mathbf{T}_3 + k\mathbf{T}_4$ is a quaternion matrix and ι is the usual imaginary complex root.

Now, if $\mathbf{V} = \bar{\mathbf{Z}}'\mathbf{T} + \bar{\mathbf{T}}'\mathbf{Z}$, say, then it follows that $\bar{\mathbf{V}}' = \bar{\mathbf{T}}'\mathbf{Z} + \bar{\mathbf{Z}}'\mathbf{T} = \mathbf{V}$ implying that $\mathbf{V}$ is Hermitian and hence, from (2.1.2), the cf may be written as

$$\begin{aligned}\phi_{\mathbf{Z}}(\mathbf{T}) =& E\left[\exp\frac{\iota}{2}\operatorname{Retr}\left(\bar{\mathbf{Z}}'\mathbf{T} + \bar{\mathbf{T}}'\mathbf{Z}\right)\right]\\ =& E\left[\exp\frac{\iota}{2}\left[\operatorname{Retr}\left(\bar{\mathbf{Z}}'\mathbf{T}\right) + \operatorname{Retr}\left(\bar{\mathbf{T}}'\mathbf{Z}\right)\right]\right]\\ =& E\left[\exp\frac{\iota}{2}\left[\operatorname{Retr}\left(\bar{\mathbf{Z}}'\mathbf{T}\right) + \operatorname{Retr}\left(\bar{\mathbf{Z}}'\mathbf{T}\right)\right]\right]\\ & \text{since the conjugate do not influence the real components}\\ =& E\left[\exp\iota\operatorname{Retr}\left(\bar{\mathbf{Z}}'\mathbf{T}\right)\right] \qquad (2.4.12)\end{aligned}$$

An expression for the cf of $\underset{p\times 1}{\underline{Z}}$ in terms of Retr can be derived in a similar fashion as in (2.4.12) and is given by

$$\phi_{\underline{Z}}(\underline{t}) = E\left[\exp\iota\operatorname{Retr}\left(\bar{\underline{Z}}'\underline{t}\right)\right].$$

The nature of the problem will determine the form of the cf used in its derivation.

It now follows that

$$\phi_{\mathbf{Z}}(\mathbf{T}) = E\left[\exp\iota\operatorname{Retr}\left(\mathbf{X}_1'\mathbf{T}_1 + \mathbf{X}_2'\mathbf{T}_2 + \mathbf{X}_3'\mathbf{T}_3 + \mathbf{X}_4'\mathbf{T}_4\right)\right]$$

such that

$$\phi_{\mathbf{Z}}(\mathbf{T}) = \phi_{\mathbf{Z}_0}(\mathbf{T}_0). \qquad (2.4.13)$$

$\phi_{\mathbf{Z}_0}(\mathbf{T}_0)$ is the cf of the associated real probability matrix, $\underset{n\times 4p}{\mathbf{Z}_0} = [\mathbf{X}_1', \mathbf{X}_2', \mathbf{X}_3', \mathbf{X}_4']$ and further is $\underset{n\times 4p}{\mathbf{T}_0} = [\mathbf{T}_1', \mathbf{T}_2', \mathbf{T}_3', \mathbf{T}_4']$ a real matrix. It is therefore clear that the cf of a quaternion probability matrix is equivalent to the cf of a $n \times 4p$-variate real probability matrix.

2.5 Summary

The basic foundations for a discussion on quaternion distribution theory have now been laid. In Section 2.2 the idea of random phenomenon, an event in a sample space were defined, along with quaternion probability variables and vectors. In Section 2.3 the pdf of a quaternion probability quantity was defined to be algebraically equal to the pdf of the real associated probability quantity. Similar algebraic equivalence results were derived for moments and characteristic functions, together with their appropriate definitions, in Section 2.4.

The quaternion normal distribution, which form the basis of a further investigation of the quaternion distribution theory, is the topic under discussion in the following chapter.

Chapter 3

The Quaternion Normal Distribution

The p-variate quaternion normal distribution forms the basis from which the quaternion distribution theory is further developed, and will be discussed in Section 3.1. The univariate and bivariate quaternion normal distributions are presented as special cases of the p-variate quaternion normal distribution. In Section 3.2 the matrix-variate quaternion normal distribution is derived using the real representation thereof. For each of these cases the probability density functions as well as their corresponding characteristic functions are derived, with special emphasis on the relationship between the quaternion and associated real cases.

3.1 The p-variate quaternion normal distribution

In this section the approach of Kabe (1976) [17], (1978) [18], (1984) [19] and Rautenbach (1983) [28] will be followed in deriving the p-variate quaternion normal distribution. Although the results in this section are in general not new, it is shown how they relate to those given by Teng and Fang (1997) [31], and with particular emphasis on the quaternion and related real characteristic functions. It should be noted that Andersson (1975) [1] first presented these results by using techniques from group theory, but these are however beyond the scope of the current discussion.

Definition 3.1.1. Let

$$\underset{p\times 1}{\underline{Z}} = [Z_1, \ldots, Z_p]' = \begin{bmatrix} X_{11} + iX_{21} + jX_{31} + kX_{41} \\ X_{12} + iX_{22} + jX_{32} + kX_{42} \\ \vdots \\ X_{1p} + iX_{2p} + jX_{3p} + kX_{4p} \end{bmatrix}$$

be a quaternion probability vector with real associated probability vector

$$\begin{aligned} \underset{4p\times 1}{\underline{Z}_0} &= [X_{11}, \ldots, X_{1p}, X_{21}, \ldots, X_{2p}, X_{31}, \ldots, X_{3p}, X_{41}, \ldots, X_{4p}]' \\ &= \left[\underset{1\times p}{\underline{X}'_1}, \underset{1\times p}{\underline{X}'_2}, \underset{1\times p}{\underline{X}'_3}, \underset{1\times p}{\underline{X}'_4}\right]'. \end{aligned}$$

Then, $\underline{Z}$ has a p-variate quaternion normal distribution if $\underline{Z}_0$ has a $4p$-variate real normal distribution.

The idea of describing the properties of a quaternion probability vector in terms of the properties of its real associated probability vector, as discussed in Chapter 2, was employed in the above definition. Teng and Fang (1997) [31] used a different matrix structure for representing quaternions by matrices. They supposed that $\underset{p\times 1}{\underline{Z}} = \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4$ may be represented by

$$\begin{aligned} \underset{4p\times 4}{\mathbf{Z}_{00}} &= \begin{bmatrix} \underset{p\times 1}{\underline{X}_1} & \underset{p\times 1}{\underline{X}_2} & \underset{p\times 1}{\underline{X}_3} & \underset{p\times 1}{\underline{X}_4} \\ -\underset{p\times 1}{\underline{X}_2} & \underset{p\times 1}{\underline{X}_1} & -\underset{p\times 1}{\underline{X}_4} & \underset{p\times 1}{\underline{X}_3} \\ -\underset{p\times 1}{\underline{X}_3} & \underset{p\times 1}{\underline{X}_4} & \underset{p\times 1}{\underline{X}_1} & -\underset{p\times 1}{\underline{X}_2} \\ -\underset{p\times 1}{\underline{X}_4} & -\underset{p\times 1}{\underline{X}_3} & \underset{p\times 1}{\underline{X}_2} & \underset{p\times 1}{\underline{X}_1} \end{bmatrix}. \\ &= \left[\underset{4p\times 1}{\underline{Y}_1}, \underset{4p\times 1}{\underline{Y}_2}, \underset{4p\times 1}{\underline{Y}_3}, \underset{4p\times 1}{\underline{Y}_4}\right]. \end{aligned}$$

Thus the conjugate $\underset{p\times 1}{\bar{\underline{Z}}} = \underline{X}_1 - i\underline{X}_2 - j\underline{X}_3 - k\underline{X}_4$ of $\underset{p\times 1}{\underline{Z}}$ may be represented as

$$\underset{4p\times 4}{\bar{\mathbf{Z}}_{00}} = \begin{bmatrix} \underset{p\times 1}{\underline{X}_1} & -\underset{p\times 1}{\underline{X}_2} & -\underset{p\times 1}{\underline{X}_3} & -\underset{p\times 1}{\underline{X}_4} \\ \underset{p\times 1}{\underline{X}_2} & \underset{p\times 1}{\underline{X}_1} & \underset{p\times 1}{\underline{X}_4} & -\underset{p\times 1}{\underline{X}_3} \\ \underset{p\times 1}{\underline{X}_3} & -\underset{p\times 1}{\underline{X}_4} & \underset{p\times 1}{\underline{X}_1} & \underset{p\times 1}{\underline{X}_2} \\ \underset{p\times 1}{\underline{X}_4} & \underset{p\times 1}{\underline{X}_3} & -\underset{p\times 1}{\underline{X}_2} & \underset{p\times 1}{\underline{X}_1} \end{bmatrix}.$$
$$= \left[\underset{4p\times 1}{\underline{Y}_1^*}, \underset{4p\times 1}{\underline{Y}_2^*}, \underset{4p\times 1}{\underline{Y}_3^*}, \underset{4p\times 1}{\underline{Y}_4^*}\right].$$

From this it is clear that $\underline{Z}_0 = \underline{Y}_1^*$. Teng and Fang (1997) [31] showed that any of the $\underline{Y}_s$, $s = 1, 2, 3, 4$ may be used to arive at the same form of the probability density function for the p-variate quaternion normal distribution.

Kabe (1976) [17] investigated the case where $\underline{Z}_0$ has a very special $4p$-variate real normal distribution, namely:

$$\begin{aligned} \underline{\mu}_0 \equiv& E\left[\underline{Z}_0\right] \\ =& \left[\underline{\mu}'_{X_1}, \underline{\mu}'_{X_2}, \underline{\mu}'_{X_3}, \underline{\mu}'_{X_4}\right]' \\ =& \left[E\left[\underline{X}'_1\right], E\left[\underline{X}'_2\right], E\left[\underline{X}'_3\right], E\left[\underline{X}'_4\right]\right]' \end{aligned}$$

and

$$\begin{aligned} \underset{4p\times 4p}{\mathbf{\Sigma}_0} =& \operatorname{cov}\left(\underline{Z}_0, \underline{Z}'_0\right) \\ =& E\left[\left(\underline{Z}_0 - \underline{\mu}_0\right)\left(\underline{Z}_0 - \underline{\mu}_0\right)'\right] \\ =& \frac{1}{4}\begin{bmatrix} \mathbf{\Sigma}_1 & -\mathbf{\Sigma}_2 & -\mathbf{\Sigma}_3 & -\mathbf{\Sigma}_4 \\ \mathbf{\Sigma}_2 & \mathbf{\Sigma}_1 & -\mathbf{\Sigma}_4 & \mathbf{\Sigma}_3 \\ \mathbf{\Sigma}_3 & \mathbf{\Sigma}_4 & \mathbf{\Sigma}_1 & -\mathbf{\Sigma}_2 \\ \mathbf{\Sigma}_4 & -\mathbf{\Sigma}_3 & \mathbf{\Sigma}_2 & \mathbf{\Sigma}_1 \end{bmatrix} \end{aligned} \tag{3.1.1}$$

where $\underset{p\times p}{\mathbf{\Sigma}_1}$ is a real symmetric matrix, and $\underset{p\times p}{\mathbf{\Sigma}_2}$, $\underset{p\times p}{\mathbf{\Sigma}_3}$ and $\underset{p\times p}{\mathbf{\Sigma}_4}$ are real skew symmetric matrices. This covariance structure of $\underline{Z}_0$ is the special property of the p-variate quaternion normal distribution and implies that:

$$\underset{p\times p}{\mathbf{\Sigma}_1} = \begin{bmatrix} \sigma_1^2 & \eta_{12}\sigma_1\sigma_2 & \dots & \eta_{1p}\sigma_1\sigma_p \\ \eta_{12}\sigma_1\sigma_2 & \sigma_2^2 & \dots & \eta_{2p}\sigma_2\sigma_p \\ \vdots & \vdots & \dots & \vdots \\ \eta_{1p}\sigma_1\sigma_p & \eta_{2p}\sigma_2\sigma_p & \dots & \sigma_p^2 \end{bmatrix} \text{ where} \quad (3.1.2)$$

$$\eta_{st} = \frac{\sigma_{1st}}{\sigma_s\sigma_t}, s,t = 1,\dots,p, s<t$$

$$\underset{p\times p}{\mathbf{\Sigma}_2} = \begin{bmatrix} 0 & \alpha_{12}\sigma_1\sigma_2 & \alpha_{13}\sigma_1\sigma_3 & \dots & \alpha_{1p}\sigma_1\sigma_p \\ -\alpha_{12}\sigma_1\sigma_2 & 0 & \alpha_{23}\sigma_2\sigma_3 & \dots & \alpha_{2p}\sigma_2\sigma_p \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\alpha_{1p}\sigma_1\sigma_p & -\alpha_{2p}\sigma_2\sigma_p & -\alpha_{3p}\sigma_3\sigma_p & \dots & 0 \end{bmatrix} \text{ where} \quad (3.1.3)$$

$$\alpha_{st} = \frac{\sigma_{2st}}{\sigma_s\sigma_t}, s,t = 1,\dots,p, s<t$$

$$\underset{p\times p}{\mathbf{\Sigma}_3} = \begin{bmatrix} 0 & \beta_{12}\sigma_1\sigma_2 & \beta_{13}\sigma_1\sigma_3 & \dots & \beta_{1p}\sigma_1\sigma_p \\ -\beta_{12}\sigma_1\sigma_2 & 0 & \beta_{23}\sigma_2\sigma_3 & \dots & \beta_{2p}\sigma_2\sigma_p \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\beta_{1p}\sigma_1\sigma_p & -\beta_{2p}\sigma_2\sigma_p & -\beta_{3p}\sigma_3\sigma_p & \dots & 0 \end{bmatrix} \text{ where} \quad (3.1.4)$$

$$\beta_{st} = \frac{\sigma_{3st}}{\sigma_s\sigma_t}, s,t = 1,\dots,p, s<t$$

$$\underset{p\times p}{\mathbf{\Sigma}_4} = \begin{bmatrix} 0 & \lambda_{12}\sigma_1\sigma_2 & \lambda_{13}\sigma_1\sigma_3 & \dots & \lambda_{1p}\sigma_1\sigma_p \\ -\lambda_{12}\sigma_1\sigma_2 & 0 & \lambda_{23}\sigma_2\sigma_3 & \dots & \lambda_{2p}\sigma_2\sigma_p \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\lambda_{1p}\sigma_1\sigma_p & -\lambda_{2p}\sigma_2\sigma_p & -\lambda_{3p}\sigma_3\sigma_p & \dots & 0 \end{bmatrix} \text{ where} \quad (3.1.5)$$

$$\lambda_{st} = \frac{\sigma_{4st}}{\sigma_s\sigma_t}, s,t = 1,\dots,p, s<t. \quad (3.1.6)$$

The relationships given in (3.1.2) – (3.1.5) are of utmost importance, since the p-variate quaternion normal distribution are hereby defined. The expected value of

$$\underline{Z} = \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4$$

now follows from Remark 2.4.5 (1)

$$\begin{aligned}\underset{p\times 1}{\underline{\mu}} =& E\left[\underline{Z}\right] \\ =& \underline{\mu}_{X_1} + i\underline{\mu}_{X_2} + j\underline{\mu}_{X_3} + k\underline{\mu}_{X_4} \end{aligned} \tag{3.1.7}$$

while the quaternion-Hermitian covariance matrix of $\underline{Z}$ follows from Remark 2.4.5 (2) and (3.1.1) as

$$\begin{aligned}
\underset{p\times p}{\mathbf{\Sigma}} =& E\left[\underline{Z} - \underline{\mu}\right]\left[\overline{\underline{Z} - \underline{\mu}}\right]' \\
=& E\left[\left(\underline{X}_1 - \underline{\mu}_{X_1}\right) + i\left(\underline{X}_2 - \underline{\mu}_{X_2}\right) + j\left(\underline{X}_3 - \underline{\mu}_{X_3}\right) + k\left(\underline{X}_4 - \underline{\mu}_{X_4}\right)\right] \\
& \times \left[\left(\underline{X}_1 - \underline{\mu}_{X_1}\right) - i\left(\underline{X}_2 - \underline{\mu}_{X_2}\right) - j\left(\underline{X}_3 - \underline{\mu}_{X_3}\right) - k\left(\underline{X}_4 - \underline{\mu}_{X_4}\right)\right]' \\
=& E\left[\underline{X}_1 - \underline{\mu}_{X_1}\right]\left[\underline{X}_1 - \underline{\mu}_{X_1}\right]' - iE\left[\underline{X}_1 - \underline{\mu}_{X_1}\right]\left[\underline{X}_2 - \underline{\mu}_{X_2}\right]' \\
& - jE\left[\underline{X}_1 - \underline{\mu}_{X_1}\right]\left[\underline{X}_3 - \underline{\mu}_{X_3}\right]' - kE\left[\underline{X}_1 - \underline{\mu}_{X_1}\right]\left[\underline{X}_4 - \underline{\mu}_{X_4}\right]' \\
& + iE\left[\underline{X}_2 - \underline{\mu}_{X_2}\right]\left[\underline{X}_1 - \underline{\mu}_{X_1}\right]' + E\left[\underline{X}_2 - \underline{\mu}_{X_2}\right]\left[\underline{X}_2 - \underline{\mu}_{X_2}\right]' \\
& - kE\left[\underline{X}_2 - \underline{\mu}_{X_2}\right]\left[\underline{X}_3 - \underline{\mu}_{X_3}\right]' + jE\left[\underline{X}_2 - \underline{\mu}_{X_2}\right]\left[\underline{X}_4 - \underline{\mu}_{X_4}\right]' \\
& + jE\left[\underline{X}_3 - \underline{\mu}_{X_3}\right]\left[\underline{X}_1 - \underline{\mu}_{X_1}\right]' + kE\left[\underline{X}_3 - \underline{\mu}_{X_3}\right]\left[\underline{X}_2 - \underline{\mu}_{X_2}\right]' \\
& + E\left[\underline{X}_3 - \underline{\mu}_{X_3}\right]\left[\underline{X}_3 - \underline{\mu}_{X_3}\right]' - iE\left[\underline{X}_3 - \underline{\mu}_{X_3}\right]\left[\underline{X}_4 - \underline{\mu}_{X_4}\right]' \\
& + kE\left[\underline{X}_4 - \underline{\mu}_{X_4}\right]\left[\underline{X}_1 - \underline{\mu}_{X_1}\right]' - jE\left[\underline{X}_4 - \underline{\mu}_{X_4}\right]\left[\underline{X}_2 - \underline{\mu}_{X_2}\right]' \\
& + iE\left[\underline{X}_4 - \underline{\mu}_{X_4}\right]\left[\underline{X}_3 - \underline{\mu}_{X_3}\right]' + E\left[\underline{X}_4 - \underline{\mu}_{X_4}\right]\left[\underline{X}_4 - \underline{\mu}_{X_4}\right]' \\
=& \frac{1}{4}\mathbf{\Sigma}_1 + i\frac{1}{4}\mathbf{\Sigma}_2 + j\frac{1}{4}\mathbf{\Sigma}_3 + k\frac{1}{4}\mathbf{\Sigma}_4 \\
& + i\frac{1}{4}\mathbf{\Sigma}_2 + \frac{1}{4}\mathbf{\Sigma}_1 + k\frac{1}{4}\mathbf{\Sigma}_4 + j\frac{1}{4}\mathbf{\Sigma}_3 \\
& + j\frac{1}{4}\mathbf{\Sigma}_3 + k\frac{1}{4}\mathbf{\Sigma}_4 + \frac{1}{4}\mathbf{\Sigma}_1 + i\frac{1}{4}\mathbf{\Sigma}_2 \\
& + k\frac{1}{4}\mathbf{\Sigma}_4 + j\frac{1}{4}\mathbf{\Sigma}_3 + i\frac{1}{4}\mathbf{\Sigma}_2 + \frac{1}{4}\mathbf{\Sigma}_1 \\
=& \mathbf{\Sigma}_1 + i\mathbf{\Sigma}_2 + j\mathbf{\Sigma}_3 + k\mathbf{\Sigma}_4.
\end{aligned} \tag{3.1.8}$$

It is clear that:

$$\begin{aligned}
\operatorname{cov}\left(\underline{X}_1, \underline{X}_1'\right) &= \operatorname{cov}\left(\underline{X}_2, \underline{X}_2'\right) \\
&= \operatorname{cov}\left(\underline{X}_3, \underline{X}_3'\right) \\
&= \operatorname{cov}\left(\underline{X}_4, \underline{X}_4'\right) \\
&= \frac{1}{4}\mathbf{\Sigma}_1, \qquad (3.1.9) \\
-\operatorname{cov}\left(\underline{X}_1, \underline{X}_2'\right) &= \operatorname{cov}\left(\underline{X}_2, \underline{X}_1'\right) \\
&= \frac{1}{4}\mathbf{\Sigma}_2, \qquad (3.1.10) \\
-\operatorname{cov}\left(\underline{X}_1, \underline{X}_3'\right) &= \operatorname{cov}\left(\underline{X}_3, \underline{X}_1'\right) \\
&= \frac{1}{4}\mathbf{\Sigma}_3, \qquad (3.1.11) \\
-\operatorname{cov}\left(\underline{X}_1, \underline{X}_4'\right) &= \operatorname{cov}\left(\underline{X}_4, \underline{X}_1'\right) \\
&= \frac{1}{4}\mathbf{\Sigma}_4. \qquad (3.1.12)
\end{aligned}$$

(See Rautenbach (1983) [28] for a detailed discussion along these lines.)

Returning to the $\underline{Y}_s$, $s = 1, 2, 3, 4$ defined by Teng and Fang (1997) [31], they showed that $\frac{1}{4}\operatorname{cov}\left(\underline{Z}, \underline{Z}'\right) \simeq \operatorname{cov}\left(\underline{Y}_s, \underline{Y}_s'\right)$, $s = 1, 2, 3, 4$.

The pdf of $\underset{4p\times 1}{\underline{Z}_0}$ is given by

$$f_{\underline{Z}_0}(\underline{z}_0) = \pi^{-\frac{1}{2}(4p)} \left\{\det\left(2\mathbf{\Sigma}_0\right)\right\}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\underline{z}_0 - \underline{\mu}_0\right)' \mathbf{\Sigma}_0^{-1}\left(\underline{z}_0 - \underline{\mu}_0\right)\right\} \quad \text{for}$$

$$\underline{z}_0 \in B_0 = \left\{\underline{z}_0 = \left[x_{11}, \ldots, x_{1p}, x_{21}, \ldots, x_{2p}, x_{31}, \ldots, x_{3p}, x_{41}, \ldots, x_{4p}\right]', \right.$$
$$\left. -\infty < x_{1s}, x_{2s}, x_{3s}, x_{4s} < \infty, s = 1, \ldots, p\right\}.$$

In order to apply the representation theory as discussed in Chapter 2 and Appendix A, i.e. by an elementwise expansion of the quaternion probability vector, the components of the real associated probability vector now have to be arranged as follows:

$$\underset{4p\times 1}{\underline{Z}_0^*} = \left[X_{11}, X_{21}, X_{31}, X_{41}, \ldots, X_{1p}, X_{2p}, X_{3p}, X_{4p}\right]'.$$

The components of $\underline{\mu}_0$ and $\mathbf{\Sigma}_0$ are rearranged accordingly in forming $\underset{4p\times 1}{\underline{\mu}_0^*}$ and $\underset{4p\times 4p}{\mathbf{\Sigma}_0^*}$

respectively, and now yield the pdf of $\underline{Z}_0^*$ as

$$f_{\underline{Z}_0^*}(\underline{z}_0^*) = (2\pi)^{-\frac{1}{2}(4p)} \left\{\det \mathbf{\Sigma}_0^*\right\}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\underline{z}_0^* - \underline{\mu}_0^*\right)' \mathbf{\Sigma}_0^{*-1}\left(\underline{z}_0^* - \underline{\mu}_0^*\right)\right\} \tag{3.1.13}$$

for

$$\underline{z}_0^* \in B_0^* = \left\{\underline{z}_0^* = [x_{11}, x_{21}, x_{31}, x_{41}, \ldots, x_{1p}, x_{2p}, x_{3p}, x_{4p}]', \right.$$
$$\left. -\infty < x_{1s}, x_{2s}, x_{3s}, x_{4s} < \infty, s = 1, \ldots, p\right\}.$$

Since $f_{\underline{Z}_0}(\underline{z}_0) = f_{\underline{Z}_0^*}(\underline{z}_0^*)$ for all $\underline{z}_0$ and corresponding $\underline{z}_0^*$, $\underline{Z}_0^*$ may be used as real associated probability vector when deriving the pdf of $\underline{Z}$.

The real covariance matrix $\underset{4p\times 4p}{\mathbf{\Sigma}_0^*}$ of $\underline{Z}_0^* = [X_{11}, X_{21}, X_{31}, X_{41}, \ldots, X_{1p}, X_{2p}, X_{3p}, X_{4p}]'$ are given by (see Rautenbach (1983) [28])

$$\underset{4p\times 4p}{\mathbf{\Sigma}_0^*} = [\mathbf{\Sigma}_{0st}^*] \tag{3.1.14}$$

where $\underset{4\times 4}{\mathbf{\Sigma}_{0st}^*}$ is of the form

$$\underset{4\times 4}{\mathbf{\Sigma}_{0ss}^*} = \frac{1}{4}\begin{bmatrix} \sigma_s^2 & 0 & 0 & 0 \\ 0 & \sigma_s^2 & 0 & 0 \\ 0 & 0 & \sigma_s^2 & 0 \\ 0 & 0 & 0 & \sigma_s^2 \end{bmatrix}$$

for $s = 1, \ldots, p$ and

$$\underset{4\times 4}{\mathbf{\Sigma}_{0st}^*} = \frac{1}{4}\begin{bmatrix} \eta_{st} & -\alpha_{st} & -\beta_{st} & -\lambda_{st} \\ \alpha_{st} & \eta_{st} & -\lambda_{st} & \beta_{st} \\ \beta_{st} & \lambda_{st} & \eta_{st} & -\alpha_{st} \\ \lambda_{st} & -\beta_{st} & \alpha_{st} & \eta_{st} \end{bmatrix}\sigma_s\sigma_t$$

for $s, t = 1, \ldots, p, s \neq t$. This covariance structure is a key feature of the p-variate quaternion normal distribution, since $\mathbf{\Sigma}_0^*$ is a matrix with elements similar to that given in Definition A.5.1. Thus, it follows that $\mathbf{\Sigma}_0^* \in M_{4p}(\mathbb{R})$, more specifically $4\mathbf{\Sigma}_0^* \in M_{4p}(\mathbb{R})$. From Definition A.5.4 it follows that operations on $M_{4p}(\mathbb{R})$, with matrices of the form

$4\mathbf{\Sigma}_0^*$, are isomorphic to operations on $M_p(\mathbb{Q})$. Thus, $M_p(\mathbb{Q})$ is a set of matrices of the form

$$\underset{p\times p}{\mathbf{\Sigma}} = [\sigma_{st}]$$

where

$$\sigma_{st} = \begin{cases} (\eta_{st} + i\alpha_{st} + j\beta_{st} + k\lambda_{st})\,\sigma_s\sigma_t, & s \neq t \\ \sigma_s^t, & s = t. \end{cases} \tag{3.1.15}$$

Hence, it follows that $4\mathbf{\Sigma}_0^* \simeq \mathbf{\Sigma}$ (see Remark A.5.6).

The pdf of $\underset{p\times 1}{\underline{Z}}$, where $\underline{Z}$ has a p-variate quaternion normal distribution is now derived.

Theorem 3.1.2. *Let $\underset{p\times 1}{\underline{Z}}$ be a quaternion probability vector that has a p-variate quaternion normal distribution, as given in Definition 3.1.1, with $E[\underline{Z}] \equiv \underline{\mu}$ and $\mathbf{\Sigma} \equiv [\sigma_{st}]$ as given in (3.1.15). The pdf of $\underset{p\times 1}{\underline{Z}} \sim \mathbb{Q}N(p;\underline{\mu},\mathbf{\Sigma})$, is given by:*

$$f_{\underline{Z}}(\underline{z}) = 2^{2p}\pi^{-2p}\left(\det\mathbf{\Sigma}\right)^{-2}\exp\left\{-2\left(\overline{\underline{z}-\underline{\mu}}\right)'\mathbf{\Sigma}^{-1}\left(\underline{z}-\underline{\mu}\right)\right\} \tag{3.1.16}$$

where

$$\underline{z} \in B = \left\{\underline{z} = [z_1,\ldots,z_p]' : z_s = x_{1s} + ix_{2s} + jx_{3s} + kx_{4s},\right.$$
$$\left. -\infty < x_{1s}, x_{2s}, x_{3s}, x_{4s} < \infty, s = 1,\ldots,p\right\}.$$

(See Rautenbach (1983) [28].)

Proof. 1. From Definition 2.3.3 it follows that

$$f_{\underline{Z}}(\underline{z}) = f_{\underline{Z}_0^*}(\underline{z}_0^*)$$

where $\underline{Z}_0^*$ is the real associated probability vector of $\underline{Z}$. The pdf of $\underline{Z}_0^*$ is given by

$$f_{\underline{Z}_0^*}(\underline{z}_0^*) = 2^{-2p}\pi^{-2p}\left\{\det\mathbf{\Sigma}_0^*\right\}^{-\frac{1}{2}}\exp\left\{-\frac{1}{2}\left(\underline{z}_0^* - \underline{\mu}_0^*\right)'\mathbf{\Sigma}_0^{*-1}\left(\underline{z}_0^* - \underline{\mu}_0^*\right)\right\}$$

for $\underline{z}_0^* \in B_0^*$.

2. It now follows that $4\mathbf{\Sigma}_0^* \simeq \mathbf{\Sigma}$. Since $\mathbf{\Sigma}_0^*$ is symmetric positive definite, it furthermore follows from Corollary A.5.15 that $\mathbf{\Sigma}$ is a positive definite quaternion

Hermitian matrix, particularly that $\mathbf{\Sigma}$ is nonsingular (see Theorem A.5.9). From Theorem A.5.12 it now follows that

$$\det\left(4\mathbf{\Sigma}_0^*\right) = \left(\det \mathbf{\Sigma}\right)^4$$
$$\therefore \det \mathbf{\Sigma}_0^* = 4^{-4p}\left(\det \mathbf{\Sigma}\right)^4 \quad (3.1.17)$$
$$\therefore \left(\det \mathbf{\Sigma}_0^*\right)^{-\frac{1}{2}} = 4^{2p}\left(\det \mathbf{\Sigma}\right)^{-2}. \quad (3.1.18)$$

3. From Theorem A.5.9 it also follows that

$$\frac{1}{4}\mathbf{\Sigma}_0^{*-1} \simeq \mathbf{\Sigma}^{-1} \quad (3.1.19)$$

so that Theorem A.5.13 yields

$$2\left(\underline{z}_0^* - \underline{\mu}_0^*\right)' \frac{1}{4}\mathbf{\Sigma}_0^{*-1}\left(\underline{z}_0^* - \underline{\mu}_0^*\right) = 2\left(\overline{\underline{z} - \underline{\mu}}\right)' \mathbf{\Sigma}^{-1}\left(\underline{z} - \underline{\mu}\right). \quad (3.1.20)$$

4. From (3.1.18) and (3.1.20) together with the expression for $f_{\underline{Z}_0^*}(\underline{z}_0^*)$ it finally follows that:

$$f_{\underline{Z}}(\underline{z}) = 2^{2p}\pi^{-2p}\left(\det \mathbf{\Sigma}\right)^{-2}\exp\left\{-2\left(\overline{\underline{z} - \underline{\mu}}\right)' \mathbf{\Sigma}^{-1}\left(\underline{z} - \underline{\mu}\right)\right\},$$

for $\underline{z} \in B$.

□

Remark 3.1.3. 1. Let $\mathrm{tr}(\cdot)$ be the trace on $M_p(\mathbb{Q})$. Rautenbach (1983) [28] argued that the trace, as used by Kabe (1976) [17], (1978) [18], and (1984) [19], in simplifying the exponent of $f_{\underline{Z}}(\underline{z})$, is not allowed, because of the noncommutativity property of quaternions, i.e.

$$\mathrm{tr}(\mathbf{AB}) \neq \mathrm{tr}(\mathbf{BA}).$$

However, from (2.1.2), and the fact that for a quaternion random vector $\underset{p\times 1}{\underline{Z}}$, it is known that $\bar{\underline{Z}}'\underline{Z} \in \mathbb{R}$, and from Theorem A.5.13 it follows that

$$\exp\left\{\bar{\underline{Z}}'\mathbf{\Sigma}^{-1}\underline{Z}\right\} = \exp\left\{\mathrm{tr}\,Re\left(\bar{\underline{Z}}'\mathbf{\Sigma}^{-1}\underline{Z}\right)\right\}$$
$$= \exp\left\{\mathrm{Retr}\left(\bar{\underline{Z}}'\mathbf{\Sigma}^{-1}\underline{Z}\right)\right\}$$
$$= \exp\left\{\mathrm{Retr}\left(\mathbf{\Sigma}^{-1}\underline{Z}\bar{\underline{Z}}'\right)\right\}. \quad (3.1.21)$$

2. It is important to note that this quaternion distribution is dependent on the assumptions made in (3.1.2) through (3.1.5), regarding the covariance structure of $\underline{Z}_0$. However, if any alternative arbitrary covariance structure was chosen for $\underline{Z}_0$, it would not have been possible to derive the corresponding pdf. Thus, the assumptions made regarding the covariance structure of $\underline{Z}_0$ naturally yielded a simple form of the derived pdf.

3. If a quaternion probability vector $\underset{p\times 1}{\underline{Z}}$ has a p-variate quaternion normal distribution, with pdf given in (3.1.16), it will be denoted by

$$\underline{Z} \sim \mathbb{Q}N(p; \underline{\mu}, \mathbf{\Sigma}),$$

where $\underset{p\times 1}{\underline{\mu}}$ and $\underset{p\times p}{\mathbf{\Sigma}}$ is the expected value and covariance matrix of $\underline{Z}$ respectively.

The univariate and bivariate quaternion normal distributions are now presented as special cases of the p-variate quaternion normal distribution.

Example 3.1.4. The pdf of the univariate quaternion normal distribution is obtained by setting $p = 1$ in (3.1.16), and is given by

$$\begin{aligned} f_Z(z) =& 4\pi^{-2}\sigma_1^{-4} \exp\left\{-\frac{2}{\sigma_1^2}\left(\overline{z-\mu}\right)(z-\mu)\right\} \qquad (3.1.22) \\ \text{for } z \in B =& \{z = x_1 + ix_2 + jx_3 + kx_4, \\ & -\infty < x_1, x_2, x_3, x_4 < \infty\}, \\ \sigma_1^2 \equiv& \operatorname{var}(Z) = \operatorname{var}(X_1) + \operatorname{var}(X_2) + \operatorname{var}(X_3) + \operatorname{var}(X_4) \\ \text{and } \mu \equiv& E[Z] = E[X_1] + iE[X_2] + jE[X_3] + kE[X_4] \end{aligned}$$

are the variance and expected value of the quaternion probability variable $Z = X_1 + iX_2 + jX_3 + kX_4 \sim \mathbb{Q}N(\mu, \sigma_1^2)$, respectively. Note that

$$\mathbf{\Sigma}_0^* = \frac{1}{4}\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 \\ 0 & 0 & 0 & \sigma_1^2 \end{bmatrix} \simeq \sigma_1^2.$$

Example 3.1.5. The pdf of the bivariate quaternion normal distribution is obtained by setting $p = 2$ in (3.1.16), and is given by

$$f_{\underline{Z}}(\underline{z}) = 2^4\pi^{-4}\left(\det \mathbf{\Sigma}\right)^{-2}\exp\left\{-2\left(\overline{\underline{z}-\underline{\mu}}\right)'\mathbf{\Sigma}^{-1}\left(\underline{z}-\underline{\mu}\right)\right\} \tag{3.1.23}$$

for

$$\begin{aligned}\underline{z} \in B = \{&\underline{z} = [z_1, z_2]', \\ &z_s = x_{1s} + ix_{2s} + jx_{3s} + kx_{4s} \\ &-\infty < x_{1s}, x_{2s}, x_{3s}, x_{4s} < \infty, s = 1, 2\},\end{aligned}$$

and

$$\underset{2\times 2}{\mathbf{\Sigma}} = \begin{bmatrix} \sigma_1^2 & (\eta_{12} + i\alpha_{12} + j\beta_{12} + k\lambda_{12})\,\sigma_1\sigma_2 \\ (\eta_{12} - i\alpha_{12} - j\beta_{12} - k\lambda_{12})\,\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

and

$$\begin{aligned}\underset{2\times 1}{\underline{\mu}} \equiv& E\left[\underline{Z}\right] \\ =& E\left[\underline{X}_1\right] + iE\left[\underline{X}_2\right] + jE\left[\underline{X}_3\right] + kE\left[\underline{X}_4\right]\end{aligned}$$

are the covariance matrix and expected value of the quaternion probability vector $\underline{Z} = \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4 \sim \mathbb{Q}N(2;\underline{\mu},\mathbf{\Sigma})$, respectively. Note that

$$4\mathbf{\Sigma}_0^* = \begin{bmatrix} \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_1^2 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & 0 \\ 0 & 0 & 0 & \sigma_1^2 \end{bmatrix} & \begin{bmatrix} \eta_{12} & -\alpha_{12} & -\beta_{12} & -\lambda_{12} \\ \alpha_{12} & \eta_{12} & -\lambda_{12} & \beta_{12} \\ \beta_{12} & \lambda_{12} & \eta_{12} & -\alpha_{12} \\ \lambda_{12} & -\beta_{12} & \alpha_{12} & \eta_{12} \end{bmatrix}\sigma_1\sigma_2 \\ \sigma_1\sigma_2\begin{bmatrix} \eta_{12} & \alpha_{12} & \beta_{12} & \lambda_{12} \\ -\alpha_{12} & \eta_{12} & \lambda_{12} & -\beta_{12} \\ -\beta_{12} & -\lambda_{12} & \eta_{12} & \alpha_{12} \\ -\lambda_{12} & \beta_{12} & -\alpha_{12} & \eta_{12} \end{bmatrix} & \begin{bmatrix} \sigma_2^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_2^2 & 0 \\ 0 & 0 & 0 & \sigma_2^2 \end{bmatrix} \end{bmatrix}$$
$$\simeq \begin{bmatrix} \sigma_1^2 & (\eta_{12} + i\alpha_{12} + j\beta_{12} + k\lambda_{12})\,\sigma_1\sigma_2 \\ (\eta_{12} - i\alpha_{12} - j\beta_{12} - k\lambda_{12})\,\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

Also note that (3.1.23) may be rewritten as

$$\begin{aligned} & f_{Z_1,Z_2}(z_1,z_2) \qquad (3.1.24)\\ =& 2^4\pi^{-4}\left[\left(1-\left(\eta_{12}^2+\alpha_{12}^2+\beta_{12}^2+\lambda_{12}^2\right)\right)\sigma_1^2\sigma_2^2\right]^{-2}\\ & \times\exp\left\{-2\left[\sigma_2^2\left|z_1-\mu_1\right|^2-2\,\mathrm{Re}\left(\left(\overline{z_1-\mu_1}\right)\left(\left(\eta_{12}+i\alpha_{12}+j\beta_{12}+k\lambda_{12}\right)\sigma_1\sigma_2\right)\left(z_2-\mu_2\right)\right)\right.\right.\\ & \left.\left.+\sigma_1^2\left|z_2-\mu_2\right|^2\right]\left[\sigma_1^2\sigma_2^2-\left(\eta_{12}^2+\alpha_{12}^2+\beta_{12}^2+\lambda_{12}^2\right)\sigma_1^2\sigma_2^2\right]^{-1}\right\}\end{aligned}$$

where $z_1, z_2 \in B$ and Re indicates that the real part of the expression is used. Furthermore,

$$\left|z_1-\mu_1\right|^2 = \left(\overline{z_1-\mu_1}\right)\left(z_1-\mu_1\right)$$

and

$$\left|z_2-\mu_2\right|^2 = \left(\overline{z_2-\mu_2}\right)\left(z_2-\mu_2\right).$$

Now, let

$$\rho = \eta_{12}+i\alpha_{12}+j\beta_{12}+k\lambda_{12}$$

be the quaternion correlation coefficient, then it follows that

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \bar{\rho}\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

(3.1.24) now becomes

$$\begin{aligned} & f_{Z_1,Z_2}(z_1,z_2) \qquad (3.1.25)\\ =& 2^4\pi^{-4}\left(\sigma_1^2\sigma_2^2(1-\bar{\rho}\rho)\right)^{-2}\\ & \times\exp\left\{-2\left[\sigma_2^2\left|z_1-\mu_1\right|^2-2\,\mathrm{Re}\left(\left(\overline{z_1-\mu_1}\right)\sigma_1\sigma_2\rho(z_2-\mu_2)\right)+\sigma_1^2\left|z_2-\mu_2\right|^2\right]\right.\\ & \left.\times\left[\sigma_1^2\sigma_2^2\left(1-\bar{\rho}\rho\right)\right]^{-1}\right\}\end{aligned}$$

where $z_1, z_2 \in B$ and where Re, $\left|z_1-\mu_1\right|^2$ and $\left|z_2-\mu_2\right|^2$ have similar meanings as in (3.1.24).

The study thus far relied heavily on a covariance structure of the form:

$$\mathbf{\Sigma}_0 = \frac{1}{4}\begin{bmatrix} \mathbf{\Sigma}_1 & -\mathbf{\Sigma}_2 & -\mathbf{\Sigma}_3 & -\mathbf{\Sigma}_4 \\ \mathbf{\Sigma}_2 & \mathbf{\Sigma}_1 & -\mathbf{\Sigma}_4 & \mathbf{\Sigma}_3 \\ \mathbf{\Sigma}_3 & \mathbf{\Sigma}_4 & \mathbf{\Sigma}_1 & -\mathbf{\Sigma}_2 \\ \mathbf{\Sigma}_4 & -\mathbf{\Sigma}_3 & \mathbf{\Sigma}_2 & \mathbf{\Sigma}_1 \end{bmatrix}$$

which implied a quaternion normal pdf with exponent of the form

$$\left(\overline{\underline{z}-\underline{\mu}}\right)'\boldsymbol{\Sigma}^{-1}\left(\underline{z}-\underline{\mu}\right).$$

If a covariance structure of the form

$$\boldsymbol{\Sigma}_{00}=\frac{1}{4}\begin{bmatrix}\boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_3 & \boldsymbol{\Sigma}_4\\ -\boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1 & -\boldsymbol{\Sigma}_4 & \boldsymbol{\Sigma}_3\\ -\boldsymbol{\Sigma}_3 & \boldsymbol{\Sigma}_4 & \boldsymbol{\Sigma}_1 & -\boldsymbol{\Sigma}_2\\ -\boldsymbol{\Sigma}_4 & -\boldsymbol{\Sigma}_3 & \boldsymbol{\Sigma}_2 & \boldsymbol{\Sigma}_1\end{bmatrix}$$

is used instead, it results in a exponent for the pdf of the p-variate quaternion normal distribution in the form

$$\left(\underline{z}-\underline{\mu}\right)'\boldsymbol{\Sigma}^{-1}\left(\overline{\underline{z}-\underline{\mu}}\right).$$

Thus, the p-variate quaternion normal pdf can be written as

$$f_{\underline{Z}}(\underline{z})=2^{2p}\pi^{-2p}\left(\det\boldsymbol{\Sigma}\right)^{-2}\exp\left\{-2\left(\overline{\underline{z}-\underline{\mu}}\right)'\boldsymbol{\Sigma}^{-1}\left(\underline{z}-\underline{\mu}\right)\right\}$$

with $\boldsymbol{\Sigma}_0$ as associated covariance matrix or

$$f_{\underline{Z}}(\underline{z})=2^{2p}\pi^{-2p}\left(\det\boldsymbol{\Sigma}\right)^{-2}\exp\left\{-2\left(\underline{z}-\underline{\mu}\right)'\boldsymbol{\Sigma}^{-1}\left(\overline{\underline{z}-\underline{\mu}}\right)\right\}$$

with $\boldsymbol{\Sigma}_{00}$ as the associated covariance matrix, which is obtained by using the representation used by Teng and Fang (1997) [31] as discussed earlier.

The remainder of this section will focus on the cf of the p-variate quaternion normal distribution.

Theorem 3.1.6. *The cf of* $\underset{p\times 1}{\underline{Z}}\sim\mathbb{Q}N(p;\underline{\mu},\boldsymbol{\Sigma})$, *is given by:*

$$\phi_{\underline{Z}}(\underline{t})=\exp\left\{\frac{\iota}{2}\left(\underline{\bar{\mu}}'\underline{t}+\underline{\bar{t}}'\underline{\mu}\right)-\frac{1}{8}\underline{\bar{t}}'\boldsymbol{\Sigma}\underline{t}\right\}\tag{3.1.26}$$

for every quaternion vector $\underset{p\times 1}{\underline{t}}$ *and* ι *the usual imaginary complex root.*

Proof. From (2.4.10) and (3.1.16) it follows that

$$\begin{aligned}\phi_{\underline{Z}}(\underline{t})=&E\left[\exp\frac{\iota}{2}\left(\underline{\bar{Z}}'\underline{t}+\underline{\bar{t}}'\underline{Z}\right)\right]\\ =&\int_B 2^{2p}\pi^{-2p}\left(\det\boldsymbol{\Sigma}\right)^{-2}\exp\left\{\frac{\iota}{2}\left(\underline{\bar{z}}'\underline{t}+\underline{\bar{t}}'\underline{z}\right)-2\left(\overline{\underline{z}-\underline{\mu}}\right)'\boldsymbol{\Sigma}^{-1}\left(\underline{z}-\underline{\mu}\right)\right\}d\underline{z},\end{aligned}$$

where $\underset{p\times 1}{\underline{t}} = \underline{t}_1 + i\underline{t}_2 + j\underline{t}_3 + k\underline{t}_4$ is a quaternion vector and

$$B = \{\underline{z} = [z_1, \ldots, z_p]' : z_s = x_{1s} + ix_{2s} + jx_{3s} + kx_{4s};$$
$$-\infty < x_{1s}, x_{2s}, x_{3s}, x_{4s} < \infty, s = 1, \ldots, p\}$$

is the region of variability for $\underline{Z}$. The exponent in this expression may be written as

$$\begin{aligned}
&-2\left\{\overline{\underline{z}}'\mathbf{\Sigma}^{-1}\underline{z} - \overline{\underline{z}}'\mathbf{\Sigma}^{-1}\underline{\mu} - \overline{\underline{\mu}}'\mathbf{\Sigma}^{-1}\underline{z} + \overline{\underline{\mu}}'\mathbf{\Sigma}^{-1}\underline{\mu} - \frac{\iota}{4}\overline{\underline{z}}'\underline{t} - \frac{\iota}{4}\overline{\underline{t}}'\underline{z}\right\} \\
=&-2\left\{\overline{\underline{z}}'\mathbf{\Sigma}^{-1}\underline{z} - \left(\overline{\underline{\mu}}' + \frac{\iota}{4}\overline{\underline{t}}'\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\underline{z} - \overline{\underline{z}}'\mathbf{\Sigma}^{-1}\left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right) + \overline{\underline{\mu}}'\mathbf{\Sigma}^{-1}\underline{\mu}\right\} \\
=&-2\left\{\left(\overline{\underline{z} - \left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)}\right)'\mathbf{\Sigma}^{-1}\left(\underline{z} - \left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)\right) + \overline{\underline{\mu}}'\mathbf{\Sigma}^{-1}\underline{\mu} - \left(\overline{\underline{\mu}}' + \frac{\iota}{4}\overline{\underline{t}}'\mathbf{\Sigma}\right)\mathbf{\Sigma}^{-1}\left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)\right\} \\
=&-2\left\{\left(\overline{\underline{z} - \left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)}\right)'\mathbf{\Sigma}^{-1}\left(\underline{z} - \left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)\right) + \overline{\underline{\mu}}'\mathbf{\Sigma}^{-1}\underline{\mu} - \overline{\underline{\mu}}'\mathbf{\Sigma}^{-1}\underline{\mu} - \frac{\iota}{4}\overline{\underline{t}}'\underline{\mu} - \frac{\iota}{4}\overline{\underline{\mu}}'\underline{t} + \frac{1}{16}\overline{\underline{t}}'\mathbf{\Sigma}\underline{t}\right\} \\
=&-2\left\{\left(\overline{\underline{z} - \left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)}\right)'\mathbf{\Sigma}^{-1}\left(\underline{z} - \left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)\right)\right\} + \frac{\iota}{2}\left(\overline{\underline{t}}'\underline{\mu} + \overline{\underline{\mu}}'\underline{t}\right) - \frac{1}{8}\overline{\underline{t}}'\mathbf{\Sigma}\underline{t}
\end{aligned}$$

From the above expression and from the definition of the p-variate quaternion normal pdf it follows that

$$\begin{aligned}
\phi_{\underline{Z}}(\underline{t}) =& \exp\left\{\frac{\iota}{2}\left(\overline{\underline{t}}'\underline{\mu} + \overline{\underline{\mu}}'\underline{t}\right) - \frac{1}{8}\overline{\underline{t}}'\mathbf{\Sigma}\underline{t}\right\}\int_B 2^{2p}\pi^{-2p}\left(\det\mathbf{\Sigma}\right)^{-2} \\
&\times\exp\left\{-2\left[\overline{\underline{z} - \left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)}\right]'\mathbf{\Sigma}^{-1}\left[\underline{z} - \left(\underline{\mu} + \frac{\iota}{4}\mathbf{\Sigma}\underline{t}\right)\right]\right\}d\underline{z} \\
=& \exp\left\{\frac{\iota}{2}\left(\overline{\underline{t}}'\underline{\mu} + \overline{\underline{\mu}}'\underline{t}\right) - \frac{1}{8}\overline{\underline{t}}'\mathbf{\Sigma}\underline{t}\right\}.
\end{aligned}$$

□

By setting $v = \overline{\underline{\mu}}'\underline{t} + \overline{\underline{t}}'\underline{\mu}$ then $\overline{v}' = \overline{\underline{t}}'\underline{\mu} + \overline{\underline{\mu}}'\underline{t} = v$, i.e. $v \in \mathbb{R}$, so that (3.1.26) may be rewritten in the form

$$\phi_{\underline{Z}}(\underline{t}) = \exp\left\{\iota\,\mathrm{Re}\left(\overline{\underline{\mu}}'\underline{t}\right) - \frac{1}{8}\overline{\underline{t}}'\mathbf{\Sigma}\underline{t}\right\} = \exp\left\{\iota\underline{\mu}_0^{*\prime}\underline{t}_0^* - \frac{1}{2}\underline{t}_0^{*\prime}\mathbf{\Sigma}_0^*\underline{t}_0^*\right\} = \phi_{\underline{Z}_0^*}(\underline{t}_0^*),$$

so that (2.4.13) holds in the multivariate case.

An additional result is now presented that is often required for application purposes.

Theorem 3.1.7. *Let* $\underline{Z} \sim \mathbb{Q}N\left(p; \underline{\mu}, \mathbf{\Sigma}\right)$ *then* $\underline{Y} = \mathbf{C}\left(\underline{Z} - \underline{\mu}\right) \sim \mathbb{Q}N\left(p; \underline{0}, \mathbf{I}_p\right)$ *if* $\mathbf{C}$ *is selected in such a way that* $\mathbf{C}\mathbf{\Sigma}\bar{\mathbf{C}}' = \mathbf{I}_p$ *and* $\mathbf{C}$ *is a symplectic matrix. (See Rautenbach (1983) [28].)*

Proof. Let $\underset{p\times p}{\mathbf{C}}$ be such that $\mathbf{C}\mathbf{\Sigma}\bar{\mathbf{C}}' = \mathbf{I}_p$. The cf of $\underline{Y}$ is given by

$$\begin{aligned}
\phi_{\underline{Y}}\left(\underline{t}\right) =& E\left[\exp\left\{\frac{\iota}{2}\left(\bar{\underline{Y}}'\underline{t} + \bar{\underline{t}}'\underline{Y}\right)\right\}\right] \\
=& \exp\left\{-\frac{\iota}{2}\left(\bar{\underline{\mu}}'\bar{\mathbf{C}}'\underline{t} + \bar{\underline{t}}'\mathbf{C}\underline{\mu}\right)\right\} E\left[\exp\left\{\frac{\iota}{2}\left(\bar{\underline{Z}}'\bar{\mathbf{C}}'\underline{t} + \bar{\underline{t}}'\mathbf{C}\underline{Z}\right)\right\}\right] \\
=& \exp\left\{-\frac{\iota}{2}\left(\bar{\underline{\mu}}'\bar{\mathbf{C}}'\underline{t} + \bar{\underline{t}}'\mathbf{C}\underline{\mu}\right)\right\} \\
& \times \exp\left\{\frac{\iota}{2}\left(\bar{\underline{\mu}}'\bar{\mathbf{C}}'\underline{t} + \bar{\underline{t}}'\mathbf{C}\underline{\mu}\right) - \frac{1}{8}\bar{\underline{t}}'\mathbf{C}\mathbf{\Sigma}\bar{\mathbf{C}}'\underline{t}\right\} \\
=& \exp\left\{-\frac{1}{8}\bar{\underline{t}}'\underline{t}\right\}
\end{aligned}$$

from which the desired result follows. □

3.2 The matrix-variate quaternion normal distribution

The matrix-variate quaternion normal distribution is now discussed by an expansion of the results presented in the previous section. Once again the aim is to emphasise the relationship between the real associated form and its resultant counterpart.

Theorem 3.2.1. *Let* $\underset{p\times 1}{\underline{Z}_\alpha}, \alpha = 1, \ldots, n$ *be* n *probability vectors each having a* p*-variate quaternion normal distribution, as given in Definition 3.1.1. Now, suppose that*

$$\begin{aligned}
\underset{n\times p}{\mathbf{Z}} &= [Z_{\alpha\beta}], \;\; \alpha = 1, \ldots, n, \;\; \beta = 1, \ldots, p \\
&= \begin{bmatrix} Z_{11} & \ldots & Z_{1p} \\ \vdots & \ddots & \vdots \\ Z_{n1} & \ldots & Z_{np} \end{bmatrix} \\
&= \begin{bmatrix} \underset{1\times p}{\underline{Z}'_1} \\ \vdots \\ \underset{1\times p}{\underline{Z}'_n} \end{bmatrix} = \begin{bmatrix} \underset{n\times 1}{\underline{Z}_{(1)}}, \ldots, \underset{n\times 1}{\underline{Z}_{(p)}} \end{bmatrix}
\end{aligned}$$

i.e. the rows of $\mathbf{Z}$ *are* $\mathbb{Q}N(p; \underline{\mu}_\alpha, \mathbf{\Sigma})$ *distributed,* $\alpha = 1, \ldots, n$ *with dependence structure given by* $\underset{n\times n}{\mathbf{R}}$ *not necessarily equal to* $\mathbf{I}_n$*. It may be assumed without loss of generality that* $\mathbf{R}$ *is real-valued. Similarly, define*

$$\underset{n\times p}{\mu} = [\mu_{\alpha\beta}], \;\; \alpha = 1, \ldots, n, \;\; \beta = 1, \ldots, p.$$

Then

$$\underset{np\times 1}{\mathrm{vec}\,\mathbf{Z}} = \begin{bmatrix} \underset{n\times 1}{\underline{Z}_{(1)}} \\ \vdots \\ \underset{n\times 1}{\underline{Z}_{(p)}} \end{bmatrix} \sim \mathbb{Q}N\left(np; \mathrm{vec}\,\mu, \mathbf{\Sigma} \otimes \mathbf{R}\right)$$

i.e. a matrix-variate quaternion normal distribution where

$$\underset{np\times 1}{\mathrm{vec}\,\mu} = \begin{bmatrix} \underset{n\times 1}{\underline{\mu}_{(1)}} \\ \vdots \\ \underset{n\times 1}{\underline{\mu}_{(p)}} \end{bmatrix}.$$

Proof. 1. In order to apply the methodology set out in Section 3.1, it is necessary to define the real associated probability vector of $\underline{Z}_{(\beta)}$. First, observe that

$$\underset{n\times 1}{\underline{Z}_{(\beta)}} = \begin{bmatrix} Z_{1\beta} \\ \vdots \\ Z_{n\beta} \end{bmatrix} = \begin{bmatrix} X_{11\beta} + iX_{21\beta} + jX_{31\beta} + kX_{41\beta} \\ \vdots \\ X_{1n\beta} + iX_{2n\beta} + jX_{3n\beta} + kX_{4n\beta} \end{bmatrix} = \underset{n\times 1}{\underline{X}_{(1\beta)}} + i\underset{n\times 1}{\underline{X}_{(2\beta)}} + j\underset{n\times 1}{\underline{X}_{(3\beta)}} + k\underset{n\times 1}{\underline{X}_{(4\beta)}}$$

with associated real counterpart given by

$$\underset{n\times 4}{\underline{Z}^*_{(0\beta)}} = \left[\underset{n\times 1}{\underline{X}_{(1\beta)}}, \underset{n\times 1}{\underline{X}_{(2\beta)}}, \underset{n\times 1}{\underline{X}_{(3\beta)}}, \underset{n\times 1}{\underline{X}_{(4\beta)}}\right].$$

From this it now follows that

$$\underset{n\times p}{\mathbf{Z}} = \left[\underset{n\times 1}{\underline{Z}_{(1)}}, \ldots, \underset{n\times 1}{\underline{Z}_{(p)}}\right] = \left[\underset{n\times 1}{\underline{X}_{(11)}} + i\underset{n\times 1}{\underline{X}_{(21)}} + j\underset{n\times 1}{\underline{X}_{(31)}} + k\underset{n\times 1}{\underline{X}_{(41)}}, \ldots, \underset{n\times 1}{\underline{X}_{(1p)}} + i\underset{n\times 1}{\underline{X}_{(2p)}} + j\underset{n\times 1}{\underline{X}_{(3p)}} + k\underset{n\times 1}{\underline{X}_{(4p)}}\right]$$

from which the real associated matrix $\underset{n\times 4p}{\mathbf{Z}^*_0}$ of $\underset{n\times p}{\mathbf{Z}}$ immediately follows as

$$\underset{n\times 4p}{\mathbf{Z}^*_0} = \left[X_{\gamma\alpha\beta}\right], \ \gamma = 1, \ldots, 4, \ \alpha = 1, \ldots, n, \ \beta = 1, \ldots, p$$
$$= \left[\underset{n\times 1}{\underline{X}_{(11)}}, \underset{n\times 1}{\underline{X}_{(21)}}, \underset{n\times 1}{\underline{X}_{(31)}}, \underset{n\times 1}{\underline{X}_{(41)}}, \ldots, \underset{n\times 1}{\underline{X}_{(1p)}}, \underset{n\times 1}{\underline{X}_{(2p)}}, \underset{n\times 1}{\underline{X}_{(3p)}}, \underset{n\times 1}{\underline{X}_{(4p)}}\right]$$
$$= \left[\underset{n\times 4}{\underline{Z}^*_{(01)}}, \ldots, \underset{n\times 4}{\underline{Z}^*_{(0p)}}\right].$$

If the supposition is made that

$$\operatorname{vec} \underset{4n\times 1}{\underline{Z}^*_{(0\beta)}} = \begin{bmatrix} \underset{n\times 1}{\underline{X}_{(1\beta)}} \\ \underset{n\times 1}{\underline{X}_{(2\beta)}} \\ \underset{n\times 1}{\underline{X}_{(3\beta)}} \\ \underset{n\times 1}{\underline{X}_{(4\beta)}} \end{bmatrix},$$

then

$$\begin{aligned} \underset{4np\times 1}{\operatorname{vec} \mathbf{Z}^*_0} &= \left[\underset{1\times n}{\underline{X}'_{(11)}}, \underset{1\times n}{\underline{X}'_{(21)}}, \underset{1\times n}{\underline{X}'_{(31)}}, \underset{1\times n}{\underline{X}'_{(41)}}, \ldots, \underset{1\times n}{\underline{X}'_{(1p)}}, \underset{1\times n}{\underline{X}'_{(2p)}}, \underset{1\times n}{\underline{X}'_{(3p)}}, \underset{1\times n}{\underline{X}'_{(4p)}} \right]' \\ &= \begin{bmatrix} \operatorname{vec} \underset{4n\times 1}{\underline{Z}^*_{(01)}} \\ \vdots \\ \operatorname{vec} \underset{4n\times 1}{\underline{Z}^*_{(0p)}} \end{bmatrix} \\ &\simeq \underset{np\times 1}{\operatorname{vec} \mathbf{Z}}. \end{aligned}$$

In a similar fashion, it can be shown that

$$\underset{4np\times 1}{\operatorname{vec} \mu^*_0} \simeq \underset{np\times 1}{\operatorname{vec} \mu}.$$

2. The problem may now be rewritten in terms of n real associated probability vectors as $\underset{4p\times 1}{\underline{Z}^*_{0\alpha}}$, $\alpha = 1, \ldots, n$ each having a $4p$-variate real normal distribution, as given in (3.1.13). The real associated quantity $\underset{4np\times 1}{\operatorname{vec} \mathbf{Z}^*_0}$ now has a density given by $N\left(\operatorname{vec} \mu^*_0, \boldsymbol{\Sigma}^*_0 \otimes \mathbf{R}\right)$, i.e. a real matrix-variate normal distribution, where $\underset{n\times n}{\mathbf{R}}$ denotes the dependence structure of the rows of $\mathbf{Z}^*_0$ and is equal to that of $\mathbf{Z}$ and since $\mathbf{R}$ is real-valued.

3. Thus, the pdf of $\mathbf{Z}^*_0$ is given by:

$$\begin{aligned} & f_{\mathbf{Z}^*_0}\left(\underset{n\times 4p}{\mu^*_0}, \underset{4p\times 4p}{\boldsymbol{\Sigma}^*_0}, \underset{n\times n}{\mathbf{R}} \right) \\ =& (2\pi)^{-2np} \det(\mathbf{R})^{-2p} \det(\boldsymbol{\Sigma}^*_0)^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2} \operatorname{tr}\left[\boldsymbol{\Sigma}^{*-1}_0 \left(\mathbf{Z}^*_0 - \mu^*_0\right)' \mathbf{R}^{-1} \left(\mathbf{Z}^*_0 - \mu^*_0\right) \right] \right\} \end{aligned}$$

and by the isomorphic relations already established above the pdf of $\mathbf{Z}$ follows as

$$f_{\mathbf{Z}}\left(\underset{n\times p}{\mu},\underset{p\times p}{\mathbf{\Sigma}},\underset{n\times n}{\mathbf{R}}\right)=\frac{2^{2np}}{\pi^{2np}\left(\det\mathbf{R}\right)^{2p}\left(\det\mathbf{\Sigma}\right)^{2n}}\exp\left\{-2\,\mathrm{Retr}\left[\mathbf{\Sigma}^{-1}\left(\overline{\mathbf{Z}-\mu}\right)'\mathbf{R}^{-1}\left(\mathbf{Z}-\mu\right)\right]\right\} \tag{3.2.1}$$

from which the desired result follows.

□

The relationship between the cf of the matrix-variate quaternion normal distribution and that of its real associated matrix-variate normal distribution is now established.

Theorem 3.2.2. *Let* $\underset{n\times p}{\mathbf{Z}}\sim\mathbb{Q}N\left(n\times p;\mu,\mathbf{\Sigma}\otimes\mathbf{R}\right)$ *as defined in (3.2.1) above. The cf of* $\mathbf{Z}$ *is given by*

$$\phi_{\mathbf{Z}}(\mathbf{T})=\exp\mathrm{Retr}\left\{\iota\bar{\mu}'\mathbf{T}-\frac{1}{8}\mathbf{\Sigma}\bar{\mathbf{T}}'\mathbf{R}\mathbf{T}\right\}, \tag{3.2.2}$$

where $\underset{n\times p}{\mathbf{T}}\in M_{n\times p}(\mathbb{Q})$ *and where* ι *is the usual complex root.*

Proof. From (2.4.12) and (3.2.1) it follows that

$$\begin{aligned}\phi_{\mathbf{Z}}(\mathbf{T})=&E\left[\exp\iota\,\mathrm{Retr}\left(\bar{\mathbf{Z}}'\mathbf{T}\right)\right]\\ =&\int_B\frac{2^{2np}}{\pi^{2np}\left(\det\mathbf{R}\right)^{2p}\left(\det\mathbf{\Sigma}\right)^{2n}}\\ &\times\exp\left\{-2\,\mathrm{Retr}\left[\mathbf{\Sigma}^{-1}\left(\overline{\mathbf{Z}-\mu}\right)'\mathbf{R}^{-1}\left(\mathbf{Z}-\mu\right)-\frac{\iota}{2}\bar{\mathbf{Z}}'\mathbf{T}\right]\right\}d\mathbf{Z}\end{aligned} \tag{3.2.3}$$

where

$$\begin{aligned}B=\{&\mathbf{Z}=[z_{st}]':z_{st}=x_{1st}+ix_{2st}+jx_{3st}+kx_{4st};\\ &-\infty<x_{1st},x_{2st},x_{3st},x_{4st}<\infty,s=1,\ldots,n,t=1,\ldots,p\}\end{aligned}$$

is the region of variability for $\mathbf{Z}$. The argument of Retr may be rewritten in the form

$$\begin{aligned}&\mathbf{\Sigma}^{-1}\left(\overline{\mathbf{Z}-\mu}\right)'\mathbf{R}^{-1}\left(\mathbf{Z}-\mu\right)-\frac{\iota}{2}\bar{\mathbf{Z}}'\mathbf{T}\\ =&\mathbf{\Sigma}^{-1}\left[\bar{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{Z}-\bar{\mathbf{Z}}'\mathbf{R}^{-1}\mu-\bar{\mu}'\mathbf{R}^{-1}\mathbf{Z}+\bar{\mu}'\mathbf{R}^{-1}\mu-\frac{\iota}{2}\mathbf{\Sigma}\bar{\mathbf{Z}}'\mathbf{T}\right]\\ =&\mathbf{\Sigma}^{-1}\left[\bar{\mathbf{Z}}'\mathbf{R}^{-1}\mathbf{Z}-\bar{\mathbf{Z}}'\mathbf{R}^{-1}\left(\mu+\frac{\iota}{4}\mathbf{R}\mathbf{T}\mathbf{\Sigma}\right)-\left(\bar{\mu}'+\frac{\iota}{4}\mathbf{\Sigma}\bar{\mathbf{T}}'\mathbf{R}\right)\mathbf{R}^{-1}\mathbf{Z}+\bar{\mu}'\mathbf{R}^{-1}\mu\right]\\ =&\mathbf{\Sigma}^{-1}\left[\left(\overline{\mathbf{Z}-\left(\mu+\frac{\iota}{4}\mathbf{R}\mathbf{T}\mathbf{\Sigma}\right)}\right)\mathbf{R}^{-1}\left(\mathbf{Z}-\left(\mu+\frac{\iota}{4}\mathbf{R}\mathbf{T}\mathbf{\Sigma}\right)\right)\right]\\ &+\mathbf{\Sigma}^{-1}\left[\bar{\mu}'\mathbf{R}^{-1}\mu-\left(\overline{\mu+\frac{\iota}{4}\mathbf{R}\mathbf{T}\mathbf{\Sigma}}\right)'\mathbf{R}^{-1}\left(\mu+\frac{\iota}{4}\mathbf{R}\mathbf{T}\mathbf{\Sigma}\right)\right].\end{aligned}$$

Noting the fact that $\mathbf{\Sigma}$ and $\mathbf{R}$ are Hermitian, that the conjugate of Retr is again Retr,and that according to (2.1.2) the arguments of Retr may be rearranged, (3.2.3) reduces to

$$\begin{aligned}\phi_{\mathbf{Z}}(\mathbf{T}) &= \exp \operatorname{Retr}\left\{-2\mathbf{\Sigma}^{-1}\left[\bar{\mu}'\mathbf{R}^{-1}\mu - \left(\overline{\mu + \frac{\iota}{4}\mathbf{RT\Sigma}}\right)'\mathbf{R}^{-1}\left(\mu + \frac{\iota}{4}\mathbf{RT\Sigma}\right)\right]\right\} \\ &= \exp \operatorname{Retr}\left\{-2\mathbf{\Sigma}^{-1}\left[\bar{\mu}'\mathbf{R}^{-1}\mu - \left(\bar{\mu}'\mathbf{R}^{-1} + \frac{\iota}{4}\mathbf{\Sigma}\bar{\mathbf{T}}'\right)\left(\mu + \frac{\iota}{4}\mathbf{RT\Sigma}\right)\right]\right\} \\ &= \exp \operatorname{Retr}\left\{-2\mathbf{\Sigma}^{-1}\left[-\frac{\iota}{4}\bar{\mu}'\mathbf{T\Sigma} - \frac{\iota}{4}\mathbf{\Sigma}\bar{\mathbf{T}}'\mu + \frac{1}{16}\mathbf{\Sigma}\bar{\mathbf{T}}'\mathbf{RT\Sigma}\right]\right\} \\ &= \exp \operatorname{Retr}\left\{\frac{\iota}{2}\bar{\mu}'\mathbf{T} + \frac{\iota}{2}\bar{\mathbf{T}}'\mu - \frac{1}{8}\bar{\mathbf{T}}'\mathbf{RT\Sigma}\right\} \\ &= \exp \operatorname{Retr}\left\{\iota\bar{\mu}'\mathbf{T} - \frac{1}{8}\mathbf{\Sigma}\bar{\mathbf{T}}'\mathbf{RT}\right\}, \quad \text{as required.}\end{aligned}$$

□

Note that

$$\phi_{\mathbf{Z}}(\mathbf{T}) = \exp \operatorname{Retr}\left\{\iota\bar{\mu}'\mathbf{T} - \frac{1}{8}\mathbf{\Sigma}\bar{\mathbf{T}}'\mathbf{RT}\right\} = \exp \operatorname{tr}\left\{\iota\mu_0^{*\prime}\mathbf{T}_0^* - \frac{1}{2}\mathbf{\Sigma}_0^*\mathbf{T}_0^{*\prime}\mathbf{RT}_0^*\right\} = \phi_{\mathbf{Z}_0^*}(\mathbf{T}_0^*)$$

satisfying (2.4.13).

3.3 Summary

In this chapter it was seen that the p-variate quaternion normal distribution and matrix-variate quaternion normal distribution are respectively algebraically equivalent to a $4p$-variate real normal and $n \times 4p$-matrix-variate real normal distribution. In the next chapter the quadratic forms of these distributions will come to the fore.

Chapter 4

The Quaternion χ^2 and Quaternion Wishart Distributions

After the foundation of the quaternion normal distribution is established, it is important to investigate the distributions of various functions in which they may appear. The sums of quadratic forms of quaternion normal variables yielding the quaternion chi-squared and quaternion Wishart distributions are respectively discussed in Sections 4.1 and 4.2.

4.1 The quaternion χ^2 distribution

What is the distribution of $\underline{\bar{Z}}'\underline{Z}$, for $\underset{p\times 1}{\underline{Z}}$ a p-variate quaternion normal distributed probability vector, from the real representation perspective? Rautenbach (1983, p 161) [28] answered this question.

Theorem 4.1.1. *Let* $\underset{p\times 1}{\underline{Z}} \sim \mathbb{Q}N(p; \underline{0}, \mathbf{I}_p)$*. If* $V = \underline{\bar{Z}}'\underline{Z}$*, then* $W = 4V \sim \chi^2_{4p}$ *(chi-squared with* $4p$ *degrees of freedom) distribution.*

Proof.

$$
\begin{aligned}
V =& \bar{\underline{Z}}'\underline{Z} \\
=& [X_{11} - iX_{21} - jX_{31} - kX_{41}, \ldots, X_{1p} - iX_{2p} - jX_{3p} - kX_{4p}] \\
& \times \begin{bmatrix} X_{11} - iX_{21} - jX_{31} - kX_{41} \\ \vdots \\ X_{1p} - iX_{2p} - jX_{3p} - kX_{4p} \end{bmatrix} \\
=& \frac{1}{4}\left[\sum_{s=1}^{p}(2X_{1s})^2 + \sum_{s=1}^{p}(2X_{2s})^2 + \sum_{s=1}^{p}(2X_{3s})^2 + \sum_{s=1}^{p}(2X_{4s})^2\right].
\end{aligned}
$$

The real associated vector of $\underline{Z}$ is

$$\underset{4p\times 1}{\underline{Z}_0^*} = [X_{11}, X_{21}, X_{31}, X_{41}, \ldots, X_{1p}, X_{2p}, X_{3p}, X_{4p}]'$$

with $\underset{4p\times 1}{\underline{Z}_0^*} \sim N(4p; \underline{0}, \frac{1}{4}\mathbf{I}_p)$. It now follows from a known result in real distribution theory that

$$\sum_{s=1}^{p}(2X_{ls})^2 \sim \chi_p^2, \; l = 1, 2, 3, 4.$$

Therefore

$$W = 4V = \sum_{l=1}^{4}\sum_{s=1}^{p}(2X_{1s})^2 \sim \chi_{4p}^2$$

with pdf

$$f_W(w) = 2^{-2p}\left[\Gamma\left(2p\right)\right]^{-1} w^{2p-1} \exp\left(-\frac{1}{2}w\right), \qquad w > 0, \tag{4.1.1}$$

where $\Gamma(\cdot)$ is the real gamma function. □

The following result is often required in applications and inference problems.

Theorem 4.1.2. *Let* $\underline{Z} \sim \mathbb{Q}N\left(p; \underline{\mu}, \mathbf{\Sigma}\right)$ *then* $4\left(\overline{\underline{Z} - \underline{\mu}}\right)' \mathbf{\Sigma}^{-1}\left(\underline{Z} - \underline{\mu}\right) \sim \chi_{4p}^2$. *(See Rautenbach (1983) [28].)*

Proof. Let $\underline{Y} = \mathbf{C}\left(\underline{Z} - \underline{\mu}\right)$ where $\mathbf{C\Sigma\bar{C}'} = \mathbf{I}_p$ with $\mathbf{C}$ a symplectic matrix. From Theorem 3.1.7 it follows that $\underline{Y} \sim \mathbb{Q}N\left(p; \underline{0}, \mathbf{I}_p\right)$ so that

$$\begin{aligned}\underline{\bar{Y}}'\underline{Y} =& \sum_{s=1}^{p} \bar{Y}_s Y_s \\ =& \frac{1}{4}\sum_{s=1}^{p}\left(4Y_{s1}^2 + 4Y_{s2}^2 + 4Y_{s3}^2 + 4Y_{s4}^2\right)\end{aligned}$$

where each $Y_{sr} \sim N(0, \frac{1}{4}), r = 1, 2, 3, 4$ so that $4Y_{sr}^2 \sim \chi_1^2, r = 1, 2, 3, 4$. Thus it follows that

$$\begin{aligned}4\underline{\bar{Y}}'\underline{Y} =& 4\left(\overline{\underline{Z} - \underline{\mu}}\right)' \bar{\mathbf{C}}'\mathbf{C}\left(\underline{Z} - \underline{\mu}\right) \\ =& 4\left(\overline{\underline{Z} - \underline{\mu}}\right)' \mathbf{\Sigma}^{-1}\left(\underline{Z} - \underline{\mu}\right) \\ \sim& \chi_{4p}^2.\end{aligned}$$

□

4.2 The quaternion Wishart distribution

Kabe (1976) [17], (1978) [18], and (1984) [19] derived the hypercomplex Wishart distribution directly from the hypercomplex normal distribution using the $\mathbb{Q}$ generalized Sverdrup's lemma. Teng and Fang (1997) [31] showed that the maximum likelihood estimator $\hat{\mathbf{\Sigma}}$ of $\mathbf{\Sigma}$ followed a quaternion Wishart distribution. They used a Fourier transform on the results given by Andersson (1975) [1] to yield explicit expressions for the probability density and characteristic functions of the quaternion Wishart distribution. The non-central quaternion Wishart distribution was discussed by Kabe (1984) [19], while Li and Xue (2010) [24] derived the singular quaternion Wishart distribution. More technical results, specifically regarding Selberg-type squared matrices, gamma and beta integrals are found in the paper by Gupta and Kabe (2008) [13].

Is it possible to find the density of the quaternion Wishart matrix from the real associated Wishart matrix? In this section, the quaternion Wishart distribution is derived from the real matrix normal distribution associated with the quaternion matrix normal distribution by which it is defined. Once again the emphasis is on the link between the characteristic functions of the quaternion and real associated Wishart distributions.

Theorem 4.2.1. *Let* $\underset{n\times p}{\mathbf{Z}} \sim \mathbb{Q}N\left(n\times p; \mathbf{0}, \mathbf{\Sigma}\otimes\mathbf{I}_n\right)$. *Then for* $n \geq p$, $\mathbf{W} = \bar{\mathbf{Z}}'\mathbf{Z}$ *is said to have the quaternion Wishart distribution with* n *degrees of freedom, i.e.* $\mathbf{W} \sim \mathbb{Q}W_p\left(\mathbf{\Sigma}, n\right)$, *with pdf given by*

$$f(\mathbf{W}) = \frac{2^{2np}}{\mathbb{Q}\Gamma_p(2n)\left(\det \mathbf{\Sigma}\right)^{2n}} \exp\left\{-2\operatorname{Retr}\left(\mathbf{\Sigma}^{-1}\mathbf{W}\right)\right\} \det\left(\mathbf{W}\right)^{2n-2p+1}, \tag{4.2.1}$$

with $\mathbf{W} = \bar{\mathbf{W}}' > \mathbf{0}$ *and where* $\mathbb{Q}\Gamma_p\left(\cdot\right)$ *is the quaternion multivariate gamma function, as given in (A.6.1).*

Proof. Let $\underset{4p\times 4p}{\mathbf{W}_0^*} = \underset{4p\times n}{\mathbf{Z}_0^{*\prime}}\underset{n\times 4p}{\mathbf{Z}_0^*}$ where $\underset{n\times 4p}{\mathbf{Z}_0^*} \sim N\left(n\times 4p; \mathbf{0}, \mathbf{\Sigma}_0^*\otimes\mathbf{I}_n\right)$ is the real associated matrix of $\mathbf{Z}$ as given in Theorem 3.2.1. Let $\mathbf{T} = (t_{ls})$, $l, s = 1, \ldots, p$ where $t_{ls} = \bar{t}_{ls}$ and $t_{ls} = t_{1ls} + it_{2ls} + jt_{3ls} + kt_{4ls}$. From (2.4.13) it follows that

$$\begin{aligned}\phi_{\mathbf{W}}(\mathbf{T}) =& \phi_{\mathbf{W}_0}(\mathbf{T}_0) \\ =& \det\left(\mathbf{I}_{4p} - 2\iota\mathbf{\Sigma}_0^*\mathbf{T}_0^*\right)^{-\frac{n}{2}},\end{aligned}$$

where $\underset{4p\times 4p}{\mathbf{T}_0^*}$ is the real associated symmetric vector of $\underset{p\times p}{\mathbf{T}}$. Let $\mathbf{Y}_0^* = \left(\mathbf{I}_{4p} - 2\iota\mathbf{\Sigma}_0^*\mathbf{T}_0^*\right)$, then $\mathbf{Y}_0^{*\prime} = \left(\mathbf{I}_{4p} - 2\iota\mathbf{T}_0^*\mathbf{\Sigma}_0^*\right) = \mathbf{Y}_0^*$ and from Theorem A.5.7 $\mathbf{Y}$ is Hermitian. Therefore

$$\begin{aligned}\phi_{\mathbf{W}}(\mathbf{T}) &= \det\left(\mathbf{Y}\right)^{-2n} \\ &= \det\left(\mathbf{I}_p - \frac{\iota}{2}\mathbf{\Sigma}\mathbf{T}\right)^{-2n}\end{aligned}$$

Let $\mathbf{S} = -\iota\mathbf{T}$, using (A.6.4) and the definition of the quaternion generalised hypergeometric function of matrix argument as given in Lemma A.6.3, it follows that

$$\begin{aligned}&\det\left(\mathbf{I}_p - \frac{\iota}{2}\mathbf{\Sigma}\mathbf{T}\right)^{-2n} \\ =& \det\left(\frac{1}{2}\mathbf{\Sigma}\right)^{-2n} \det(\mathbf{S})^{-2n}\det\left(\mathbf{I}_p + 2\mathbf{\Sigma}^{-1}\mathbf{S}^{-1}\right)^{-2n} \\ =& \det\left(\frac{1}{2}\mathbf{\Sigma}\right)^{-2n} \det\left(\mathbf{S}\right)^{-2n}\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(2n)_\kappa\mathbb{Q}C_\kappa\left(-2\mathbf{\Sigma}^{-\frac{1}{2}}\mathbf{S}^{-1}\mathbf{\Sigma}^{-\frac{1}{2}}\right)}{k!}\end{aligned}$$

Using the inverse Laplace transformation given in Lemma A.6.9 and (A.6.3), it follows that

$$
\begin{aligned}
f(\mathbf{W}) =& \frac{2^{2p(p-1)}}{(2\pi\iota)^{2p(p-1)+p}} 2^{2np} \det(\mathbf{\Sigma})^{-2n} \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(2n)_\kappa}{k!} \\
& \times \int\limits_{\mathbf{S}-\mathbf{S}_0 \in \Phi} \operatorname{etr}(\mathbf{WS}) \det(\mathbf{S})^{-2n} \mathbb{Q}C_\kappa \left(-2\mathbf{\Sigma}^{-\frac{1}{2}} \mathbf{S}^{-1} \mathbf{\Sigma}^{-\frac{1}{2}}\right) d\mathbf{S} \\
=& 2^{2np} \det(\mathbf{\Sigma})^{-2n} \sum_{k=0}^{\infty} \sum_{\kappa} \frac{(2n)_\kappa}{k! \mathbb{Q}\Gamma_p(2n, \kappa)} \det(\mathbf{W})^{2n-2(p-1)-1} \mathbb{Q}C_\kappa \left(-2\mathbf{W}\mathbf{\Sigma}^{-1}\right) \\
=& \frac{2^{2np}}{\mathbb{Q}\Gamma_p(2n)} \det(\mathbf{\Sigma})^{-2n} \det(\mathbf{W})^{2n-2(p-1)-1} {}_0\mathbb{Q}F_0 \left(-2\mathbf{\Sigma}^{-1}\mathbf{W}\right)
\end{aligned}
$$

from which the desired result follows. □

Remark 4.2.2. From Kabe (1984) [19] it follows that $\mathbf{W}$ has the non-central quaternion Wishart distribution with n degrees of freedom, i.e. $\mathbf{W} \sim \mathbb{Q}W_p(\mathbf{\Sigma}, n, \mathbf{\Omega})$ with pdf given by

$$
f(\mathbf{W}) = \frac{2^{2np} \exp\{-2 \operatorname{Retr}(\mathbf{\Omega})\}}{\mathbb{Q}\Gamma_p(2n) \det(\mathbf{\Sigma})^{2n}} \exp\left\{-2 \operatorname{Retr}\left(\mathbf{\Sigma}^{-1}\mathbf{W}\right)\right\} \det(\mathbf{W})^{2n-2p+1} {}_0\mathbb{Q}F_1 \left(2n, 4\mathbf{\Omega}\mathbf{\Sigma}^{-1}\mathbf{W}\right) \tag{4.2.2}
$$

where $\mathbf{W} = \bar{\mathbf{W}}' > \mathbf{0}$ and where ${}_0\mathbb{Q}F_1(\cdot)$ is the quaternion hypergeometric function with a matrix argument, (see (A.6.2)).

4.3 Summary

In this chapter the quaternion chi-squared and quaternion Wishart distributions were respectively derived using the results from the previous chapter, i.e. the real matrix and multivariate normal distributions associated with their quaternion counterparts. The final chapter will focus on some applications of the study thus far.

Chapter 5

Applications Illustrating the Role of the Quaternion Normal Distribution in Hypothesis Testing

In this chapter two applications of quaternion normal related distributions are presented. Section 5.1 investigates how a hypothesis of the form $H_{01} : \underline{\mu} = \underline{0}$, with $\mathbf{\Sigma}$ known, may be tested while Section 5.2 focusses on the derivation of the probability density function for Wilks's statistic in the quaternion case.

5.1 Quaternion hypothesis testing

Suppose that $\underset{2\times 1}{\underline{Z}} = \underline{X}_1 + i\underline{X}_2 + j\underline{X}_3 + k\underline{X}_4 \sim \mathbb{Q}N\left(2; \underline{\mu}, \mathbf{\Sigma}\right)$ where

$$\underset{2\times 1}{\underline{\mu}} = [\mu_1, \mu_2]' = \underline{\mu}_{X_1} + i\underline{\mu}_{X_2} + j\underline{\mu}_{X_3} + k\underline{\mu}_{X_4}$$

and

$$\underset{2\times 2}{\mathbf{\Sigma}} = \begin{bmatrix} \sigma_1^2 & \xi_{11} + i\xi_{12} + j\xi_{13} + k\xi_{14} \\ \xi_{11} - i\xi_{12} - j\xi_{13} - k\xi_{14} & \sigma_2^2 \end{bmatrix}$$

is a positive definite Hermitian matrix. How may the quaternion null hypothesis

$$H_{01} : \underline{\mu} = \underline{0}, \quad \mathbf{\Sigma} \text{ known}$$

based upon a random sample, $\underline{Z}_1, \ldots, \underline{Z}_n$, against

$$H_{a1} : \underline{\mu} \neq \underline{0}, \quad \boldsymbol{\Sigma} \text{ known},$$

be tested? In order to derive a test criterion for such a test, consider the following likelihood function

$$L = \left(\frac{2}{\pi}\right)^{4n} \left[\sigma_1^2\sigma_2^2 - \left(\xi_{11}^2 + \xi_{12}^2 + \xi_{13}^2 + \xi_{14}^2\right)\right]^{-2n}$$
$$\times \exp\left\{-2\left(\sigma_1^2\sigma_2^2 - \left(\xi_{11}^2 + \xi_{12}^2 + \xi_{13}^2 + \xi_{14}^2\right)\right)^{-1}\left[\sigma_2^2\sum_{s=1}^{n}\left(\overline{z_{1s} - \mu_1}\right)\left(z_{1s} - \mu_1\right)\right.\right.$$
$$\left.\left.-2\,\mathrm{Re}\sum_{s=1}^{n}\left(\overline{z_{1s} - \mu_1}\right)\left(\xi_{11} + i\xi_{12} + j\xi_{13} + k\xi_{14}\right)\left(z_{2s} - \mu_2\right) + \sigma_1^2\sum_{s=1}^{n}\left(\overline{z_{2s} - \mu_2}\right)\left(z_{2s} - \mu_2\right)\right)\right]\right\}.$$

The likelihood ratio criterion is given by

$$\Lambda^* = \frac{\max\limits_{H_{01}} L\left(\underline{\mu}, \boldsymbol{\Sigma}\right)}{\max\limits_{H_{a1}} L\left(\underline{\mu}, \boldsymbol{\Sigma}\right)}.$$

Under H_0 we have

$$\max_{H_{01}} L\left(\underline{\mu}, \boldsymbol{\Sigma}\right) \tag{5.1.1}$$
$$= \left(\frac{2}{\pi}\right)^{4n}\left(\sigma_1^2\sigma_2^2 - \left(\xi_{11}^2 + \xi_{12}^2 + \xi_{13}^2 + \xi_{14}^2\right)\right)^{-2n}$$
$$\times \exp\left\{-2\left(\sigma_1^2\sigma_2^2 - \left(\xi_{11}^2 + \xi_{12}^2 + \xi_{13}^2 + \xi_{14}^2\right)\right)^{-1}\right.$$
$$\left.\times\left[\sigma_2^2\sum_{s=1}^{n}\bar{z}_{1s}z_{1s} - 2\,\mathrm{Re}\sum_{s=1}^{n}\bar{z}_{1s}\left(\xi_{11} + i\xi_{12} + j\xi_{13} + k\xi_{14}\right)z_{2s} + \sigma_1^2\sum_{s=1}^{n}\bar{z}_{2s}z_{2s}\right]\right\}$$

Rautenbach (1983) [28] Theorem 6.3.2 showed that the maximum likelihood estimate of $\underline{\mu}$ is given by

$$\hat{\underline{\mu}} = \frac{1}{n}\sum_{s=1}^{n}\underline{z}_s = \left(\operatorname{avg} z_1, \operatorname{avg} z_2\right)'$$

such that

$$\begin{aligned}
&\max_{H_{a1}} L\left(\underline{\mu}, \mathbf{\Sigma}\right) \\
=&L\left(\hat{\underline{\mu}}, \mathbf{\Sigma}\right) \\
=&\left(\frac{2}{\pi}\right)^{4n}\left(\sigma_1^2\sigma_2^2-\left(\xi_{11}^2+\xi_{12}^2+\xi_{13}^2+\xi_{14}^2\right)\right)^{-2n}\exp\left\{-2\left(\sigma_1^2\sigma_2^2-\left(\xi_{11}^2+\xi_{12}^2+\xi_{13}^2+\xi_{14}^2\right)\right)^{-1}\right. \\
&\times\left[\sigma_2^2\sum_{s=1}^{n}\left(\overline{z_{1s}-\operatorname{avg} z_1}\right)\left(z_{1s}-\operatorname{avg} z_1\right)\right. \\
&-2\operatorname{Re}\sum_{s=1}^{n}\left(\overline{z_{1s}-\operatorname{avg} z_1}\right)\left(\xi_{11}+i\xi_{12}+j\xi_{13}+k\xi_{14}\right)\left(z_{2s}-\operatorname{avg} z_2\right) \\
&\left.\left.+\sigma_1^2\sum_{s=1}^{n}\left(\overline{z_{2s}-\operatorname{avg} z_2}\right)\left(z_{2s}-\operatorname{avg} z_2\right)\right]\right\}
\end{aligned} \tag{5.1.2}$$

From (5.1.1) and (5.1.2) it now follows that

$$\begin{aligned}
\Lambda^* =& \exp\left\{-2n\left(\sigma_1^2\sigma_2^2-\left(\xi_{11}^2+\xi_{12}^2+\xi_{13}^2+\xi_{14}^2\right)^{-1}\left[\sigma_1^2\operatorname{avg}\bar{z}_1\operatorname{avg} z_1\right.\right.\right. \\
&\left.\left.-2\operatorname{Re}\operatorname{avg}\bar{z}_1\left(\xi_{11}+i\xi_{12}+j\xi_{13}+k\xi_{14}\right)\operatorname{avg} z_2+\sigma_2^2\operatorname{avg}\bar{z}_2\operatorname{avg} z_2\right]\right\} \\
=& \exp\left\{-2n\operatorname{avg}\underline{\bar{z}}'\mathbf{\Sigma}^{-1}\operatorname{avg}\underline{z}\right\}.
\end{aligned}$$

The null hypothesis, H_{01}, is now rejected at the $100\alpha\%$ significance level, in favour of H_{a1}, if

$$\exp\left\{-2n\operatorname{avg}\underline{\bar{z}}'\mathbf{\Sigma}^{-1}\operatorname{avg}\underline{z}\right\}\leq\lambda_\alpha^*$$

where the constant λ_α^* is such that $P\left[\Lambda^*\leq\lambda_\alpha^*|H_{01}\right]=\alpha$. Thus, H_{01} is rejected if

$$y=4n\operatorname{avg}\underline{\bar{z}}'\mathbf{\Sigma}^{-1}\operatorname{avg}\underline{z}\geq-2\ln\lambda_\alpha^*=\lambda_\alpha'.$$

Under H_{01} $\operatorname{avg}\underline{z}\sim\mathbb{Q}N\left(2;\underline{0},\frac{1}{n}\mathbf{\Sigma}\right)$ and from Theorem 4.1.2 it follows that $Y=4n\operatorname{avg}\underline{\bar{Z}}'\mathbf{\Sigma}^{-1}\operatorname{avg}\underline{Z}\sim\chi_8^2$. Since $P\left[y\geq\chi_{8,1-\alpha}^2\right]=\alpha$ the null hypothesis is rejected if $y\geq\chi_{8,1-\alpha}^2$ where $\chi_{8,1-\alpha}^2$ is the $100\left(1-\alpha\right)^{th}$ percentile of χ_8^2.

An alternative approach in deriving a test criterion in this case involves the use of the real associated probability vector of $\underline{z}$. It is known that $\underset{8\times1}{\underline{z}_0^*}\sim N\left(8;\underline{\mu}_0^*,\mathbf{\Sigma}_0^*\right)$, i.e. a real multivariate normal distribution, where $\underset{8\times1}{\underline{\mu}_0^*}=\left[\mu_{11},\mu_{21},\mu_{31},\mu_{41},\mu_{12},\mu_{22},\mu_{23},\mu_{24}\right]'$

and

$$\mathbf{\Sigma}_0^* = \frac{1}{4}\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 & \xi_{11} & -\xi_{12} & -\xi_{13} & -\xi_{14} \\ 0 & \sigma_1^2 & 0 & 0 & \xi_{12} & \xi_{11} & -\xi_{14} & \xi_{13} \\ 0 & 0 & \sigma_1^2 & 0 & \xi_{13} & \xi_{14} & \xi_{11} & -\xi_{12} \\ 0 & 0 & 0 & \sigma_1^2 & \xi_{14} & -\xi_{13} & \xi_{12} & \xi_{11} \\ \xi_{11} & \xi_{12} & \xi_{13} & \xi_{14} & \sigma_2^2 & 0 & 0 & 0 \\ -\xi_{12} & \xi_{11} & \xi_{14} & -\xi_{13} & 0 & \sigma_2^2 & 0 & 0 \\ -\xi_{13} & -\xi_{14} & \xi_{11} & \xi_{12} & 0 & 0 & \sigma_2^2 & 0 \\ -\xi_{14} & \xi_{13} & -\xi_{12} & \xi_{11} & 0 & 0 & 0 & \sigma_2^2 \end{bmatrix}.$$

The test criterion for testing $H_{01}^* : \underline{\mu}_0^* = \underline{0}$, $\mathbf{\Sigma}_0^*$ known, against $H_{a1}^* : \underline{\mu}_0^* \neq \underline{0}$, $\mathbf{\Sigma}_0^*$ known, based upon a random sample $\underline{Z}_{01}^*, \ldots, \underline{Z}_{0n}^*$ of $\underline{Z}_0^*$ is given by

$$Y_0 = n\left(\text{avg}\, \underline{Z}_0^{*\prime} \mathbf{\Sigma}_0^{*-1} \text{avg}\, \underline{Z}_0^*\right) \sim \chi_8^2$$

where $\text{avg}\, \underline{Z}_0^* = [\text{avg}\, X_{11}, \text{avg}\, X_{21}, \text{avg}\, X_{31}, \text{avg}\, X_{41}, \text{avg}\, X_{12}, \text{avg}\, X_{22}, \text{avg}\, X_{32}, \text{avg}\, X_{42}]'$ such that the null hypothesis is rejected if $y_0 \geq \chi_{8,1-\alpha}^2$ where $\chi_{8,1-\alpha}^2$ is the $100(1-\alpha)^{th}$ percentile of χ_8^2.

From the above discussion it is once again clear that two different approaches exist in order to test $H_{01} : \underline{\mu} = \underline{0}$ against $H_{a1} : \underline{\mu} \neq \underline{0}$ with $\mathbf{\Sigma}$ known. We may either conduct an analysis using quaternion quantities directly with y or, by utilising the real associated quantity, y_0, as test criterion respectively. From this it is clear that the quaternion hypothesis $H_{01} : \underline{\mu} = \underline{0}$ against $H_{a1} : \underline{\mu} \neq \underline{0}$ may also be expressed in terms of real quantities, i.e. $H_{01}^* : \underline{\mu}_0^* = \underline{0}$ against $H_{a1}^* : \underline{\mu}_0^* \neq \underline{0}$.

5.2 Quaternion matrix-variate beta type I and Wilks's statistic

Let the rows of $\underset{n_1 \times p}{\mathbf{X}}$ and $\underset{n_2 \times p}{\mathbf{Y}}$ be independently $\mathbb{Q}N(p; \underline{0}, \mathbf{\Sigma})$ and $\mathbb{Q}N(p; \underline{\mu}, \mathbf{\Sigma})$ distributed, respectively. From Theorem 3.2.1 we know that $\mathbf{X} \sim \mathbb{Q}N(n_1 \times p, \mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{I}_{n_1})$ and $\mathbf{Y} \sim \mathbb{Q}N(n_2 \times p; \mu, \mathbf{\Sigma} \otimes \mathbf{I}_{n_2})$, and from Theorem 4.2.1 it furthermore follows that $\mathbf{X}'\mathbf{X} \sim \mathbb{Q}W_p(\mathbf{\Sigma}, n_1)$ and $\mathbf{Y}'\mathbf{Y} \sim \mathbb{Q}W_p(\mathbf{\Sigma}, n_2, \mathbf{\Omega})$ (i.e. is the non-central quaternion Wishart distribution, see Kabe (1984) [19]), as was discussed in Remark 4.2.2 with $\mathbf{\Omega} = \mathbf{\Sigma}^{-1}\mu\bar{\mu}'$. Wilks's statistic

$$\Lambda = \frac{\det\left(\mathbf{X}'\mathbf{X}\right)}{\det\left(\mathbf{X}'\mathbf{X} + \mathbf{Y}'\mathbf{Y}\right)}$$

can be used as a likelihood ratio criterion for testing whether the matrix mean μ is equal to zero or not.

The solution of this problem depends on the derivation of the distribution of Λ, and to this end, we consider the quaternion matrix-variate beta type I distribution.

Various references to the quaternion matrix-variate beta family of distributions are available in the literature. Dumitriu and Koev (2008) [8], for instance, derived explicit expressions for the distributions of the extreme eigenvalues of the quaternion Jacobi ensemble, while Gupta and Kabe (2008) [13] and Díaz-García and Gutiérrez Jáimez (2009) [7] explored further properties of this family. Li and Xue (2010) [24] dealt with singular quaternion matrix-variate beta distributions. In Theorems 5.2.1 and 5.2.2, that now follow, the central and non-central quaternion matrix-variate beta type I distributions are respectively discussed.

Theorem 5.2.1. *Let* $\mathbf{A} \sim \mathbb{Q}W_p\left(\mathbf{\Sigma}, n_1\right)$, $(n_1 \geq p)$ *and* $\mathbf{B} \sim \mathbb{Q}W_p\left(\mathbf{\Sigma}, n_2\right)$, $(n_2 \geq p)$ *be independent, then* $\mathbf{U} = \left(\mathbf{A}+\mathbf{B}\right)^{-\frac{1}{2}}\mathbf{A}\left(\mathbf{A}+\mathbf{B}\right)^{-\frac{1}{2}}$ *is said to have the quaternion matrix-variate beta type I distribution with* n_1 *and* n_2 *degrees of freedom, i.e.* $\mathbf{U} \sim \mathbb{Q}B1(m, n_1, n_2)$ *with pdf given by*

$$f(\mathbf{U}) = \frac{\det\left(\mathbf{U}\right)^{2n_1-2p+1}\det\left(\mathbf{I}_p - \mathbf{U}\right)^{2n_2-2p+1}}{\mathbb{Q}B_p(2n_1, 2n_2)}, \quad \mathbf{0} < \mathbf{U} = \bar{\mathbf{U}}' < \mathbf{I}_p,$$

where $\mathbb{Q}B_p(\cdot)$ *is given in Lemma A.6.2.*

Proof. From (4.2.1) the joint pdf of $\mathbf{A}$ and $\mathbf{B}$ is given by

$$f(\mathbf{A}, \mathbf{B}) = \frac{\det\left(\frac{1}{2}\mathbf{\Sigma}\right)^{-2n_1-2n_2}}{\mathbb{Q}\Gamma_p(2n_1)\mathbb{Q}\Gamma_p(2n_2)}\det\left(\mathbf{A}\right)^{2n_1-2p+1}\det\left(\mathbf{B}\right)^{2n_2-2p+1}\exp\left\{-2\,\mathrm{Retr}\left(\mathbf{\Sigma}^{-1}(\mathbf{A}+\mathbf{B})\right)\right\}.$$

Let $\mathbf{V} = \mathbf{A} + \mathbf{B}$, then $\mathbf{U} = \mathbf{V}^{-\frac{1}{2}}\mathbf{A}\mathbf{V}^{-\frac{1}{2}}$ whose Jacobian is given by $J\left(\mathbf{A}, \mathbf{B} \rightarrow \mathbf{V}, \mathbf{U}\right) =$

$\det(\mathbf{V})^{2p-1}$ (see Lemma A.6.6). Therefore

$$f(\mathbf{U}) = \frac{\det\left(\frac{1}{2}\boldsymbol{\Sigma}\right)^{-2n_1-2n_2}}{\mathbb{Q}\Gamma_p(2n_1)\mathbb{Q}\Gamma_p(2n_2)} \int\limits_{\mathbf{V}=\bar{\mathbf{V}}'>\mathbf{0}} \det\left(\mathbf{V}^{\frac{1}{2}}\mathbf{U}\mathbf{V}^{\frac{1}{2}}\right)^{2n_1-2p+1} \det\left(\mathbf{V}-\mathbf{V}^{\frac{1}{2}}\mathbf{U}\mathbf{V}^{\frac{1}{2}}\right)^{2n_2-2p+1}$$
$$\times \exp\left\{-2\,\mathrm{Retr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{V}\right)\right\}\det(\mathbf{V})^{2p-1}\,\mathrm{d}\mathbf{V}$$
$$= \frac{\det\left(\frac{1}{2}\boldsymbol{\Sigma}\right)^{-2n_1-2n_2}}{\mathbb{Q}\Gamma_p(2n_1)\mathbb{Q}\Gamma_p(2n_2)} \det(\mathbf{U})^{2n_1-2p+1}\det(\mathbf{I}_p-\mathbf{U})^{2n_2-2p+1}$$
$$\times \int\limits_{\mathbf{V}=\bar{\mathbf{V}}'>\mathbf{0}} \det(\mathbf{V})^{2n_1+2n_2-2p+1}\exp\left\{-2\,\mathrm{Retr}\left(\boldsymbol{\Sigma}^{-1}\mathbf{V}\right)\right\}\mathrm{d}\mathbf{V}$$
$$= \frac{\mathbb{Q}\Gamma_p(2(n_1+n_2))}{\mathbb{Q}\Gamma_p(2n_1)\mathbb{Q}\Gamma_p(2n_2)} \det(\mathbf{U})^{2n_1-2p+1}\det(\mathbf{I}_p-\mathbf{U})^{2n_2-2p+1} \quad \text{(see Lemma A.6.11).}$$

□

Theorem 5.2.2. *Let* $\mathbf{A} \sim \mathbb{Q}W_p(\boldsymbol{\Sigma}, n_1)$, $(n_1 \geq p)$ *and* $\mathbf{B} \sim \mathbb{Q}W_p(\boldsymbol{\Sigma}, n_2, \boldsymbol{\Omega})$, $(n_2 \geq p)$ *be independent, then*

$$\mathbf{U} = \left(\mathbf{I}_p + \mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}\right)^{-\frac{1}{2}} \mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}} \left(\mathbf{I}_p + \mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}\right)^{-\frac{1}{2}}, \tag{5.2.1}$$

or equivalently, $\mathbf{U} = (\mathbf{A}+\mathbf{B})^{-\frac{1}{2}}\mathbf{A}(\mathbf{A}+\mathbf{B})^{-\frac{1}{2}}$ *is said to have the non-central quaternion matrix-variate beta type I distribution with* n_1 *and* n_2 *degrees of freedom, i.e.* $\mathbf{U} \sim \mathbb{Q}B1(p, n_1, n_2, \boldsymbol{\Omega})$ *with pdf given by*

$$f(\mathbf{U}) = \frac{\exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\}\det(\mathbf{U})^{2n_1-2p+1}\det(\mathbf{I}_p-\mathbf{U})^{2n_2-2p+1}}{\mathbb{Q}B_p(2n_1, 2n_2)}$$
$$\times\, {}_1\mathbb{Q}F_1\left(2(n_1+n_2), 2n_2, 2\boldsymbol{\Omega}(\mathbf{I}_p-\mathbf{U})\right), \quad \mathbf{0} < \mathbf{U} = \bar{\mathbf{U}}' < \mathbf{I}_p,$$

where $\mathbb{Q}B_p(\cdot)$ *is defined in Lemma A.6.2, and* $n_1 > (p-1)$, $n_2 > (p-1)$ *and* ${}_1\mathbb{Q}F_1(\cdot)$ *is the quaternion confluent hypergeometric function with a matrix argument (see Lemma A.6.3).*

Proof. From (4.2.1) and (4.2.2) the joint pdf of $(\mathbf{A}, \mathbf{B})$ is given by

$$K\exp\left\{-2\,\mathrm{Retr}\left(\boldsymbol{\Sigma}^{-1}(\mathbf{A}+\mathbf{B})\right)\right\}\det(\mathbf{A})^{2n_1-2p+1}\det(\mathbf{B})^{2n_2-2p+1}$$
$$\times \exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\}\,{}_0\mathbb{Q}F_1\left(2n_2; 4\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\mathbf{B}\right) \tag{5.2.2}$$

where $K^{-1} = \mathbb{Q}\Gamma_p(2n_1)\,\mathbb{Q}\Gamma_p(2n_2)\det\left(\frac{1}{2}\mathbf{\Sigma}\right)^{2n_1+2n_2}$. The transformation

$$\mathbf{V} = \mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}$$

has Jacobian (see Lemma A.6.6).

$$J(\mathbf{A},\mathbf{B} \to \mathbf{V},\mathbf{B}) = \det(\mathbf{B})^{2p-1}$$

By substituting this into (5.2.2), and by integrating over $\mathbf{B}$ the pdf of $\mathbf{V}$ is obtained as

$$\begin{aligned} f(\mathbf{V}) =& K\exp\{\mathrm{Retr}(-2\mathbf{\Omega})\}\det(\mathbf{V})^{2n_1-2p+1} \\ &\times \int\limits_{\mathbf{B}=\bar{\mathbf{B}}'>\mathbf{0}} \exp\left\{-2\,\mathrm{Retr}\left(-\mathbf{\Sigma}^{-1}\mathbf{B}^{\frac{1}{2}}(\mathbf{I}_p+\mathbf{V})\mathbf{B}^{\frac{1}{2}}\right)\right\}\det(\mathbf{B})^{2(n_1+n_2)-2p+1} \\ &\times {}_0\mathbb{Q}F_1\left(2n_2; 4\mathbf{\Omega}\mathbf{\Sigma}^{-1}\mathbf{B}\right)\mathrm{d}\mathbf{B}. \end{aligned} \tag{5.2.3}$$

Now, making the transformation $\mathbf{V} \to \mathbf{H}\mathbf{V}\bar{\mathbf{H}}'$, where $\underset{p\times p}{\mathbf{H}}$ is symplectic, and by substituting this into (5.2.3) gives

$$\begin{aligned} f\left(\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right) =& K\exp\{\mathrm{Retr}(-2\mathbf{\Omega})\}\det\left(\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)^{2n_1-2p+1} \\ &\times \int\limits_{\mathbf{B}=\bar{\mathbf{B}}'>\mathbf{0}} \det(\mathbf{B})^{2(n_1+n_2)-2p+1}\exp\left\{-2\,\mathrm{Retr}\left(-\mathbf{\Sigma}^{-1}\mathbf{B}^{\frac{1}{2}}\left(\mathbf{I}_p+\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)\mathbf{B}^{\frac{1}{2}}\right)\right\} \\ &\times {}_0\mathbb{Q}F_1\left(2n_2; 4\mathbf{\Omega}\mathbf{\Sigma}^{-1}\mathbf{B}\right)\mathrm{d}\mathbf{B}. \end{aligned} \tag{5.2.4}$$

Consider the symmetrised pdf , see Appendix A.7, of $\mathbf{V}$, i.e.

$$f_s(\mathbf{V}) = \int\limits_{O(p)} f\left(\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)\mathrm{d}\mathbf{H},$$

where $\mathbf{H} \in O(p)$, $\mathrm{d}\mathbf{H}$ is the normalised invariant measure on $O(p)$ and

$$O(p) = \left\{\mathbf{A} \in M_p(\mathbb{Q})\,\middle|\,\bar{\mathbf{A}}'\mathbf{A} = \mathbf{A}\bar{\mathbf{A}}' = \mathbf{I}_p\right\}.$$

Hence, from (5.2.4), Lemma A.6.5 and Lemma A.6.10, it follows that

$$
\begin{aligned}
& f_s(\mathbf{V}) \\
=& K \exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\} \det(\mathbf{V})^{2n_1-2p+1} \\
& \times \int\limits_{\mathbf{B}=\bar{\mathbf{B}}'>\mathbf{0}} \det(\mathbf{B})^{2(n_1+n_2)-2p+1} {}_0\mathbb{Q}F_1\left(2n_2; 4\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\mathbf{B}\right) \\
& \times \int\limits_{O(p)} \exp\left\{-2\,\mathrm{Retr}\left(-\boldsymbol{\Sigma}^{-1}\mathbf{B}^{\frac{1}{2}}\left(\mathbf{I}_p+\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)\mathbf{B}^{\frac{1}{2}}\right)\right\} d\mathbf{H}d\mathbf{B} \\
=& K \exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\} \det(\mathbf{V})^{2n_1-2p+1} \\
& \times \int\limits_{\mathbf{B}=\bar{\mathbf{B}}'>\mathbf{0}} \det(\mathbf{B})^{2(n_1+n_2)-2p+1} {}_0\mathbb{Q}F_1\left(2n_2; 4\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\mathbf{B}\right) \\
& \times \int\limits_{O(p)} \exp\left\{-2\,\mathrm{Retr}\left(\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{B}\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)\bar{\mathbf{H}}'\right)\right\} d\mathbf{H}d\mathbf{B} \\
=& K \exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\} \det(\mathbf{V})^{2n_1-2p+1} \\
& \times \int\limits_{O(p)} \int\limits_{\mathbf{B}=\bar{\mathbf{B}}'>\mathbf{0}} \det(\mathbf{B})^{2(n_1+n_2)-2p+1} {}_0\mathbb{Q}F_1\left(2n_2; 4\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\mathbf{B}\right) \\
& \times \exp\left\{-2\,\mathrm{Retr}\left(-\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)\bar{\mathbf{H}}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{B}\right)\right\} d\mathbf{B}d\mathbf{H} \\
=& K \exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\} \det(\mathbf{V})^{2n_1-2p+1} \mathbb{Q}\Gamma_p\left(2\left(n_1+n_2\right)\right) \\
& \times \int\limits_{O(p)} \det\left(2\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)\bar{\mathbf{H}}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\right)^{-2(n_1+n_2)} \\
& \times {}_1\mathbb{Q}F_1\left(2\left(n_1+n_2\right); 2n_2; 2\boldsymbol{\Sigma}^{\frac{1}{2}}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)^{-1}\bar{\mathbf{H}}'\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-1}\right) d\mathbf{H} \\
=& \left\{\mathbb{Q}B_p\left(2n_1, 2n_2\right)\right\}^{-1} \exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\} \det(\mathbf{V})^{2n_1-2p+1} \det\left(\mathbf{I}_p+\mathbf{V}\right)^{-2(n_1+n_2)} \\
& \times \int\limits_{O(p)} {}_1\mathbb{Q}F_1\left(2\left(n_1+n_2\right); 2n_2; 2\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Omega}\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)^{-1}\bar{\mathbf{H}}'\right) d\mathbf{H}.
\end{aligned} \tag{5.2.5}
$$

Since $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}\mu\bar{\mu}'$, it follows from Lemma A.6.3 and Lemma A.6.4 that the integral in

(5.2.5) can be rewritten as

$$\begin{aligned}
&\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(2(n_1+n_2))_\kappa}{(2n_2)_\kappa}\frac{1}{k!}\int\limits_{O(p)}\mathbb{Q}C_\kappa\left[2\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Sigma}^{-1}\mu\bar{\mu}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)^{-1}\bar{\mathbf{H}}'\right]\mathrm{d}\mathbf{H}\\
&=\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(2(n_1+n_2))_\kappa}{(2n_2)_\kappa}\frac{1}{k!}\frac{\mathbb{Q}C_\kappa\left(2\boldsymbol{\Sigma}^{-\frac{1}{2}}\mu\bar{\mu}'\boldsymbol{\Sigma}^{-\frac{1}{2}}\right)\mathbb{Q}C_\kappa\left[\left(\mathbf{I}_p+\mathbf{V}\right)^{-1}\right]}{\mathbb{Q}C_\kappa\left(\mathbf{I}_p\right)}\\
&=\sum_{k=0}^{\infty}\sum_{\kappa}\frac{(2(n_1+n_2))_\kappa}{(2n_2)_\kappa}\frac{1}{k!}\int\limits_{O(p)}\mathbb{Q}C_\kappa\left[2\boldsymbol{\Omega}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)^{-1}\bar{\mathbf{H}}'\right]\mathrm{d}\mathbf{H}\\
&=\int\limits_{O(p)}{}_1\mathbb{Q}F_1\left(2(n_1+n_2);2n_2;2\boldsymbol{\Omega}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)^{-1}\bar{\mathbf{H}}'\right)\mathrm{d}\mathbf{H}.
\end{aligned}\tag{5.2.6}$$

Substituting (5.2.6) into (5.2.5) yields

$$\begin{aligned}
f_s(\mathbf{V})=&\left\{\mathbb{Q}B_p(2n_1,2n_2)\right\}^{-1}\exp\left\{\mathrm{Retr}\left(-2\boldsymbol{\Omega}\right)\right\}\det(\mathbf{V})^{2n_1-2p+1}\\
&\times\det\left(\mathbf{I}_p+\mathbf{V}\right)^{-2(n_1+n_2)}\\
&\times\int\limits_{O(p)}{}_1\mathbb{Q}F_1\left(2(n_1+n_2);2n_2;2\boldsymbol{\Omega}\mathbf{H}\left(\mathbf{I}_p+\mathbf{V}\right)^{-1}\bar{\mathbf{H}}'\right)\mathrm{d}\mathbf{H}.
\end{aligned}\tag{5.2.7}$$

Since $f_s(\mathbf{V})=\int\limits_{O(p)} f\left(\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)d\mathbf{H}$ it follows from (5.2.7) that

$$\begin{aligned}
f\left(\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)=&\left\{\mathbb{Q}B_p(2n_1,2n_2)\right\}^{-1}\exp\left\{\mathrm{Retr}\left(-2\boldsymbol{\Omega}\right)\right\}\det\left(\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)^{2n_1-2p+1}\\
&\times\det\left(\mathbf{I}_p+\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)^{-2(n_1+n_2)}\\
&\times{}_1\mathbb{Q}F_1\left(2(n_1+n_2);2n_2;2\boldsymbol{\Omega}\left(\mathbf{I}_p+\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\right)^{-1}\right)\mathrm{d}\mathbf{H}.
\end{aligned}$$

The transformation $\mathbf{H}\mathbf{V}\bar{\mathbf{H}}'\to\mathbf{V}$ yields the pdf of $\mathbf{V}$ as

$$\begin{aligned}
&\left\{\mathbb{Q}B_p(2n_1,2n_2)\right\}^{-1}\exp\left\{\mathrm{Retr}\left(-2\boldsymbol{\Omega}\right)\right\}\det(\mathbf{V})^{2n_1-2p+1}\\
&\times\det\left(\mathbf{I}_p+\mathbf{V}\right)^{-2(n_1+n_2)}{}_1\mathbb{Q}F_1\left(2(n_1+n_2);2n_2;2\left(\mathbf{I}_p+\mathbf{V}\right)^{-1}\boldsymbol{\Omega}\right).
\end{aligned}\tag{5.2.8}$$

Consider the transformation in (5.2.1) written in terms of $\mathbf{V}$, i.e.

$$\begin{aligned}
\mathbf{U}&=\left(\mathbf{I}_p+\mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}\right)^{-\frac{1}{2}}\mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}\left(\mathbf{I}_p+\mathbf{B}^{-\frac{1}{2}}\mathbf{A}\mathbf{B}^{-\frac{1}{2}}\right)^{-\frac{1}{2}}\\
&=\left(\mathbf{I}_p+\mathbf{V}\right)^{-\frac{1}{2}}\mathbf{V}\left(\mathbf{I}_p+\mathbf{V}\right)^{-\frac{1}{2}}.
\end{aligned}$$

Since $\mathbf{V}$ commutes with any rationale function, it follows that

$$\begin{aligned}\mathbf{U} &= (\mathbf{I}_p + \mathbf{V})^{-\frac{1}{2}} \mathbf{V} (\mathbf{I}_p + \mathbf{V})^{-\frac{1}{2}} \\ &= (\mathbf{I}_p + \mathbf{V})^{-1} \mathbf{V}\end{aligned}$$

whose Jacobian is given by

$$J(\mathbf{V} \to \mathbf{U}) = \det(\mathbf{I}_p - \mathbf{U})^{-2(2p-1)}$$

and $\mathbf{V} = \mathbf{U}(\mathbf{I}_p - \mathbf{U})^{-1}$. By substituting this into (5.2.8) yields

$$\begin{aligned}f(\mathbf{U}) &= \{\mathbb{Q}B_p(2n_1, 2n_2)\}^{-1} \exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\} \det(\mathbf{I}_p - \mathbf{U})^{-2(2p-1)} \\ &\quad \times \det\left((\mathbf{I}_p - \mathbf{U})^{-\frac{1}{2}} \mathbf{U} (\mathbf{I}_p - \mathbf{U})^{-\frac{1}{2}}\right)^{2n_1-2p+1} \\ &\quad \times \det\left(\mathbf{I}_p + (\mathbf{I}_p - \mathbf{U})^{-\frac{1}{2}} \mathbf{U} (\mathbf{I}_p - \mathbf{U})^{-\frac{1}{2}}\right)^{-2(n_1+n_2)} \\ &\quad \times {}_1\mathbb{Q}F_1\left(2(n_1+n_2); 2n_2; 2\left[\mathbf{I}_p + (\mathbf{I}_p - \mathbf{U})^{-\frac{1}{2}} \mathbf{U} (\mathbf{I}_p - \mathbf{U})^{-\frac{1}{2}}\right]^{-1} \boldsymbol{\Omega}\right) \\ &= \{\mathbb{Q}B_p(2n_1, 2n_2)\}^{-1} \exp\{\mathrm{Retr}(-2\boldsymbol{\Omega})\} \det(\mathbf{U})^{2n_1-2p+1} \\ &\quad \times \det(\mathbf{I}_p - \mathbf{U})^{2n_2-2p+1} {}_1\mathbb{Q}F_1(2(n_1+n_2); 2n_2; 2(\mathbf{I}_p - \mathbf{U})\boldsymbol{\Omega}).\end{aligned}$$

□

What is the corresponding density expression for Wilks's statistic in the case of quaternion Wishart matrices? Mehta (2004) [27] provides many useful results for quaternion random matrices, for instance, on p282 the pdf for the determinant of a $n \times n$ random Hermitian matrix taken from the Gaussian unitary ensemble is calculated.

Let $\underset{p\times p}{\mathbf{A}}$ and $\underset{p\times p}{\mathbf{B}}$ be two independent quaternion Wishart matrices, i.e. $\mathbf{A} \sim \mathbb{Q}W_p(n_1, \boldsymbol{\Sigma})$ and $\mathbf{B} \sim \mathbb{Q}W_p(n_2, \boldsymbol{\Sigma}, \boldsymbol{\Omega})$, and $n_1, n_2 \geq p$. What is the exact expression for the density function of

$$\Lambda = \frac{\det(\mathbf{A})}{\det(\mathbf{A} + \mathbf{B})} = \det(\mathbf{U})?$$

The present work proposes the distribution of Wilks's statistic based on Meijer's G-function (see Mathai (1993) [25]) in a numerical feasible form. Since Λ is real according to Theorem A.5.12 (also see Mehta (2004, p284) [27]), the result follows similarly as that given in Bekker, Roux and Arashi (2010) [4].

From Theorem 5.2.2 and Lemma A.6.3

$$
\begin{aligned}
&E\left[\det\left(\mathbf{U}\right)^{h-1}\right]\\
=&\frac{\exp\left\{\mathrm{Retr}(-2\mathbf{\Omega})\right\}}{\mathbb{Q}B_p(2n_1,2n_2)}\\
&\times\int\limits_{\mathbf{0}<\mathbf{U}=\bar{\mathbf{U}}'<\mathbf{I}_p}\det\left(\mathbf{U}\right)^{h+2n_1-2p}\det\left(\mathbf{I}_p-\mathbf{U}\right)^{2n_2-2p+1}{}_1\mathbb{Q}F_1\left(2(n_1+n_2);2n_2;2\mathbf{\Omega}(\mathbf{I}_p-\mathbf{U})\right)\mathrm{d}\mathbf{U}\\
=&\frac{\exp\left\{\mathrm{Retr}(-2\mathbf{\Omega})\right\}}{\mathbb{Q}B_p(2n_1,2n_2)}\sum_{k=0}^{\infty}\sum_{\kappa}\frac{\left(2(n_1+n_2)\right)_\kappa}{k!\left(2n_2\right)_\kappa}\\
&\times\int\limits_{\mathbf{0}<\mathbf{U}=\bar{\mathbf{U}}'<\mathbf{I}_p}\det\left(\mathbf{U}\right)^{h+2n_1-2p}\det\left(\mathbf{I}_p-\mathbf{U}\right)^{2n_2-2p+1}\mathbb{Q}C_\kappa\left(2\mathbf{\Omega}(\mathbf{I}_p-\mathbf{U})\right)\mathrm{d}\mathbf{U}
\end{aligned}
$$

Let $\mathbf{T}=(\mathbf{I}_p-\mathbf{U})$, after applying Lemma A.6.12 to the above expression (using $\mathbb{Q}\Gamma_p\left(a,\kappa\right)=\left(a\right)_\kappa\mathbb{Q}\Gamma_p\left(a\right)$), and then simplifying, we obtain

$$
\begin{aligned}
&E\left[\det\left(\mathbf{U}\right)^{h-1}\right]\\
=&\frac{\exp\left\{\mathrm{Retr}(-2\mathbf{\Omega})\right\}}{\mathbb{Q}\Gamma_p(2n_1)}\sum_{k=0}^{\infty}\sum_{\kappa}\frac{\mathbb{Q}\Gamma_p(2(n_1+n_2),\kappa)\mathbb{Q}\Gamma_p(2n_1+h-1)}{k!\mathbb{Q}\Gamma_p(2(n_1+n_2)+h-1,\kappa)}\mathbb{Q}C_\kappa\left(2\mathbf{\Omega}\right)
\end{aligned}
$$

From Lemma A.6.1 it follows that

$$
\begin{aligned}
\mathbb{Q}\Gamma_p\left(2n_1+h-1\right)=&\pi^{p(p-1)}\prod_{\alpha=1}^{p}\Gamma(2n_1+h-1-2(\alpha-1))\\
\mathbb{Q}\Gamma_p\left(2\left(n_1+n_2\right)+h-1,\kappa\right)=&\pi^{p(p-1)}\prod_{\alpha=1}^{p}\Gamma\left(2\left(n_1+n_2\right)+h-1+k_\alpha-2(\alpha-1)\right).
\end{aligned}
$$

Therefore the pdf of $\Lambda=\det(\mathbf{U})$ is uniquely determined by the inverse Mellin transform as given in Mathai (1993, Definition 1.8, p23) [25] (see Appendix A.8). Thus

$$
\begin{aligned}
&f_\Lambda\left(\lambda\right)\\
=&\frac{\exp\left\{\mathrm{Retr}(-2\mathbf{\Omega})\right\}}{\mathbb{Q}\Gamma_p(2n_1)}\sum_{k=0}^{\infty}\sum_{\kappa}\mathbb{Q}C_\kappa\left(2\mathbf{\Omega}\right)\frac{\mathbb{Q}\Gamma_p(2(n_1+n_2),\kappa)}{k!}G_{p,p}^{p,0}\left(\lambda\,\middle|\,\begin{matrix}a_1,\ldots,a_p\\ b_1,\ldots,b_p\end{matrix}\right)
\end{aligned}
$$

where $a_\alpha=2(n_1+n_2-\alpha)+k_\alpha+1$, $\alpha=1,\ldots,p$ and $b_\alpha=2n_1-2\alpha+1$, $\alpha=1,\ldots,p$, and where $G(\cdot)$ denotes Meijer's G-function as given in Appendix A.8.

If $\mathbf{\Omega} = \mathbf{0}$, using Theorem 5.2.1 the distribution of Wilks's statistic, under the null hypothesis, is given by

$$f_{\Lambda}(\lambda) = \frac{\mathbb{Q}\Gamma_p(2(n_1+n_2))}{\mathbb{Q}\Gamma_p(2n_1)} G_{p,p}^{p,0}\left(\lambda \,\middle|\, \begin{matrix} a_1,\ldots,a_p \\ b_1,\ldots,b_p \end{matrix}\right)$$

where $a_\alpha = 2(n_1+n_2-\alpha)+1$, $\alpha = 1,\ldots,p$ and $b_\alpha = 2n_1 - 2\alpha + 1$, $\alpha = 1,\ldots,p$.

Remark 5.2.3. Hotelling's T^2 statistic is given by $T^2 = n\underline{Y}'\mathbf{A}^{-1}\underline{Y}$ where $\underset{p\times 1}{\underline{Y}} \sim \mathbb{Q}N(p;\underline{0},\mathbf{\Sigma})$ and which is independently distributed of $\underset{p\times p}{\mathbf{A}} \sim \mathbb{Q}W_p(\mathbf{\Sigma}, n)$. From Theorem A.5.13 this is equal to $n\underline{Y}_0^{*\prime}\mathbf{A}_0^{-1}\underline{Y}_0$ where $\underset{4p\times 1}{\underline{Y}_0^*} \sim N(4p;\underline{0},\mathbf{\Sigma}_0)$ (from Definition 3.1.1) and which is independently distributed of $\underset{4p\times 4p}{\mathbf{A}_0} \sim W_{4p}(\mathbf{\Sigma}_0, n)$ (from Theorem 4.2.1). Once again, this problem reduces to a problem in the real space, and familiar techniques and inference procedures in the real distribution theory may be applied.

5.3 Summary

In this chapter we saw that a quaternion hypothesis reduces to a real hypothesis and that one may either conduct the testing procedure by working with the quaternion quantities directly, or conduct the testing procedure by working with the real associated counterparts thereof; yielding similar results in either case.

In the final section the central and non-central quaternion matrix-variate beta type I distributions were explored in order to assist in the derivation of the pdf for Wilks's statistic in the quaternion space.

Chapter 6

Conclusions

This final chapter concludes with a summary of the objectives met in this work and suggests possible avenues for future pursuits.

6.1 Summary of Conclusions

The contributions of several prominent researchers on quaternions in distribution theory such as the papers of Kabe (1976) [17], (1978) [18] and (1984) [19], Rautenbach (1983) [28], and Teng and Fang (1997) [31] were highlighted, all of whom made use of the representation theory.

It was shown how the p-variate quaternion normal distribution, along with its real associated counterpart, forms the basis from which the quaternion distribution theory may be further explored. For the first time the quaternion Wishart distribution was derived from the real associated Wishart via the characteristic function.

Finally, two applications were presented, illustrating the role of the quaternion normal distribution in hypothesis testing.

6.2 Future Work

The fact that a quaternion pdf may be algebraically equivalent to its real associated pdf poses interesting possibilities, specifically in the area of computation and simmulation.

Rautenbach (1983) [28] devoted an entire chapter on hypothesis testing and inference procedures in the quaternionic space. Many of these ideas may be expanded to the matrix-variate cases, i.e. tests involving Wishart matrices.

There appear to be a gap in the literature regarding Bayesian analysis involving quaternions. The real representation approach may in this case, shed an interesting light on the relationships between quaternion prior and posterior density functions, and may lead to new types of loss functions.

The possibilities in applications in quantum mechanics were covered in quite some detail in Rautenbach (1983) [28], however many other areas, such as quantum finance (see Baaquie (2004) [2]), and rotational problems within molecular modelling (see Karney (2007) [20]), may benefit equally well from the subsequent development of the quaternion distribution theory.

Hopefully this dissertation will stimulate further research in this area...

Bibliography

[1] S. Andersson. Invariant normal models. *The Annals of Statistics*, 3(1):132–154, 1975.

[2] B.E. Baaquie. *Quantum finance: path integrals and Hamiltonians for options and interest rates.* Cambridge Univercity Press, 2004.

[3] L.J. Bain and M. Engelhardt. *Introduction to Probability and Mathematical Statistics.* Duxbury, Mass, 1992.

[4] A. Bekker, J.J.J. Roux, and M. Arashi. Exact nonnull distribution of Wilks's statistic: Ratio and product of independent components. *Accepted in Journal of Multivariate Analysis, 10.1016/j.jmva.2010.11.005*, 2010.

[5] J.A. Díaz-García. Special functions: Integral properties of jack polynomials, hypergeometric functions and invariant polynomials. *ArXiv:0909.1988v1 [math.ST]*, 2009.

[6] J.A. Díaz-García and R. Gutiérrez-Jáimez. Singular noncentral and doubly noncentral matrix variate beta distribution. *Comunicación Técnica, No. I-06-#(PE/CIMAT), Guanajuato, México.*

[7] J.A. Díaz-García and R. Gutiérrez Jáimez. Random matrix theory and multivariate statistics. *ArXiv:0907.1064v1 [math.ST]*, 2009.

[8] I. Dumitriu and P. Koev. Distributions of the extreme eigenvalues of beta-jacobi random matrices. *SIAM J. Matrix Anal. Appl*, 30:1–6, 2008.

[9] R. Ehlers, A. Bekker, and J.J.J. Roux. The central and non-central matrix variate Dirichlet type III distribution. *South African Statistical Journal*, 43(2):97–116, 2009.

[10] N.R. Goodman. Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction). *Annals of Mathematical Statistics*, 34:152–177, 1963.

[11] M.J. Greenacre. Symmetrized multivariate distributions. *South African Statistical Journal*, 7:95–101, 1973.

[12] K.I. Gross and D.S.P. Richards. Special functions of matrix argument. i: Algebraic induction, zonal polynomials, and hypergeometric functions. *Transactions of the American Mathematical Society*, 301(2):781–811, 1987.

[13] A.K. Gupta and D.G. Kabe. Selberg-type squared matrices gamma and beta integrals. *European Journal of Pure and Applied Mathematics*, 1(1):197–201, 2008.

[14] A.K. Gupta and D.K. Nagar. *Contributions to complex matrix variate distributions theory*. Universidad de Antioquia, 2009.

[15] H. Halberstam and R.E. Ingram. *The Mathematical Papers of Sir William Rowan Hamilton*, volume III: Algebra. Camebridge: University Press, 1967.

[16] A.J. Hanson. *Visualizing quaternions*. Morgan Kaufmann Pub, 2006.

[17] D.G. Kabe. Q generalized Sverdrup's lemma. *The Canadian Journal of Statistics/La Revue Canadienne de Statistique*, 4(2):221–226, 1976.

[18] D.G. Kabe. On some inequalities satisfied by beta and gamma functions. *Suid-Afrikaanse Statistiese Tydskrif*, 12:25–31, 1978.

[19] D.G. Kabe. Classical statistical analysis based on a certain hypercomplex multivariate normal distribution. *Metrika*, 31(1):63–76, 1984.

[20] C.F.F. Karney. Quaternions in molecular modeling. *Journal of Molecular Graphics and Modelling*, 25(5):595–604, 2007.

[21] C.G. Khatri. Classical statistical analysis based on a certain multivariate complex gaussian distribution. *The Annals of Mathematical Statistics*, 36(1):98–114, 1965.

[22] M. Kline. *Mathematical Thought from Ancient to Modern Times.* Oxford University Press, New York, 1972.

[23] F. Li and Y. Xue. Zonal polynomials and hypergeometric functions of quaternion matrix argument. *Communications in Statistics-Theory and Methods*, 38(8):1184–1206, 2009.

[24] F. Li and Y. Xue. The density functions of the singular quaternion normal matrix and the singular quaternion Wishart matrix. *Communications in Statistics-Theory and Methods*, 39(18):3316–3331, 2010.

[25] A.M. Mathai. *A handbook of generalized special functions for statistical and physical sciences.* Oxford University Press, USA, 1993.

[26] M.L. Mehta. *Random Matrices and the Statistical Theory of Energy Levels.* New York: Academic Press, 1967.

[27] M.L. Mehta. *Random matrices.* New York: Academic Press, 2004.

[28] H.M. Rautenbach. *Kwaternione – Ontwikkeling met Spesiale Verwysing na die Kwaternioonnormaalverdeling.* Doctoral thesis, University of South Africa, University of South Africa, Department of Statistics, South Africa, June 1983. Unpublished doctoral thesis.

[29] H.M Rautenbach and J.J.J. Roux. Statistical analysis based on quaternion normal random variables. Technical Report 85/9, University of South Africa, Pretoria, 1985.

[30] K. Rektorys. Survey of applicable mathematics. 1969.

[31] C.Y. Teng and K.T. Fang. Statistical analysis based on normal distribution of quaternion. *1997 International Symposium on Contemporary Multivariate Analysis and Its Applications, May 1997.*

[32] B.L. van der Waerden. *Modern Algebra, Vol II.* F. Ungar, New York, 1950.

[33] F. Zhang. Quaternions and matrices of quaternions. *Linear algebra and its applications*, 251:21–57, 1997.

Appendix A

Some Useful Mathematical Results

A.1 Introduction

In this Appendix some useful mathematical results are given that are assumed throughout this work.

In Section A.2 the representation theory is discussed, i.e. the replacement of the elements of an abstract division algebra by matrices. In order to facilitate the construction of a representation that maps division algebra elements into matrices, it is necessary to know when sets of matrices vorm a devision algebra.

After the discussion of some fundemental mathematical concepts it is shown that the complex (Section A.3), as well as the quaternion numbers (Section A.4), an associative division algebra over the real number field form. This result is then used to show how these numbers are represented by matrices with specific structures. Most of the results in these sections are a summary of that given by Rautenbach (1983) [28].

The algebraic results given in Section A.5 as discussed in Rautenbach (1983) [28] and are used in the derivation of the p-variate quaternion normal distribution.

Some mathematical functions and polynomials are given in Section A.6 that are used in the derivation of additional results in the quaternion distribution theory, while Sections A.7 and A.8 respectively discuss the symmetrised density function (as presented by Greenacre (1973) [11]), and the Mellin transform along with Meijer's G-function.

A.2 Division algebra representation

Definition A.2.1. A ring is a set $\mathcal{R}$ with two binary operations, addition and multiplication such that $\mathbb{R}$ is closed over them and satisfies:

1. $(a+b)+c = a+(b+c)\ \forall a,b,c \in \mathcal{R}$.
2. $\exists 0 \in \mathcal{R} : a+0 = 0+a = a\ \forall a \in \mathcal{R}$.
3. $\forall a \in \mathcal{R} \exists -a \in \mathcal{R} : a+(-a) = 0$.
4. $a+b = b+a\ \forall a,b \in \mathcal{R}$.
5. $(ab)c = a(bc)\ \forall a,b,c \in \mathcal{R}$.
6. $a(b+c) = ab+ac\ \forall a,b,c \in \mathcal{R}$.
7. $(a+b)c = ac+bc\ \forall a,b,c \in \mathcal{R}$.

(See Rektorys (1969) [30] as cited in Rautenbach (1983) [28].)

Definition A.2.2. A ring $\mathcal{R}$ is called commutative if it satisfies:

$$ab = ba\ \forall \mathbb{R}$$

(See Rektorys (1969) [30] as cited in Rautenbach (1983) [28].)

Definition A.2.3. if

$$\exists\ 1 \in \mathcal{R} : a \times 1 = 1 \times a = a\ \forall a \in \mathcal{R}$$

then $\mathcal{R}$ is called a unit ring (or a ring with identity). (See Rektorys (1969) [30] as cited in Rautenbach (1983) [28].)

Definition A.2.4. If

$$\exists\ a^{-1} \in \mathcal{R} : aa^{-1} = a^{-1}a = 1, 1, a \in \mathcal{R}$$

(where 1 is the unit element defined in $\mathcal{R}$), then $\mathcal{R}$ is called a division ring. Moreover, if a division ring is commutative, it is called a field. (See Rektorys (1969) [30] as cited in Rautenbach (1983) [28].)

Definition A.2.5. Let S be a set and $\mathcal{R}$ a division ring. S is a collective operator (scalar) of $\mathcal{R}$ if:

1. $a + b)\alpha = a\alpha + b\alpha \ \forall \alpha \in S, a, b \in \mathcal{R}$.
2. $(ab)\alpha = a(b\alpha) = (a\alpha)b \ \forall \alpha \in S, a, b \in \mathcal{R}$.

(See Rautenbach (1983) [28].)

Definition A.2.6. A division algebra over the field $\mathbb{P}$ is a division ring $\mathcal{R}$, with the field $\mathbb{P}$ as ring operator. (See Rautenbach (1983) [28].)

Remark A.2.7.

1. The above mentioned definitions follow similarly in the case where S is a set of scalars of $\mathcal{R}$ in such a way that the scalar multiplication is defined from the left.
2. The set of real numbers $\mathbb{R}$ together with the usual operations of addition and multiplication, and with scalar multiplication, forms a division algebra over the field of real numbers.
3. The set of $n \times n$ matrices with real entries, $M_n(\mathbb{R})$, forms a non-commutative division ring, with the usual matrix addition and multiplication. This set furthermore forms a division algebra over $\mathbb{R}$.

Definition A.2.8. Let D and D^* be two division algebras over $\mathbb{R}$, the field of real numbers. A function $f : D \to D^*$ is then called a homomorphism if:

1. $f(a + b) = f(a) + f(b) \ \forall a, b \in \mathbb{R}$.
2. $f(ab) = f(a)f(b) \ \forall a, b \in \mathbb{R}$.
3. $f(a\alpha) = f(a)\alpha \ \forall \alpha \in \mathbb{R}, \ \forall a \in \mathbb{R}$.

(See van der Wearden (1950) [32] as cited in Rautenbach (1983) [28].)

Definition A.2.9. A Homomorphism $f : D \to D^*$ is a monomorphism if:

$$f(a) \neq f(b) \implies a \neq b.$$

(See Rautenbach (1983) [28].)

Definition A.2.10. A homomorphism $f : D \to D^*$ is a epimorphism if:

$$\exists\ a^* \in D^* \exists!\ a \in D : f(a) = a^*.$$

(See Rautenbach (1983) [28].)

Definition A.2.11. A homomorphism $f : D \to D^*$ which is both a monomorphism and a epimorphism, is called an isomorphism. If there exists an isomorphism from D unto D^* then it is said that D and D^* are isomorphic and is denoted by $D \simeq D^*$. (See Rautenbach (1983) [28].)

Definition A.2.12. Let G be a arbitrary division algebra over a field of all real numbers and $M_n(\mathbb{R})$ a matrix division algebra over a field of real numbers. A homomorphism $f : G \to M_n(\mathbb{R})$ is called a representation of G. (See Rautenbach (1983) [28].)

Definition A.2.13. A representation which is an isomorphism, is called a faithful representation. (See Rautenbach (1983) [28].)

Remark A.2.14. Hence, a faithful representation of G replaces the elements of G with $n \times n$ matrices, and the operations with matrix operations. It can therefore be thought of as a type of function between a particular division algebra, and a matrix division algebra. The main advantage of representations lie in the fact that operations with matrices can easily be carried out. (See Rautenbach (1983) [28].)

A.3 Complex numbers

In this section it is shown that every complex number can be represented by a 2×2 matrix with a specific structure.

Definition A.3.1. The set $\mathbb{C}$ of complex numbers consists of elements of the form:

$$c = a + ib$$

where $a, b \in \mathbb{R}$ and $i^2 = -1$. a and b are called the real and imaginary parts, of the complex number c, respectively. (See Rektorys (1969) [30] as cited in Rautenbach (1983) [28].)

Theorem A.3.2. *Let*

$$C = \{a + ib | a, b \in \mathbb{R}\}$$

and let the operations be defined as:

1. $(a + ib) + (c + id) = (a + c) + i(b + d)\ \forall a, b, c, d \in \mathbb{R}.$

2. $(a + ib)(c + id) = (ac - bd) + i(bc + ad)\ \forall a, b, c, d \in \mathbb{R}.$

3. $[(a + ib) + (c + id)]\alpha = (a + ib)\alpha + (c + id)\alpha\ \forall a, b, c, d \in \mathbb{R},\ \forall \alpha \in \mathbb{R}.$

4.
$$\begin{aligned} &[(a + ib)(c + id)]\alpha \\ =&(a + ib)[(c + id)\alpha] \\ =&[(a + ib)\alpha](c + id) \\ &\forall a, b, c, d \in \mathbb{R},\ \forall \alpha \in \mathbb{R}. \end{aligned}$$

The set $\mathbb{C}$ *together with these operations form a division algebra over the field of real numbers. (See Rautenbach (1983) [28].)*

Proof. Let $a + ib, c + id, e + if \in \mathbb{C}$, then it follows that:

1. $(a + ib) + (c + id) = a + c + i(b + d) \in \mathbb{C}.$

2. $(a + ib)(c + id) = ac - bd + i(bc + ad) \in \mathbb{C}.$

3. $[(a + ib) + (c + id)] + (e + if) = (a + ib) + [(c + id) + (e + if)].$

4. $\exists\ 0 + i0 \in \mathbb{C} : (a + ib) + (0 + i0) = (0 + i0) + (a + ib) = a + ib.$

5. $\forall a + ib \in \mathbb{C} \exists!\ -a + i(-b) \in \mathbb{C} : (a + ib) + (-a + i(-b)) = 0 + i0.$

6. $(a + ib) + (c + id) = (c + id) + (a + ib).$

7. $[(a + ib)(c + id)](e + if) = (a + ib)[(c + id)(e + if)].$

8. $(a + ib)[(c + id) + (e + if)] = (a + ib)(c + id) + (a + ib)(e + if).$

9. $[(a+ib)+(c+id)](e+if) = (a+ib)(e+if)+(c+id)(e+if)$.

10. $\exists!\ 1+i0 \in \mathbb{C} : (a+ib)(1+i0) = (1+i0)(a+ib) = (a+ib)\ \forall a+ib \in \mathbb{C}$.

11. $\forall c = a+ib \in \mathbb{C} \exists!\ \frac{a-ib}{a^2+b^2} \in \mathbb{C}$, such that:

$$\begin{aligned} &(a+ib)\,\frac{(a-ib)}{a^2+b^2} \\ =&\frac{(a-ib)}{a^2+b^2}\,(a+ib) \\ =&1+i0, \end{aligned}$$

where $a-ib$ is called the conjugate and is denoted by $\bar{c} = a-ib$.

12. $\forall \alpha \in \mathbb{R}$ it follows that:

$$[(a+ib)+(c+id)]\alpha = (a+ib)\alpha+(c+id)\alpha$$

and

$$\begin{aligned} &[(a+ib)(c+id)]\alpha \\ =&(a+ib)[(c+id)\alpha] \\ =&[(a+ib)\alpha](c+id). \end{aligned}$$

□

Theorem A.3.3. *Let*

$$M_2(\mathbb{R}) = \left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} | a, b \in \mathbb{R} \right\}$$

and let the following operations be defined:

1. $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} a+c & -b-d \\ b+d & a+c \end{bmatrix}\ \forall a, b, c, d \in \mathbb{R}$.

2. $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac-bd & -bc-ad \\ bc+ad & ac-bd \end{bmatrix}\ \forall a, b, c, d \in \mathbb{R}$.

3. $\left\{\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix}\right\}\alpha = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\alpha + \begin{bmatrix} c & -d \\ d & c \end{bmatrix}\alpha, \ \forall a, b, c, d \in \mathbb{R}, \alpha \in \mathbb{R}.$

4.

$$\begin{aligned}\left\{\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} c & -d \\ d & c \end{bmatrix}\right\}\alpha &= \begin{bmatrix} a & -b \\ b & a \end{bmatrix}\left\{\begin{bmatrix} c & -d \\ d & c \end{bmatrix}\alpha\right\} \\ &= \left\{\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\alpha\right\}\begin{bmatrix} c & -d \\ d & c \end{bmatrix} \\ &\forall a, b, c, d \in \mathbb{R}, \alpha \in \mathbb{R}.\end{aligned}$$

The set $M_2(\mathbb{R})$ together with these operations form a division algebra over the field of real numbers. (See Rautenbach (1983) [28].)

Proof. Let $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}, \begin{bmatrix} c & -d \\ d: & c \end{bmatrix}, \begin{bmatrix} e & -f \\ f & e \end{bmatrix} \in M_2(\mathbb{R})$, then it follows that:

1. $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} a+c & -b-d \\ b+d & a+c \end{bmatrix} \in M_2(\mathbb{R})$.

2. $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} ac-bd & -bc-ad \\ bc+ad & ac-bd \end{bmatrix} \in M_2(\mathbb{R})$.

3.

$$\begin{aligned}&\left\{\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix}\right\} + \begin{bmatrix} e & -f \\ f & e \end{bmatrix} \\ &= \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \left\{\begin{bmatrix} c & -d \\ d & c \end{bmatrix} + \begin{bmatrix} e & -f \\ f & e \end{bmatrix}\right\}.\end{aligned}$$

4. The set $M_2(\mathbb{R})$ has a element $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ such that:

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

5. For each element $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \in M_2(\mathbb{R})$ there exists an element $\begin{bmatrix} -a & b \\ -b & -a \end{bmatrix}$ in $M_2(\mathbb{R})$ such that:

$$\begin{bmatrix} -a & b \\ -b & -a \end{bmatrix} + \begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} -a & b \\ -b & -a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

6.

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} = \begin{bmatrix} c & -d \\ d & c \end{bmatrix} + \begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

7.

$$\left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \right\} \begin{bmatrix} e & -f \\ f & e \end{bmatrix}$$
$$= \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \left\{ \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \begin{bmatrix} e & -f \\ f & e \end{bmatrix} \right\}$$

8.

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \left\{ \begin{bmatrix} c & -d \\ d & c \end{bmatrix} + \begin{bmatrix} e & -f \\ f & e \end{bmatrix} \right\}$$
$$= \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} c & -d \\ d & c \end{bmatrix} + \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} e & -f \\ f & e \end{bmatrix}$$

9.

$$\left\{ \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \right\} \begin{bmatrix} e & -f \\ f & e \end{bmatrix}$$
$$= \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} e & -f \\ f & e \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \begin{bmatrix} e & -f \\ f & e \end{bmatrix}$$

10. The set $M_2(\mathbb{R})$ has an element $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ such that:

$$\begin{aligned} &\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ =&\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \\ =&\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \end{aligned}$$

11. For each $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \in M_2(\mathbb{R})$ there exists an element $\begin{bmatrix} \frac{a}{a^2+b^2} & \frac{b}{a^2+b^2} \\ \frac{-b}{a^2+b^2} & \frac{a}{a^2+b^2} \end{bmatrix} \in M_2(\mathbb{R})$ such that:

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} \frac{a}{a^2+b^2} & \frac{b}{a^2+b^2} \\ \frac{-b}{a^2+b^2} & \frac{a}{a^2+b^2} \end{bmatrix} = \begin{bmatrix} \frac{a}{a^2+b^2} & \frac{b}{a^2+b^2} \\ \frac{-b}{a^2+b^2} & \frac{a}{a^2+b^2} \end{bmatrix}\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

12. $\forall \alpha \in \mathbb{R}$ it follows that:

$$\left\{\begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix}\right\}\alpha = \begin{bmatrix} a\alpha & -b\alpha \\ b\alpha & a\alpha \end{bmatrix} + \begin{bmatrix} c\alpha & -d\alpha \\ d\alpha & c\alpha \end{bmatrix}$$

and

$$\begin{aligned} &\left\{\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} c & -d \\ d & c \end{bmatrix}\right\}\alpha \\ =&\begin{bmatrix} a & -b \\ b & a \end{bmatrix}\begin{bmatrix} c\alpha & -d\alpha \\ d\alpha & c\alpha \end{bmatrix} \\ =&\begin{bmatrix} a\alpha & -b\alpha \\ b\alpha & a\alpha \end{bmatrix}\begin{bmatrix} c & -d \\ d & c \end{bmatrix} \end{aligned}$$

□

Remark A.3.4. In a similar fashion as in Theorem A.3.3 it can be shown that the set $M_2'(\mathbb{R})$:

$$\left\{\begin{bmatrix} a & b \\ -b & a \end{bmatrix} | a, b \in \mathbb{R}\right\}$$

together with the operations:

1.
$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} + \begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \begin{bmatrix} a+c & b+d \\ -b-d & a+c \end{bmatrix}$$

2.
$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} c & d \\ -d & c \end{bmatrix} = \begin{bmatrix} ac-bd & ad+bc \\ -ad-bc & ac-bd \end{bmatrix}$$

3.
$$\left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} + \begin{bmatrix} c & d \\ -d & a \end{bmatrix} \right\} \alpha$$
$$= \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \alpha + \begin{bmatrix} c & d \\ -d & c \end{bmatrix} \alpha$$

and

$$\left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \begin{bmatrix} c & d \\ -d & c \end{bmatrix} \right\} \alpha$$
$$= \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \left\{ \begin{bmatrix} c & d \\ -d & c \end{bmatrix} \alpha \right\}$$
$$= \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \alpha \right\} \begin{bmatrix} c & d \\ -d & c \end{bmatrix}$$

also form a division algebra over the field of real numbers.

Definition A.3.5. Define a mapping $f : \mathbb{C} \to M_2(\mathbb{R})$ as:

$$f(a+ib) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \quad \forall a+bi \in \mathbb{C}$$

Theorem A.3.6. *f is a faithful representation. (See Rautenbach (1983) [28].)*

Proof. 1. The mapping $f : \mathbb{C} \to M_2(\mathbb{R})$ is a homomorphism since:

(a)

$$\begin{aligned} f\{(a+ib)+(c+id)\} =& f\{(a+c)+i(b+d)\} \\ =& \begin{bmatrix} a+c & -b-d \\ b+d & a+c \end{bmatrix} \\ =& \begin{bmatrix} a & -b \\ b & a \end{bmatrix} + \begin{bmatrix} c & -d \\ d & c \end{bmatrix} \\ =& f(a+ib)+f(c+id). \end{aligned}$$

(b)

$$\begin{aligned} f\{(a+ib)(c+id)\} =& f\{(ac-bd)+i(bc+ad)\} \\ =& \begin{bmatrix} ac-bd & -bc-ad \\ bc+ad & ac-bd \end{bmatrix} \\ =& \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} b & -d \\ d & b \end{bmatrix} \\ =& f(a+ib)f(c+id), \quad \text{and} \end{aligned}$$

(c)

$$\begin{aligned} & f\{(a+ib)\alpha\} \\ =& f(a\alpha+ib\alpha) \\ =& \begin{bmatrix} a\alpha & -b\alpha \\ b\alpha & a\alpha \end{bmatrix} \\ =& \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \alpha \\ =& f(a+ib)\alpha. \end{aligned}$$

This proves that f is a homomorphism.

2. Furthermore f is a monomorphism since if

$$f(a+id) \neq f(c+id)$$

then

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \neq \begin{bmatrix} c & -d \\ d & c \end{bmatrix}.$$

3. Finally, it follows that for each $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} \in M_2(\mathbb{R})$ there exists a $a + ib \in (C)$ such that:

$$f(a+ib) = \begin{bmatrix} a & -b \\ b & a \end{bmatrix}.$$

4. Thus it follows that the mapping $f : \mathbb{C} \to M_2(\mathbb{R})$ is an isomorphism.

5. It can therefore be concluded that f is a faithful representation.

□

Remark A.3.7. 1. Consequently, the representation f now substitutes the elements of $\mathbb{C}$ with 2×2 matrices with a specific structure of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ and the operations are substituted by matrix operations. From this it follows that:

$$0 + i \to \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

and

$$1 + i0 \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

which is called the base matrices of the complex numbers.

2. In a similar fashion as in Theorem A.3.6 it can be shown that a representation f' can be constructed such that $f' : \mathbb{C} \to M_2(\mathbb{R})$ is a faithful representation. Hence, the representation f' now substitutes the elements of $\mathbb{C}$ with 2×2 matrices with a specific structure, i.e. of the form $\begin{bmatrix} a & b \\ -b & a \end{bmatrix}$.

A.4 Quaternions

In this section, a similar approach is followed to that employed in Section A.3 in order to show how the quaternions may be represented by 4×4 matrices with specific structures.

Definition A.4.1. The set $\mathbb{Q}$ of quaternions consists of elements of the form:

$$q = a_1 + ia_2 + ja_3 + ka_4$$

where $a_1, a_2, a_3, a_4 \in \mathbb{R}$ and:

$$i^2 = j^2 = k^2 = -1$$

and

$$\begin{aligned} ij &= k & jk &= i & ki &= j \\ ji &= -k & kj &= -i & ik &= -j \end{aligned}$$

(See Halberstam and Ingram (1967) [15] as cited in Rautenbach (1983) [28].)

Remark A.4.2. From Definition A.4.1 the following multiplication table can easily be constructed:

Table A.1: Quaternion multiplication table illustrating the relationships between i, j and k.

	1	i	j	k
1	1	i	j	k
i	i	-1	k	$-j$
j	j	$-k$	-1	i
k	k	j	$-i$	-1

Theorem A.4.3. *Let*

$$\mathbb{Q} = \{a_1 + ia_2 + ja_3 + ka_4 | a_1, a_2, a_3, a_4 \in \mathbb{R}\}$$

and define the following operations:

1.

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4) + (b_1 + ib_2 + jb_3 + kb_4)\\ =&(a_1 + b_1) + i(a_2 + b_2) + j(a_3 + b_3) + k(a_4 + b_4)\end{aligned}$$

2.

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4)(b_1 + ib_2 + jb_3 + kb_4)\\ =&(a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4) + i(a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)\\ &+ j(a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2) + k(a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1)\end{aligned}$$

3.

$$\begin{aligned}&[(a_1 + ia_2 + ja_3 + ka_4) + (b_1 + ib_2 + jb_3 + kb_4)]\alpha\\ =&(a_1 + ia_2 + ja_3 + ka_4)\alpha + (b_1 + ib_2 + jb_3 + kb_4)\alpha\end{aligned}$$

4.

$$\begin{aligned}&[(a_1 + ia_2 + ja_3 + ka_4)(b_1 + ib_2 + jb_3 + kb_4)]\alpha\\ =&(a_1 + ia_2 + ja_3 + ka_4)[(b_1 + ib_2 + jb_3 + kb_4)\alpha]\\ =&[(a_1 + ia_2 + ja_3 + ka_4)\alpha](b_1 + ib_2 + jb_3 + kb_4)\end{aligned}$$

$\forall a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4 \in \mathbb{R}$ *and* $\alpha \in \mathbb{R}$. *The set* $\mathbb{Q}$ *together with these operations form a division algebra over the field of real numbers. (See Rautenbach (1983) [28].)*

Proof. Let $a_1 + ia_2 + ja_3 + ka_4$, $b_1 + ib_2 + jb_3 + kb_4$ and $c_1 + ic_2 + jc_3 + kc_4 \in \mathbb{Q}$, then it follows that:

1.

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4) + (b_1 + ib_2 + jb_3 + kb_4)\\ =&(a_1 + b_1) + i(a_2 + b_2) + j(a_3 + b_3) + k(a_4 + b_4) \in \mathbb{Q}.\end{aligned}$$

2.

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4)(b_1 + ib_2 + jb_3 + kb_4)\\=&(a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4) + i(a_1b_2 + a_2b_1 + a_3b_4 - a_4b_3)\\&+ j(a_1b_3 - a_2b_4 + a_3b_1 + a_4b_2) + k(a_1b_4 + a_2b_3 - a_3b_2 + a_4b_1) \in \mathbb{Q}.\end{aligned}$$

3.

$$\begin{aligned}&[(a_1 + ia_2 + ja_3 + ka_4) + (b_1 + ib_2 + jb_3 + kb_4)] + (c_1 + ic_2 + jc_3 + kc_4)\\=&(a_1 + ia_2 + ja_3 + ka_4) + [(b_1 + ib_2 + jb_3 + kb_4) + (c_1 + ic_2 + jc_3 + kc_4)].\end{aligned}$$

4. The set $\mathbb{Q}$ has an element $0 + i0 + j0 + k0$ such that:

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4) + (0 + i0 + j0 + k0)\\=&(0 + i0 + j0 + k0) + (a_1 + ia_2 + ja_3 + ka_4)\\=&(a_1 + ia_2 + ja_3 + ka_4).\end{aligned}$$

5. For every $a_1 + ia_2 + ja_3 + ka_4 \in \mathbb{Q}$ there exists an element $-a_1 + i(-a_2) + j(-a_3) + k(-a_4) \in \mathbb{Q}$ such that:

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4) + (-a_1 + i(-a_2) + j(-a_3) + k(-a_4))\\=&0 + i0 + j0 + k0.\end{aligned}$$

6.

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4) + (b_1 + ib_2 + jb_3 + kb_4)\\=&(b_1 + ib_2 + jb_3 + kb_4) + (a_1 + ia_2 + ja_3 + ka_4).\end{aligned}$$

7.

$$\begin{aligned}&[(a_1 + ia_2 + ja_3 + ka_4)(b_1 + ib_2 + jb_3 + kb_4)](c_1 + ic_2 + jc_3 + kc_4)\\=&(a_1 + ia_2 + ja_3 + ka_4)[(b_1 + ib_2 + jb_3 + kb_4)(c_1 + ic_2 + jc_3 + kc_4)].\end{aligned}$$

8.

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4)[(b_1 + ib_2 + jb_3 + kb_4) + (c_1 + ic_2 + jc_3 + kc_4)]\\=&(a_1 + ia_2 + ja_3 + ka_4)(b_1 + ib_2 + jb_3 + kb_4)\\&+ (a_1 + ia_2 + ja_3 + ka_4)(c_1 + ic_2 + jc_3 + kc_4).\end{aligned}$$

9.

$$\begin{aligned}&[(a_1 + ia_2 + ja_3 + ka_4) + (b_1 + ib_2 + jb_3 + kb_4)](c_1 + ic_2 + jc_3 + kc_4)\\=&(a_1 + ia_2 + ja_3 + ka_4)(c_1 + ic_2 + jc_3 + kc_4)\\&+ (b_1 + ib_2 + jb_3 + kb_4)(c_1 + ic_2 + jc_3 + kc_4).\end{aligned}$$

10. The set $\mathbb{Q}$ has an element $1 + i0 + j0 + k0$ such that:

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4)(1 + i0 + j0 + k0)\\=&(1 + i0 + j0 + k0)(a_1 + ia_2 + ja_3 + ka_4)\\=&a_1 + ia_2 + ja_3 + ka_4 \ \forall a_1 + ia_2 + ja_3 + ka_4 \in \mathbb{Q}.\end{aligned}$$

11. For every $q = a_1 + ia_2 + ja_3 + ka_4 \in \mathbb{Q}$ there exists an element

$$\frac{a_1 - ia_2 - ja_3 - ka_4}{a_1^2 + a_2^2 + a_3^2 + a_4^2} \in \mathbb{Q}$$

such that:

$$\begin{aligned}&(a_1 + ia_2 + ja_3 + ka_4)\frac{(a_1 - ia_2 - ja_3 - ka_4)}{a_1^2 + a_2^2 + a_3^2 + a_4^2}\\=&\frac{(a_1 - ia_2 - ja_3 - ka_4)}{a_1^2 + a_2^2 + a_3^2 + a_4^2}(a_1 + ia_2 + ja_3 + ka_4)\\=&1 + i0 + j0 + k0\end{aligned}$$

where $a_1 - ia_2 - ja_3 - ka_4$ is called the conjugate and is denoted by $\bar{q} = a_1 - ia_2 - ja_3 - ka_4$.

12. For every $\alpha \in \mathbb{R}$ it follows that:

$$\begin{aligned}&[(a_1 + ia_2 + ja_3 + ka_4) + (b_1 + ib_2 + jb_3 + kb_4)]\alpha\\=&(a_1 + ia_2 + ja_3 + ka_4)\alpha + (b_1 + ib_2 + jb_3 + kb_4)\alpha\end{aligned}$$

and

$$\begin{aligned}&[(a_1 + ia_2 + ja_3 + ka_4)(b_1 + ib_2 + jb_3 + kb_4)]\alpha\\ =&(a_1 + ia_2 + ja_3 + ka_4)[(b_1 + ib_2 + jb_3 + kb_4)\alpha]\\ =&[(a_1 + ia_2 + ja_3 + ka_4)\alpha](b_1 + ib_2 + jb_3 + kb_4).\end{aligned}$$

□

Theorem A.4.4. *Let*

$$M_4(\mathbb{R}) = \left\{ \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} | a_1, a_2, a_3, a_4 \in \mathbb{R} \right\}$$

and define the following operations:

1.

$$\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} + \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}$$

$$= \begin{bmatrix} a_1+b_1 & -a_2-b_2 & -a_3-b_3 & -a_4-b_4 \\ a_2+b_2 & a_1+b_1 & -a_4-b_4 & a_3+b_3 \\ a_3+b_3 & a_4+b_4 & a_1+b_1 & -a_2-b_2 \\ a_4+b_4 & -a_3-b_3 & a_2+b_2 & a_1+b_1 \end{bmatrix}$$

2.

$$\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}$$

$$= \begin{bmatrix} a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4 & -a_1b_2 - a_2b_1 - a_3b_4 + a_4b_3 & -a_1b_3 + a_2b_4 - a_3b_1 - a_4b_2 & -a_1b_4 - a_2b_3 + a_3b_2 - a_4b_1 \\ a_2b_1 + a_1b_2 - a_4b_3 + a_3b_4 & -a_2b_2 + a_1b_1 - a_4b_4 - a_3b_3 & -a_2b_3 - a_1b_4 - a_4b_1 + a_3b_2 & -a_2b_4 + a_1b_3 + a_4b_2 + a_3b_1 \\ a_3b_1 + a_4b_2 + a_1b_3 - a_2b_4 & -a_3b_2 + a_4b_1 + a_1b_4 + a_2b_3 & -a_3b_3 - a_4b_4 + a_1b_1 - a_2b_2 & -a_3b_4 + a_4b_3 - a_1b_2 - a_2b_1 \\ a_4b_1 - a_3b_2 + a_2b_3 + a_1b_4 & -a_4b_2 - a_3b_1 + a_2b_4 - a_1b_3 & -a_4b_3 + a_3b_4 + a_2b_1 + a_1b_2 & -a_4b_4 - a_3b_3 - a_2b_2 + a_1b_1 \end{bmatrix}$$

3.

$$\left\{\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} + \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}\right\}\alpha$$

$$= \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}\alpha + \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}\alpha$$

4.

$$\left\{\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}\right\}\alpha$$

$$= \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \left\{\begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}\alpha\right\}$$

$$= \left\{\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}\alpha\right\} \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}$$

$\forall a_1, a_2, a_3, a_4, b_1, b_2, b_3, b_4 \in \mathbb{R}$ *and* $\alpha \in \mathbb{R}$. *The set* $M_4(\mathbb{R})$, *together with these operations, form a division algebra over the field of real numbers. (See Rautenbach (1983) [28].)*

Proof. Let $\mathbf{A} = \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}$ and

$$\mathbf{C} = \begin{bmatrix} c_1 & -c_2 & -c_3 & -c_4 \\ c_2 & c_1 & -c_4 & c_3 \\ c_3 & c_4 & c_1 & -c_2 \\ c_4 & -c_3 & c_2 & c_1 \end{bmatrix} \in M_4(\mathbb{R}).$$ Then it follows that:

1.

$$\begin{aligned} &\mathbf{A}+\mathbf{B} \\ &= \begin{bmatrix} a_1+b_1 & -a_2-b_2 & -a_3-b_3 & -a_4-b_4 \\ a_2+b_2 & a_1+b_1 & -a_4-b_4 & a_3+b_3 \\ a_3+b_3 & a_4+b_4 & a_1+b_1 & -a_2-b_2 \\ a_4+b_4 & -a_3-b_3 & a_2+b_2 & a_1+b_1 \end{bmatrix} \in M_4(\mathbb{R}). \end{aligned}$$

2.

$$\begin{aligned} &\mathbf{AB} \\ &= \begin{bmatrix} a_1b_1-a_2b_2-a_3b_3-a_4b_4 & -a_1b_2-a_2b_1-a_3b_4+a_4b_3 & -a_1b_3+a_2b_4-a_3b_1-a_4b_2 & -a_1b_4-a_2b_3+a_3b_2-a_4b_1 \\ a_2b_1+a_1b_2-a_4b_3+a_3b_4 & -a_2b_2+a_1b_1-a_4b_4-a_3b_3 & -a_2b_3-a_1b_4-a_4b_1+a_3b_2 & -a_2b_4+a_1b_3+a_4b_2+a_3b_1 \\ a_3b_1+a_4b_2+a_1b_3-a_2b_4 & -a_3b_2+a_4b_1+a_1b_4+a_2b_3 & -a_3b_3-a_4b_4+a_1b_1-a_2b_2 & -a_3b_4+a_4b_3-a_1b_2-a_2b_1 \\ a_4b_1-a_3b_2+a_2b_3+a_1b_4 & -a_4b_2-a_3b_1+a_2b_4-a_1b_3 & -a_4b_3+a_3b_4+a_2b_1+a_1b_2 & -a_4b_4-a_3b_3-a_2b_2+a_1b_1 \end{bmatrix} \\ &\in M_4(\mathbb{R}). \end{aligned}$$

3.

$$\{\mathbf{A}+\mathbf{B}\}+\mathbf{C} = \mathbf{A}+\{\mathbf{B}+\mathbf{C}\}.$$

4. The set $M_4(\mathbb{R})$ contains an element $\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ such that:

$$\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}.$$

5. For every element $\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \in M_4(\mathbb{R})$ there exists an element in

$M_4(\mathbb{R})$ namely $\begin{bmatrix} -a_1 & a_2 & a_3 & a_4 \\ -a_2 & -a_1 & a_4 & -a_3 \\ -a_3 & -a_4 & -a_1 & a_2 \\ -a_4 & a_3 & -a_2 & -a_1 \end{bmatrix}$ such that:

$$\begin{aligned} &\begin{bmatrix} -a_1 & a_2 & a_3 & a_4 \\ -a_2 & -a_1 & a_4 & -a_3 \\ -a_3 & -a_4 & -a_1 & a_2 \\ -a_4 & a_3 & -a_2 & -a_1 \end{bmatrix} + \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \\ = &\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} + \begin{bmatrix} -a_1 & a_2 & a_3 & a_4 \\ -a_2 & -a_1 & a_4 & -a_3 \\ -a_3 & -a_4 & -a_1 & a_2 \\ -a_4 & a_3 & -a_2 & -a_1 \end{bmatrix} \\ = &\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

6.
$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.$$

7.
$$\{\mathbf{AB}\}\,\mathbf{C} = \mathbf{A}\,\{\mathbf{BC}\}\,.$$

8.
$$\mathbf{A}\,\{\mathbf{B} + \mathbf{C}\} = \mathbf{AB} + \mathbf{AC}.$$

9.
$$\{\mathbf{A} + \mathbf{B}\}\,\mathbf{C} = \mathbf{AC} + \mathbf{BC}.$$

10. The set $M_4(\mathbb{R})$ contains an element namely $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ such that:

$$\begin{aligned}
&\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
= &\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \\
= &\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}.
\end{aligned}$$

11. For each $\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \in M_4(\mathbb{R})$ there exists an element

$$\frac{1}{(a_1^2 + a_2^2 + a_3^2 + a_4^2)} \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ -a_2 & a_1 & a_4 & -a_3 \\ -a_3 & -a_4 & a_1 & a_2 \\ -a_4 & a_3 & -a_2 & a_1 \end{bmatrix} \in M_4(\mathbb{R})$$

such that:

$$
\begin{aligned}
&\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \frac{1}{(a_1^2 + a_2^2 + a_3^2 + a_4^2)} \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ -a_2 & a_1 & a_4 & -a_3 \\ -a_3 & -a_4 & a_1 & a_2 \\ -a_4 & a_3 & -a_2 & a_1 \end{bmatrix} \\
&= \frac{1}{(a_1^2 + a_2^2 + a_3^2 + a_4^2)} \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ -a_2 & a_1 & a_4 & -a_3 \\ -a_3 & -a_4 & a_1 & a_2 \\ -a_4 & a_3 & -a_2 & a_1 \end{bmatrix} \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.
\end{aligned}
$$

12. $\forall \alpha \in \mathbb{R}$ it follows that:

$$\{\mathbf{A} + \mathbf{B}\}\,\alpha = \mathbf{A}\alpha + \mathbf{B}\alpha$$

and

$$
\begin{aligned}
\{\mathbf{AB}\}\,\alpha =& \mathbf{A}\,\{\mathbf{B}\alpha\} \\
=& \{\mathbf{A}\alpha\}\,\mathbf{B}.
\end{aligned}
$$

$\square$

Remark A.4.5. In a similar fashion as in Theorem A.4.4 it can be shown that the sets

$$M_4^1(\mathbb{R}) = \left\{ \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ -a_2 & a_1 & -a_4 & a_3 \\ -a_3 & a_4 & a_1 & -a_2 \\ -a_4 & -a_3 & a_2 & a_1 \end{bmatrix} | a_1, a_2, a_3, a_4 \in \mathbb{R} \right\}$$

$$M_4^2(\mathbb{R}) = \left\{ \begin{bmatrix} a_1 & -a_2 & -a_3 & a_4 \\ a_2 & a_1 & -a_4 & -a_3 \\ a_3 & a_4 & a_1 & a_2 \\ -a_4 & a_3 & -a_2 & a_1 \end{bmatrix} | a_1, a_2, a_3, a_4 \in \mathbb{R} \right\}$$

$$M_4^3(\mathbb{R}) = \left\{ \begin{bmatrix} a_1 & -a_2 & a_3 & -a_4 \\ a_2 & a_1 & a_4 & a_3 \\ -a_3 & -a_4 & a_1 & a_2 \\ a_4 & -a_3 & -a_2 & a_1 \end{bmatrix} | a_1, a_2, a_3, a_4 \in \mathbb{R} \right\}$$

$$M_4^4(\mathbb{R}) = \left\{ \begin{bmatrix} a_1 & a_2 & -a_3 & -a_4 \\ -a_2 & a_1 & a_4 & -a_3 \\ a_3 & -a_4 & a_1 & -a_2 \\ a_4 & a_3 & a_2 & a_1 \end{bmatrix} | a_1, a_2, a_3, a_4 \in \mathbb{R} \right\}$$

together with the appropriate addition, multiplication and scalar multiplication operators, all form division algebras over the field of real numbers. (See Rautenbach (1983) [28].)

Definition A.4.6. Define a mapping $f : \mathbb{Q} \to M_4(\mathbb{R})$ as:

$$f(a_1 + ia_2 + ja_3 + ka_4) = \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}, \ \forall a_1 + ia_2 + ja_3 + ka_4 \in \mathbb{Q}.$$

(See Rautenbach (1983) [28].)

Theorem A.4.7. *f is a faithful representation. (See Rautenbach (1983) [28].)*

Proof. 1. The mapping $f : \mathbb{Q} \to M_4(\mathbb{R})$ is a homomorphism since:

(a)

$$
\begin{aligned}
& f\left\{(a_1 + ia_2 + ja_3 + ka_4) + (b_1 + ib_2 + jb_3 + kb_4)\right\} \\
=& f\left\{(a_1 + b_1) + i(a_2 + b_2) + j(a_3 + b_3) + k(a_4 + b_4)\right\} \\
=& \begin{bmatrix} a_1 + b_1 & -(a_2 + b_2) & -(a_3 + b_3) & -(a_4 + b_4) \\ a_2 + b_2 & a_1 + b_1 & -(a_4 + b_4) & a_3 + b_3 \\ a_3 + b_3 & a_4 + b_4 & a_1 + b_1 & -(a_2 + b_2) \\ a_4 + b_4 & -(a_3 + b_3) & a_2 + b_2 & a_1 + b_1 \end{bmatrix} \\
=& \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} + \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix} \\
=& f(a_1 + ia_2 + ja_3 + ka_4) + f(b_1 + ib_2 + jb_3 + kb_4).
\end{aligned}
$$

(b)

$$
\begin{aligned}
& f\left\{(a_1 + ia_2 + ja_3 + ka_4)(b_1 + ib_2 + jb_3 + kb_4)\right\} \\
=& \begin{bmatrix} a_1b_1 - a_2b_2 - a_3b_3 - a_4b_4 & -a_1b_2 - a_2b_1 - a_3b_4 + a_4b_3 & -a_1b_3 + a_2b_4 - a_3b_1 - a_4b_2 & -a_1b_4 - a_2b_3 + a_3b_2 - a_4b_1 \\ a_2b_1 + a_1b_2 - a_4b_3 + a_3b_4 & -a_2b_2 + a_1b_1 - a_4b_4 - a_3b_3 & -a_2b_3 - a_1b_4 - a_4b_1 + a_3b_2 & -a_2b_4 + a_1b_3 + a_4b_2 + a_3b_1 \\ a_3b_1 + a_4b_2 + a_1b_3 - a_2b_4 & -a_3b_2 + a_4b_1 + a_1b_4 + a_2b_3 & -a_3b_3 - a_4b_4 + a_1b_1 - a_2b_2 & -a_3b_4 + a_4b_3 - a_1b_2 - a_2b_1 \\ a_4b_1 - a_3b_2 + a_2b_3 + a_1b_4 & -a_4b_2 - a_3b_1 + a_2b_4 - a_1b_3 & -a_4b_3 + a_3b_4 + a_2b_1 + a_1b_2 & -a_4b_4 - a_3b_3 - a_2b_2 + a_1b_1 \end{bmatrix} \\
=& \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix} \\
=& f(a_1 + ia_2 + ja_3 + ka_4)f(b_1 + ib_2 + jb_3 + kb_4)
\end{aligned}
$$

(c)

$$\begin{aligned}
& f\{(a_1 + ia_2 + ja_3 + ka_4)\alpha\} \\
=& \begin{bmatrix} a_1\alpha & -a_2\alpha & -a_3\alpha & -a_4\alpha \\ a_2\alpha & a_1\alpha & -a_4\alpha & a_3\alpha \\ a_3\alpha & a_4\alpha & a_1\alpha & -a_2\alpha \\ a_4\alpha & -a_3\alpha & a_2\alpha & a_1\alpha \end{bmatrix} \\
=& \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \alpha \\
=& f(a_1 + ia_2 + ja_3 + ka_4)\alpha.
\end{aligned}$$

2. f is a monomorphism since if

$$f(a_1 + ia_2 + ja_3 + ka_4) \neq f(b_1 + ib_2 + jb_3 + kb_4)$$

then it follows that

$$\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \neq \begin{bmatrix} b_1 & -b_2 & -b_3 & -b_4 \\ b_2 & b_1 & -b_4 & b_3 \\ b_3 & b_4 & b_1 & -b_2 \\ b_4 & -b_3 & b_2 & b_1 \end{bmatrix}.$$

3. Finally, it follows that

$$\forall \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix} \in M_4(\mathbb{R})$$

there exists a

$$(a_1 + ia_2 + ja_3 + ka_4) \in \mathbb{Q}$$

such that

$$f(a_1 + ia_2 + ja_3 + ka_4) = \begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}.$$

4. Thus it follows that the mapping $f : \mathbb{Q} \to M_4(\mathbb{R})$ is an isomorphism.

5. It can therefore be concluded that f is a faithful representation.

□

Remark A.4.8. 1. Consequently, the mapping $f : \mathbb{Q} \to M_4(\mathbb{R})$ now substitutes the elements of $\mathbb{Q}$ with 4×4 matrices with specific structures of the form $\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}$, and the operations are substituted by matrix operations.

2. In a similar fashion as in Theorem A.4.7 it can be shown that the four structures of $M_4^1(\mathbb{R})$, $M_4^2(\mathbb{R})$, $M_4^3(\mathbb{R})$ and $M_4^4(\mathbb{R}$, given in Remark A.4.5, may also be used to construct faithful representations.

A.5 Algebraic results for quaternions

In this section some useful algebraic results are given to aid in the construction and development of the quaternion distribution theory.

Definition A.5.1. Let

$$\underset{4\times 4_{st}}{\mathbf{R}} = \begin{bmatrix} a_{1st} & -a_{2st} & -a_{3st} & -a_{4st} \\ a_{2st} & a_{1st} & -a_{4st} & a_{3st} \\ a_{3st} & a_{4st} & a_{1st} & -a_{2st} \\ a_{4st} & -a_{3st} & a_{2st} & a_{1st} \end{bmatrix}$$

and

$$q_{st} = a_{1st} + ia_{2st} + ja_{3st} + ka_{4st} \; \forall a_{1st}, a_{2st}, a_{3st}, a_{4st} \in \mathbb{R}$$

Now define the real matrix $\underset{4p\times 4p}{\mathbf{R}}$ and the quaternion matrix $\underset{p\times p}{\mathbf{Q}}$ respectively to be:

$$\underset{4p\times 4p}{\mathbf{R}} = [\mathbf{R}_{st}]$$

and

$$\underset{p\times p}{\mathbf{Q}} = [q_{st}] \,.$$

Theorem A.5.2. *Let*

$$M_{4p}(\mathbb{R}) = \left\{ \underset{4p\times 4p}{\mathbf{R}} | \mathbf{R}_{st} \in M_4(\mathbb{R}), s,t = 1,\dots,p \right\}$$

and let the following operations be defined:

1. $\mathbf{R}_1 + \mathbf{R}_2 \in M_{4p}(\mathbb{R})\ \forall \mathbf{R}_1, \mathbf{R}_2 \in M_{4p}(\mathbb{R})$.

2. $\mathbf{R}_1\mathbf{R}_2 \in M_{4p}(\mathbb{R})\ \forall \mathbf{R}_1, \mathbf{R}_2 \in M_{4p}(\mathbb{R})$.

3. $\{\mathbf{R}_1 + \mathbf{R}_2\}\,\alpha = \mathbf{R}_1\alpha + \mathbf{R}_2\alpha\ \forall \mathbf{R}_1, \mathbf{R}_2 \in M_{4p}(\mathbb{R}), \alpha \in \mathbb{R}$.

4.
$$\begin{aligned}\{\mathbf{R}_1\mathbf{R}_2\}\,\alpha =& \mathbf{R}_1\{\mathbf{R}_2\alpha\} \\ =& \{\mathbf{R}_1\alpha\}\,\mathbf{R}_2\ \forall \mathbf{R}_1, \mathbf{R}_2 \in M_{4p}(\mathbb{R}), \alpha \in \mathbb{R}.\end{aligned}$$

The set $M_{4p}(\mathbb{R})$ together with these operations form a division algebra over the field of real numbers. (See Rautenbach (1983) [28].)

Proof. The proof is analogous to that given in Theorem A.4.4. □

Theorem A.5.3. *Let*

$$M_p(\mathbb{Q}) = \left\{ \underset{p\times p}{\mathbf{Q}} | q_{st} \in \mathbb{Q}, s,t = 1,\dots,p \right\}$$

and let the following operations be defined:

1. $\mathbf{Q}_1 + \mathbf{Q}_2 \in M_p(\mathbb{Q})\ \forall \mathbf{Q}_1, \mathbf{Q}_2 \in M_p(\mathbb{Q})$.

2. $\mathbf{Q}_1\mathbf{Q}_2 \in M_p(\mathbb{Q})\ \forall \mathbf{Q}_1, \mathbf{Q}_2 \in M_p(\mathbb{Q})$.

3. $\{\mathbf{Q}_1 + \mathbf{Q}_2\}\,\alpha = \mathbf{Q}_1\alpha + \mathbf{Q}_2\alpha \ \forall \mathbf{Q}_1, \mathbf{Q}_2 \in M_p(\mathbb{Q}), \alpha \in \mathbb{R}.$

4.

$$\begin{aligned}\{\mathbf{Q}_1\mathbf{Q}_2\}\,\alpha =& \mathbf{Q}_1\{\mathbf{Q}_2\alpha\}\\ =& \{\mathbf{Q}_1\alpha\}\,\mathbf{Q}_2 \ \forall \mathbf{Q}_1, \mathbf{Q}_2 \in M_p(\mathbb{Q}), \alpha \in \mathbb{R}.\end{aligned}$$

The set $M_p(\mathbb{Q})$ together with these operations form a division algebra over the field of real numbers. (See Rautenbach (1983) [28].)

Proof. The proof is analogous to that given in Theorem A.4.3. □

Definition A.5.4. Define the mapping $f : M_{4p}(\mathbb{R}) \to M_p(\mathbb{Q})$ as:

$$f\left(\underset{4p\times 4p}{\mathbf{R}}\right) = \underset{p\times p}{\mathbf{Q}} \ \forall \mathbf{R} \in M_{4p}(\mathbb{R}), \mathbf{Q} \in \mathbb{Q}.$$

Theorem A.5.5. *f is a faithful representation. (See Rautenbach (1983) [28].)*

Proof. The proof is analogous to that given in Theorem A.4.7. □

Remark A.5.6.

1. If $\underset{4p\times 4p}{\mathbf{R}} \in M_{4p}(\mathbb{R})$ and $f(\mathbf{R}) = \mathbf{Q} \in M_p(\mathbb{Q})$ then it will be indicated as

$$\mathbf{R} \simeq \mathbf{Q}.$$

2. A matrix is called a quaternion Hermitian matrix if:

$$\bar{\mathbf{Q}}' = \mathbf{Q}.$$

(See Rautenbach (1983) [28].)

Theorem A.5.7. *Suppose $\underset{4p\times 4p}{\mathbf{R}} \in M_{4p}(\mathbb{R})$, $\underset{p\times p}{\mathbf{Q}} \in M_p(\mathbb{Q})$ and $\mathbf{R} \simeq \mathbf{Q}$. Then $\mathbf{R}$ is symmetric if and only if $\mathbf{Q}$ is a quaternion Hermitian matrix. (See Rautenbach and Roux (1983) [28].)*

Proof.

1. First, suppose that $\underset{4p\times 4p}{\mathbf{R}}$ is symmetric. It follows that:

$$\begin{aligned}\mathbf{R} =& \mathbf{R}'\\ \therefore [r_{st}] =& [r_{ts}]\end{aligned}$$

and so $a_{1st} = a_{1ts}$, $a_{2st} = -a_{2ts}$, $a_{3st} = -a_{3ts}$, and $a_{4st} = -a_{4ts}$. Therefore

$$\begin{aligned} q_{st} =& a_{1st} + ia_{2st} + ja_{3st} + ka_{4st} \\ =& a_{1ts} - ia_{2ts} - ja_{3ts} - ka_{4ts} \\ =& \overline{a_{1ts} + ia_{2ts} + ja_{3ts} + ka_{4ts}} \\ =& \bar{q}_{st}. \end{aligned}$$

Thus it follows that:

$$\mathbf{Q} = \bar{\mathbf{Q}}'.$$

2. Conversely, if $\mathbf{Q} = \bar{\mathbf{Q}}'$ then it follows that:

$$q_{st} = a_{1st} + ia_{2st} + ja_{3st} + ka_{4st}$$

and

$$\bar{q}_{ts} = a_{1ts} - ia_{2ts} - ja_{3ts} - ka_{4ts}$$

Consequently, it follows that $a_{1st} = a_{1ts}$, $a_{2st} = -a_{2ts}$, $a_{3st} = -a_{3ts}$, and $a_{4st} = -a_{4ts}$. from which it finally follows that

$$\mathbf{R} = \mathbf{R}'.$$

□

Theorem A.5.8. *Suppose* $\underset{4p\times 4p}{\mathbf{R}} \in M_{4p}(\mathbb{R})$, $\underset{p\times p}{\mathbf{Q}} \in M_p(\mathbb{Q})$ *and* $\mathbf{R} \simeq \mathbf{Q}$. *Then* $\mathbf{R}$ *is orthogonal if and only if* $\mathbf{Q}$ *is symplectic. (See Rautenbach and Roux (1983) [28].)*

Proof. 1. First, suppose that $\underset{4p\times 4p}{\mathbf{R}}$ is orthogonal such that:

$$\mathbf{R}' = \mathbf{R}^{-1},$$

thus

$$\left[r_{st}\right]' = \left[r^{st}\right].$$

It consequently follows that:

$$\begin{aligned} a_{1st} =& a_1^{st} \\ a_{2st} =& -a_2^{st} \\ a_{3st} =& -a_3^{st} \\ a_{4st} =& -a_4^{st} \end{aligned}$$

such that

$$\begin{aligned}\left[q^{st}\right] &= \left[a_1^{st} + ia_2^{st} + ja_3^{st} + ka_4^{st}\right] \\ &= \left[a_{1ts} - ia_{2ts} - ja_{3ts} - ka_{4ts}\right] \\ &= \left[\bar{q}_{ts}\right] \\ &= \left[\bar{q}'_{st}\right] \\ \therefore \mathbf{Q}^{-1} &= \bar{\mathbf{Q}}',\end{aligned}$$

i.e. $\mathbf{Q}$ is symplectic.

2. Conversely, suppose that $\underset{p\times p}{\mathbf{Q}}$ is symplectic. Then it follows in exactly the same way, as shown above, that $\mathbf{R}$ is orthogonal.

□

Theorem A.5.9. *Suppose* $\underset{4p\times 4p}{\mathbf{R}} \in M_{4p}(\mathbb{R})$, $\underset{p\times p}{\mathbf{Q}} \in M_p(\mathbb{Q})$ *and* $\mathbf{R} \simeq \mathbf{Q}$. *Then* $\mathbf{R}$ *is nonsingular with inverse*

$$\begin{aligned}\underset{4p\times 4p}{\mathbf{R}}^{-1} &= \left[r^{st}\right] \\ &= \left[\begin{bmatrix} a_1^{st} & -a_2^{st} & -a_3^{st} & -a_4^{st} \\ a_2^{st} & a_1^{st} & -a_4^{st} & a_3^{st} \\ a_3^{st} & a_4^{st} & a_1^{st} & -a_2^{st} \\ a_4^{st} & -a_3^{st} & a_2^{st} & a_1^{st} \end{bmatrix}\right]\end{aligned}$$

if and only if $\mathbf{Q}$ *is nonsingular with inverse*

$$\begin{aligned}\underset{p\times p}{\mathbf{Q}}^{-1} &= \left[q^{st}\right] \\ &= \left[a_1^{st} + ia_2^{st} + ja_3^{st} + ka_4^{st}\right].\end{aligned}$$

(See Rautenbach and Roux (1985) [29].)

Proof. 1. First, suppose that $\underset{p\times p}{\mathbf{Q}}$ is nonsingular with inverse

$$\underset{p\times p}{\mathbf{Q}}^{-1} = \left[q^{st}\right] \in M_p(\mathbb{Q})$$

such that

$$\mathbf{Q}\mathbf{Q}^{-1} = \mathbf{I}_p.$$

Since there exists a faithful representation between $M_p(\mathbb{Q})$ and $M_{4p}(\mathbb{R})$ in such a way that there exists $\mathbf{R}^{-1} \in M_{4p}(\mathbb{R})$ it follows that

$$\mathbf{Q}^{-1} \simeq \mathbf{R}^{-1}$$

and hence

$$\mathbf{R}\mathbf{R}^{-1} = \mathbf{I}_{4p}.$$

Thus, $\mathbf{R}$ is nonsingular with inverse

$$\underset{4p\times 4p}{\mathbf{R}}^{-1} = \left[\begin{bmatrix} a_1^{st} & -a_2^{st} & -a_3^{st} & -a_4^{st} \\ a_2^{st} & a_1^{st} & -a_4^{st} & a_3^{st} \\ a_3^{st} & a_4^{st} & a_1^{st} & -a_2^{st} \\ a_4^{st} & -a_3^{st} & a_2^{st} & a_1^{st} \end{bmatrix}\right].$$

2. Conversely, suppose that $\mathbf{R}$ is nonsingular with inverse as defined above. In exactly the same way as shown above, there exists a $\mathbf{Q}^{-1}$ such that

$$\mathbf{Q}\mathbf{Q}^{-1} = \mathbf{I}_p.$$

□

Theorem A.5.10. *Let* $\underset{p\times p}{\mathbf{Q}}$ *with quaternion entries. If the* p *quaternion eigenvectors of* $\mathbf{Q}$ *are all distinct, then there exists a matrix* $\underset{p\times p}{\mathbf{U}}$ *with quaternion entries such that*

$$\mathbf{Q} = \mathbf{U}\mathbf{E}_{\lambda_q}\mathbf{U}^{-1}$$

where

$$\mathbf{E}_{\lambda_q} = \text{diag}\left(\lambda_{q_1}, \ldots, \lambda_{q_p}\right)$$

with λ_{q_s} $(s = 1, \ldots, p)$ *the quaternion characteristic root of* $\mathbf{Q}$.

Proof. See Mehta (1967) [26] for a proof of this. □

Theorem A.5.11. *Let* $\underset{p\times p}{\mathbf{Q}}$*, with quaternion entries, be a Hermitian matrix. Then there exists a symplectic matrix* $\mathbf{H}$ *such that:*

$$\begin{aligned}\mathbf{Q} =& \mathbf{H}\mathbf{D}_\lambda\mathbf{H}^{-1}\\ =& \mathbf{H}\mathbf{D}_\lambda\bar{\mathbf{H}}'\end{aligned}$$

where

$$\mathbf{D}_\lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_p)$$

with

$$\lambda_s = \lambda_{s\mathbf{R}} + i0 + j0 + k0, s = 1, \ldots, p$$

the real characteristic roots of $\mathbf{Q}$*.*

Proof. See Mehta (1967) [26] for a proof of this. □

Theorem A.5.12. *Suppose* $\underset{4p\times 4p}{\mathbf{R}} \in M_{4p}(\mathbb{R})$*,* $\underset{p\times p}{\mathbf{Q}} \in M_p(\mathbb{Q})$ *and* $\mathbf{R} \simeq \mathbf{Q}$*. If* $\mathbf{R}$ *is symmetric, then*

$$\det \mathbf{R} = (\det \mathbf{Q})^4 .$$

(See Rautenbach and Roux (1985) [29].)

Proof. From Theorem A.5.7 it follows that if $\mathbf{R}$ is symmetric, then $\mathbf{Q}$ is a quaternion Hermitian. From Theorem A.5.11 it follows that there exists a symplectic matrix, $\mathbf{H}$ such that

$$\mathbf{Q} = \mathbf{H}\mathbf{D}_\lambda\bar{\mathbf{H}}'$$

with $\mathbf{D}_\lambda$ a diagonal matrix with elements equal to the characteristic roots of $\mathbf{Q}$. Now suppose that $\underset{4p\times 4p}{\mathbf{M}} \in M_{4p}(\mathbb{R})$ is an orthogonal matrix such that $\mathbf{M} \simeq \mathbf{H}$ (see Theorem A.5.8). Let $\underset{4p\times 4p}{\mathbf{R_D}} \in M_{4p}(\mathbb{R})$ be such that $\mathbf{R_D} \simeq \mathbf{D}_\lambda$. Then it follows that

$$\begin{aligned}\mathbf{Q} =& \mathbf{H}\mathbf{D}_\lambda\bar{\mathbf{H}}'\\ \simeq& \mathbf{M}\mathbf{R_D}\mathbf{M}'\\ =& \mathbf{R}.\end{aligned}$$

The matrix $\mathbf{R_D}$ is given by:

$$\underset{4p\times 4p}{\mathbf{R_D}} = \begin{bmatrix} \lambda_{1\mathbf{R}} & 0 & 0 & 0 & \dots & \dots & \dots & \dots & 0 & 0 & 0 & 0 \\ 0 & \lambda_{1\mathbf{R}} & 0 & 0 & \dots & \dots & \dots & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_{1\mathbf{R}} & 0 & \dots & \dots & \dots & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_{1\mathbf{R}} & \dots & \dots & \dots & \dots & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots & \dots & \dots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots & \dots & \dots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots & \dots & \dots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \dots & \dots & \dots & \dots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & \lambda_{p\mathbf{R}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & 0 & \lambda_{p\mathbf{R}} & 0 & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & 0 & 0 & \lambda_{p\mathbf{R}} & 0 \\ 0 & 0 & 0 & 0 & \dots & \dots & \dots & \dots & 0 & 0 & 0 & \lambda_{p\mathbf{R}} \end{bmatrix}.$$

It now follows that

$$\begin{aligned} \det \mathbf{R} &= \det \left(\mathbf{M}\mathbf{R_D}\mathbf{M}'\right) \\ &= \det \mathbf{R_D}, \text{ since } \mathbf{M} \text{ is orthogonal} \\ &= \prod_{s=1}^{p} \det \begin{bmatrix} \lambda_{s\mathbf{R}} & 0 & 0 & 0 \\ 0 & \lambda_{s\mathbf{R}} & 0 & 0 \\ 0 & 0 & \lambda_{s\mathbf{R}} & 0 \\ 0 & 0 & 0 & \lambda_{s\mathbf{R}} \end{bmatrix} \\ &= \prod_{s=1}^{p} (\lambda_s)^4, \ \lambda_s = \lambda_{s\mathbf{R}} + i0 + j0 + k0 \\ &= (\det \mathbf{D}_\lambda)^4 \\ &= (\det \mathbf{Q})^4. \end{aligned}$$

□

Theorem A.5.13. *Let* $\underset{4p\times 1}{\underline{q}_0} = [b_{11}, b_{21}, b_{31}, b_{41}, \dots, b_{1p}, b_{2p}, b_{3p}, b_{4p}]'$ *and* $\underset{p\times 1}{\underline{q}} = [q_1, \dots, q_p]'$ *with* $q_s = b_{1s} + ib_{2s} + jb_{3s} + kb_{4s}, s = 1, \dots, p$. *Suppose* $\underset{4p\times 4p}{\mathbf{R}} \in M_{4p}(\mathbb{R})$, $\underset{p\times p}{\mathbf{Q}} \in M_p(\mathbb{Q})$ *and*

$\mathbf{R} \simeq \mathbf{Q}$. *Let* $\mathbf{R}$ *be symmetric, such that* $\mathbf{Q}$ *is quaternion Hermitian. It now follows that*

$$\underline{q}_0'\mathbf{R}\underline{q}_0 = \bar{\underline{q}}'\mathbf{Q}\underline{q}.$$

(See Rautenbach and Roux (1985) [29].)

Proof. It is clear that

$$\underline{q}_0'\mathbf{R}\underline{q}_0 = \sum_{s,t=1}^{p} [b_{s1}, b_{s2}, b_{s3}, b_{s4}] \begin{bmatrix} a_{1st} & -a_{2st} & -a_{3st} & -a_{4st} \\ a_{2st} & a_{1st} & -a_{4st} & a_{3st} \\ a_{3st} & a_{4st} & a_{1st} & -a_{2st} \\ a_{4st} & -a_{3st} & a_{2st} & a_{1st} \end{bmatrix} \begin{bmatrix} b_{t1} \\ b_{t2} \\ b_{t3} \\ b_{t4} \end{bmatrix}$$

$$\begin{aligned} &= \sum_{s,t=1}^{p} [a_{1st}(b_{s1}b_{t1} + b_{s2}b_{t2} + b_{s3}b_{t3} + b_{s4}b_{t4}) \\ &\quad + a_{2st}(b_{s2}b_{t1} - b_{s1}b_{t2} + b_{s4}b_{t3} - b_{s3}b_{t4}) \\ &\quad + a_{3st}(b_{s3}b_{t1} - b_{s4}b_{t2} - b_{s1}b_{t3} + b_{s2}b_{t4}) \\ &\quad + a_{4st}(b_{s4}b_{t1} + b_{s3}b_{t2} - b_{s2}b_{t3} - b_{s1}b_{t4})]. \end{aligned} \tag{A.5.1}$$

It also follows that

$$\begin{aligned}
\underline{\bar{q}}'\mathbf{Q}\underline{q} &= \sum_{s,t=1}^{p} \bar{q}_s q_{st} q_t \\
&= \sum_{s,t=1}^{p} \left(b_{s1} - ib_{s2} - jb_{s3} - kb_{s4}\right)\left(a_{1st} + ia_{2st} + ja_{3st} + ka_{4st}\right)\left(b_{t1} + ib_{t2} + jb_{t3} + kb_{t4}\right) \\
&= \sum_{s,t=1}^{p} \left[a_{1st}\left(b_{s1}b_{t1} + b_{s2}b_{t2} + b_{s3}b_{t3} + b_{s4}b_{t4}\right)\right. \\
&\quad + a_{2st}\left(b_{s2}b_{t1} - b_{s1}b_{t2} + b_{s4}b_{t3} - b_{s3}b_{t4}\right) \\
&\quad + a_{3st}\left(b_{s3}b_{t1} - b_{s4}b_{t2} - b_{s1}b_{t3} + b_{s2}b_{t4}\right) \\
&\quad \left. + a_{4st}\left(b_{s4}b_{t1} + b_{s3}b_{t2} - b_{s2}b_{t3} - b_{s1}b_{t4}\right)\right] \\
&\quad + i\sum_{s,t=1}^{p} \left[a_{1st}\left(-b_{s2}b_{t1} + b_{s1}b_{t2} + b_{s4}b_{t3} - b_{s3}b_{t4}\right)\right. \\
&\quad + a_{2st}\left(b_{s1}b_{t1} + b_{s2}b_{t2} - b_{s3}b_{t3} - b_{s4}b_{t4}\right) \\
&\quad + a_{3st}\left(b_{s4}b_{t1} + b_{s3}b_{t2} + b_{s2}b_{t3} + b_{s1}b_{t4}\right) \\
&\quad \left. + a_{4st}\left(-b_{s3}b_{t1} + b_{s4}b_{t2} - b_{s1}b_{t3} + b_{s2}b_{t4}\right)\right] \\
&\quad + j\sum_{s,t=1}^{p} \left[a_{1st}\left(-b_{s3}b_{t1} - b_{s4}b_{t2} + b_{s1}b_{t3} + b_{s2}b_{t4}\right)\right. \\
&\quad + a_{2st}\left(-b_{s4}b_{t1} + b_{s3}b_{t2} + b_{s2}b_{t3} - b_{s1}b_{t4}\right) \\
&\quad + a_{3st}\left(b_{s1}b_{t1} - b_{s2}b_{t2} + b_{s3}b_{t3} - b_{s4}b_{t4}\right) \\
&\quad \left. + a_{4st}\left(b_{s2}b_{t1} + b_{s1}b_{t2} + b_{s4}b_{t3} + b_{s3}b_{t4}\right)\right] \\
&\quad + k\sum_{s,t=1}^{p} \left[a_{1st}\left(-b_{s4}b_{t1} + b_{s3}b_{t2} - b_{s2}b_{t3} + b_{s1}b_{t4}\right)\right. \\
&\quad + a_{2st}\left(b_{s3}b_{t1} + b_{s4}b_{t2} + b_{s1}b_{t3} + b_{s2}b_{t4}\right) \\
&\quad + a_{3st}\left(-b_{s2}b_{t1} - b_{s1}b_{t2} + b_{s4}b_{t3} + b_{s3}b_{t4}\right) \\
&\quad \left. + a_{4st}\left(b_{s1}b_{t1} + b_{s2}b_{t2} - b_{s3}b_{t3} + b_{s4}b_{t4}\right)\right].
\end{aligned}$$

Given that $\mathbf{R}$ is symmetric, it follows that

$$\begin{aligned} a_{1st} &= a_{1ts} \\ a_{2st} &= -a_{2ts} \\ a_{3st} &= -a_{3ts} \\ a_{4st} &= -a_{4ts} \end{aligned}$$

from which it follows that

$$\begin{aligned} \underline{\bar{q}}'\mathbf{Q}\underline{q} = \sum_{s,t=1}^{p} & [a_{1st}\left(b_{s1}b_{t1} + b_{s2}b_{t2} + b_{s3}b_{t3} + b_{s4}b_{t4}\right) \\ & + a_{2st}\left(b_{s2}b_{t1} - b_{s1}b_{t2} + b_{s4}b_{t3} - b_{s3}b_{t4}\right) \\ & + a_{3st}\left(b_{s3}b_{t1} - b_{s4}b_{t2} - b_{s1}b_{t3} + b_{s2}b_{t4}\right) \\ & + a_{4st}\left(b_{s4}b_{t1} + b_{s3}b_{t2} - b_{s2}b_{t3} - b_{s1}b_{t4}\right)] . \end{aligned} \tag{A.5.2}$$

From (A.5.1) and (A.5.2) the required result follows. □

Corollary A.5.14. *Let* $\mathbf{R}$ *be a symmetric matrix with elements* $[R_{st}]$ *where*

$$R_{st} = \begin{bmatrix} a_{1st} & a_{2st} & a_{3st} & a_{4st} \\ -a_{2st} & a_{1st} & -a_{4st} & a_{3st} \\ -a_{3st} & a_{4st} & a_{1st} & -a_{2st} \\ -a_{4st} & -a_{3st} & a_{2st} & a_{1st} \end{bmatrix}$$

then it follows that

$$\underline{q}_0'\mathbf{R}\underline{q}_0 = \underline{q}'\mathbf{Q}\underline{\bar{q}}.$$

(See Rautenbach (1983) [28].)

Proof. The proof is analygous to that given in Theorem A.5.13. □

Corollary A.5.15. *Suppose* $\underset{4p\times 4p}{\mathbf{R}} \in M_{4p}(\mathbb{R})$*,* $\underset{p\times p}{\mathbf{Q}} \in M_p(\mathbb{Q})$ *and* $\mathbf{R} \simeq \mathbf{Q}$*. Then* $\mathbf{R}$ *is symmetric positive definite if and only if* $\mathbf{Q}$ *is a positive definite quaternion Hermitian matrix. (See Rautenbach and Roux (1985) [29].)*

Proof. 1. First suppose that $\mathbf{R}$ is symmetric positive definite. Then it follows that

$$\underline{q}_0'\mathbf{R}\underline{q}_0 > 0 \; \forall \underset{4p\times 1}{\underline{q}_0} \neq \underline{0},$$

and hence it follows from Theorem A.5.13 that

$$\bar{\underline{q}}'\mathbf{Q}\underline{q} > 0 \; \forall \underset{p\times 1}{\underline{q}} \neq \underline{0}.$$

If $\mathbf{R}$ is symmetric $\Rightarrow$ $\mathbf{Q}$ is quaternion Hermitian such that if $\mathbf{R}$ is symmetric positive definite $\Rightarrow$ $\mathbf{Q}$ is positive definite quaternion Hermitian.

2. The converse follows in a similar fashion.

□

Remark A.5.16. 1. Matrices of the form $\begin{bmatrix} a_1 & -a_2 & -a_3 & a_4 \\ a_2 & a_1 & -a_4 & -a_3 \\ a_3 & a_4 & a_1 & a_2 \\ -a_4 & a_3 & -a_2 & a_1 \end{bmatrix}$, $\begin{bmatrix} a_1 & -a_2 & a_3 & -a_4 \\ a_2 & a_1 & a_4 & a_3 \\ -a_3 & -a_4 & a_1 & a_2 \\ a_4 & -a_3 & -a_2 & a_1 \end{bmatrix}$ and $\begin{bmatrix} a_1 & a_2 & -a_3 & -a_4 \\ -a_2 & a_1 & a_4 & -a_3 \\ a_3 & -a_4 & a_1 & -a_2 \\ a_4 & a_3 & a_2 & a_1 \end{bmatrix}$ do not have the properties given in Theorem A.5.13 and Corollary A.5.14, however, all the other properties discussed in this section, apply to them as well.

2. For the purposes of this discussion, the matrices that are most frequently used are:

(a) $\begin{bmatrix} a_1 & -a_2 & -a_3 & -a_4 \\ a_2 & a_1 & -a_4 & a_3 \\ a_3 & a_4 & a_1 & -a_2 \\ a_4 & -a_3 & a_2 & a_1 \end{bmatrix}$, and

(b) $\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ -a_2 & a_1 & -a_4 & a_3 \\ -a_3 & a_4 & a_1 & -a_2 \\ -a_4 & -a_3 & a_2 & a_1 \end{bmatrix}$.

A.6 Functions and polynomials for quaternions

A number of scattered results that are frequently used are briefly presented below.

Lemma A.6.1. *The quaternion multivariate gamma function is defined by*

$$\begin{aligned}\mathbb{Q}\Gamma_p(a) =& \int\limits_{\mathbf{A}=\bar{\mathbf{A}}'>\mathbf{0}} \det(\mathbf{A})^{a-2(p-1)-1} \operatorname{etr}(-\mathbf{A})\mathrm{d}\mathbf{A} \\ =&\pi^{p(p-1)} \prod_{\alpha=1}^{p} \Gamma\left(a-2(\alpha-1)\right), \quad \operatorname{Re}(a) > 2(p-1), \end{aligned} \tag{A.6.1}$$

and $\mathbb{Q}\Gamma_p(a,\kappa) = (a)_\kappa \mathbb{Q}\Gamma_p(a)$ *is the quaternion generalised multivariate gamma function of weight* κ*, where the quaternion generalized hypergeometric coefficient* $(a)_\kappa$ *is defined by*

$$(a)_\kappa = \prod_{\alpha=1}^{p} \left(a-2(\alpha-1)\right)_{k_\alpha}$$

where $(a)_\alpha = a(a+1)\cdots(a-2\alpha+2), \alpha = 1,2,\ldots$ *with* $(a)_0 = 1$*. (See Gross and Richards, (1987) [12].)*

Lemma A.6.2. *The quaternion multivariate beta function is given by*

$$\begin{aligned}\mathbb{Q}B_p(a,b) =& \int\limits_{\mathbf{0}<\mathbf{A}=\bar{\mathbf{A}}'<\mathbf{I}_p} \det(\mathbf{A})^{a-2(p-1)-1} \det(\mathbf{I}_p-\mathbf{A})^{b-2(p-1)-1} \mathrm{d}\mathbf{A} \\ =&\frac{\mathbb{Q}\Gamma_p(a)\mathbb{Q}\Gamma_p(b)}{\mathbb{Q}\Gamma_p(a+b)} = \mathbb{Q}B_p(b,a), \quad \operatorname{Re}(a,b) > 2(p-1).\end{aligned}$$

(See Kabe (1984, equations 58 and 59) [19].)

Lemma A.6.3. *The quaternion generalized hypergeometric function with one matrix argument is defined by*

$${}_r\mathbb{Q}F_s(a_1,\ldots,a_r;b_1,\ldots,b_s;\mathbf{A}) = \sum_{k=0}^{\infty}\sum_{\kappa} \frac{(a_1)_\kappa\cdots(a_r)_\kappa}{(b_1)_\kappa\cdots(b_s)_\kappa} \frac{\mathbb{Q}C_\kappa(\mathbf{A})}{k!}, \tag{A.6.2}$$

where a_α*,* $\alpha = 1,\ldots,r$*,* b_β*,* $\beta = 1,\ldots,s$ *are arbitrary quaternion numbers,* $\underset{p\times p}{\mathbf{A}}$ *is a quaternion Hermitian matrix,* $\mathbb{Q}C_\kappa(\mathbf{A})$ *is the zonal polynomial of the quaternion Hermitian matrix* $\underset{p\times p}{\mathbf{A}}$ *corresponding to the partition* $\kappa = (k_1,\ldots,k_p), k_1 \geq \ldots \geq k_p \geq 0,$

$k_1 + \ldots + k_p = k$ and $\sum_\kappa$ denotes summation over all partitions κ. Conditions for convergence of the series are available in the literature, see Gross and Richards (1987) [12], and Li and Xue (2009) [23].

From (A.6.2) the following special cases follow:

$$_0\mathbb{Q}F_0(a;\mathbf{A}) = \operatorname{etr}(\mathbf{A}) \tag{A.6.3}$$

and

$$_1\mathbb{Q}F_0(a;\mathbf{A}) = \sum_{k=0}^{\infty}\sum_{\kappa}\frac{(a)_\kappa \mathbb{Q}C_\kappa(\mathbf{A})}{k!} = \det\left(\mathbf{I}_p - \mathbf{A}\right)^{-a}, \quad \|\mathbf{A}\| < 1. \tag{A.6.4}$$

For properties and further results on these functions, as well as their proofs the reader is referred to Gross and Richards (1987) [12] and Li and Xue (2009) [23].

Lemma A.6.4. *Let $\mathbf{A},\mathbf{B} \in M_p(\mathbb{Q})$ be Hermitian matrices with $\mathbf{A} > \mathbf{0}$, then*

$$\int\limits_{O(p)} \mathbb{Q}C_\kappa\left(\mathbf{AHB\bar{H}}'\right)\mathrm{d}\mathbf{H} = \frac{\mathbb{Q}C_\kappa\left(\mathbf{A}\right)\mathbb{Q}C_\kappa\left(\mathbf{B}\right)}{\mathbb{Q}C_\kappa\left(\mathbf{I}_p\right)}$$

where $\mathrm{d}\mathbf{H}$ is the normalised Haar invariant measure on $O\left(p\right)$ and

$$O\left(p\right) = \left\{\mathbf{A} \in M_p\left(\mathbb{Q}\right) \middle| \bar{\mathbf{A}}'\mathbf{A} = \mathbf{A}\bar{\mathbf{A}}' = \mathbf{I}_p\right\}.$$

(See Li and Xue (2009, Theorem 3.1) [23].)

Lemma A.6.5. *Let $\mathbf{X},\mathbf{A},\mathbf{B} \in M_p(\mathbb{Q})$, where $\mathbf{X}$ is a Hermitian matrix and $\mathbf{A},\mathbf{B} > \mathbf{0}$, then*

$$\int\limits_{O(p)} \mathbb{Q}C_\kappa\left(\mathbf{A}^{\frac{1}{2}}\mathbf{B}\mathbf{A}^{\frac{1}{2}}\mathbf{HX\bar{H}}'\right)\mathrm{d}\mathbf{H} = \int\limits_{O(p)} \mathbb{Q}C_\kappa\left(\mathbf{B}^{\frac{1}{2}}\mathbf{A}\mathbf{B}^{\frac{1}{2}}\mathbf{HX\bar{H}}'\right)\mathrm{d}\mathbf{H}$$

where $\mathrm{d}\mathbf{H}$ is the normalised Haar invariant measure on $O\left(p\right)$ and

$$O\left(p\right) = \left\{\mathbf{A} \in M_p\left(\mathbb{Q}\right) \middle| \bar{\mathbf{A}}'\mathbf{A} = \mathbf{A}\bar{\mathbf{A}}' = \mathbf{I}_p\right\}.$$

(This is analogous to Ehlers, Bekker and Roux (2009, Lemma 5) [9].)

Lemma A.6.6. *Let* $\mathbf{X}, \mathbf{Y} \in M_p(\mathbb{Q})$ *be positive definite Hermitian matrices, and let* $\mathbf{Y} = \mathbf{A}\mathbf{X}\bar{\mathbf{A}}' + \mathbf{C}$*, where* $\mathbf{A}$ *and* $\mathbf{C} \in M_p(\mathbb{Q})$ *are constant matrices. Then*

$$J(\mathbf{Y} \to \mathbf{X}) = \det\left(\bar{\mathbf{A}}'\mathbf{A}\right)^{2p-1}$$

(See Díaz-García (2009, equation (2.10)) [5].)

Lemma A.6.7. *Let* $\mathbf{U} \in M_p(\mathbb{Q})$ *be a positive definite, Hermitian matrix . If*

$$\mathbf{V} = (\mathbf{I}_p - \mathbf{U})^{-\frac{1}{2}}\mathbf{U}(\mathbf{I}_p - \mathbf{U})^{-\frac{1}{2}},$$

then

$$J(\mathbf{V} \to \mathbf{U}) = \det(\mathbf{I}_p - \mathbf{U})^{-2(2p-1)}.$$

Lemma A.6.8. *If* $f(\mathbf{A})$ *is a function of the positive definite quaternion matrix* $\underset{p\times p}{\mathbf{A}}$*, the Laplace transform of* $f(\mathbf{A})$ *is defined to be*

$$g(\mathbf{S}) = \mathcal{L}(f(\mathbf{A})) = \int_{\mathbf{A}>\mathbf{0}} \operatorname{etr}(\mathbf{A}\mathbf{S})f(\mathbf{A})\mathrm{d}\mathbf{A}$$

which is absolutely convergent for $\mathbf{S} \in \Phi$*, the generalized right half-plane. (See Li and Xue (2009) [23], Definition 3.2.)*

The inverse Laplace transformation as used by Díaz-García (2009) [5], equation (4.13) is given below.

Lemma A.6.9. *Let* $\mathbf{S}, \mathbf{A}, \mathbf{U} \in M_p(\mathbb{Q})$ *be Hermitian matrices, and* $\operatorname{Re}(a) > a_0$*. Then*

$$\begin{aligned}&\frac{2^{2p(p-1)}}{(2\pi\iota)^{2p(p-1)+p}} \int_{\mathbf{S}-\mathbf{S}_0\in\Phi} \operatorname{etr}(\mathbf{A}\mathbf{S})\det(\mathbf{S})^{-a}\,\mathbb{Q}C_\kappa\left(\mathbf{U}\mathbf{S}^{-1}\right)\mathrm{d}\mathbf{S}\\ =&\frac{1}{\mathbb{Q}\Gamma_p(a,\kappa)}\det(\mathbf{A})^{a-2(p-1)-1}\mathbb{Q}C_\kappa(\mathbf{A}\mathbf{U})\end{aligned}$$

where the Hermitian matrix $\mathbf{S}_0 \in M_p(\mathbb{Q})$*, and* Φ *is called the generalized right half-plane. (See Díaz-García (2009, equation (4.13)) [5].)*

The Laplace transform of the hypergeometric function is presented below.

Lemma A.6.10. *Assume* $r \leq s$, $\mathrm{Re}(a) > 2(p-1)$ *and* $\mathbf{B}$ *is a Hermitian matrix. Then*

$$\int\limits_{\mathbf{S}=\bar{\mathbf{S}}'>\mathbf{0}} \operatorname{etr}(-\mathbf{SA}) \det(\mathbf{S})^{a-2p+1} {}_r\mathbb{Q}F_s(a_1,\ldots,a_r;b_1,\ldots,b_s;\mathbf{SB})\,\mathrm{d}\mathbf{S}$$
$$= \mathbb{Q}\Gamma_p(a)\det(\mathbf{A})^{-a} {}_{r+1}\mathbb{Q}F_s\left(a_1,\ldots,a_r,a;b_1,\ldots,b_r;\mathbf{BA}^{-1}\right).$$

When $r < s$, *the integral above converges absolutely for all* $\mathbf{A} \in \Phi$, *and for* $r = s$, *the integral converges absolutely for all Hermitian matrices* $\mathbf{A}$, *such that* $\left\|(\mathrm{Re}\,\mathbf{A})^{-1}\right\| < 1$. *(See Díaz-García (2009, equation (4.6)) [5].)*

Lemma A.6.11.

$$\int\limits_{\mathbf{B}=\bar{\mathbf{B}}'>\mathbf{0}} \operatorname{etr}(-\mathbf{AZ})\det(\mathbf{A})^{a-2p+1}\,\mathrm{d}\mathbf{A} = \mathbb{Q}\Gamma_p(a)\det(\mathbf{Z})^{-a}.$$

(See Díaz-García (2009, equation (3.17)) [5].)

Lemma A.6.12. *Let* $\mathbf{A},\mathbf{U} \in M_p(\mathbb{Q})$ *be Hermitian matricesand* $\mathrm{Re}(a,b) > 2(p-1)$. *Then*

$$\int\limits_{\mathbf{0}<\mathbf{A}=\bar{\mathbf{A}}'<\mathbf{I}_p} \det(\mathbf{A})^{a-2(p-1)-1}\det(\mathbf{I}_p-\mathbf{A})^{b-2(p-1)-1}\,\mathbb{Q}C_\kappa(\mathbf{AU})\mathrm{d}\mathbf{A}$$
$$= \frac{\mathbb{Q}\Gamma_p(a,\kappa)\mathbb{Q}\Gamma_p(b)}{\mathbb{Q}\Gamma_p(a+b,\kappa)}\mathbb{Q}C_\kappa(\mathbf{U}).$$

(See Díaz-García (2009, equation (3.13).)

A.7 The symmetrised density function

The symmetrised density function was defined by Greenacre (1973) [11] and is adapted for the quaternionic space below. This result is required frequently in applications involving transformations of quaternion-valued variables, in obtaining exact expressions for pdf's.

Definition A.7.1. The symmetrised function $f_s(\mathbf{A})$, of a given function $f(\mathbf{A})$, is defined as

$$f_s(\mathbf{A}) = \int\limits_{O(p)} f\left(\mathbf{HA}\bar{\mathbf{H}}'\right)\mathrm{d}\mathbf{H},\ \ \mathbf{H} \in O(p)$$

where $d\mathbf{H}$ denotes the normalised Haar invariant measure on $O(p)$, and

$$O(p) = \left\{ \mathbf{A} \in M_p(\mathbb{Q}) \left| \bar{\mathbf{A}}'\mathbf{A} = \mathbf{A}\bar{\mathbf{A}}' = \mathbf{I}_p \right. \right\}.$$

Díaz-García and Gutiérrez-Jáimez (2006) [6] proposed the inverse application of the symmetrised pdf defined by Greenacre (1973) [11], i.e. if the symmetrised pdf of $\mathbf{A}$ is given by Definition A.7.1 then the pdf of $\mathbf{A}$ can be obtained from $f\left(\mathbf{H}\mathbf{A}\bar{\mathbf{H}}'\right)$ by making the transformation $\mathbf{H}\mathbf{A}\bar{\mathbf{H}}' \to \mathbf{A}$.

A.8 The Mellin transform and Meijer's G-function

Definition A.8.1. If $f(x)$ is a real-valued function which is single-valued almost everywhere for $x \geq 0$, and if the integral

$$\int_0^\infty x^{k-1} \left| f(x) \right| dx$$

converges for some value of k, then the Mellin transform of $f(x)$ is defined as

$$M_f(s) = \int_0^\infty x^{s-1} f(x)\, dx \tag{A.8.1}$$

where $M_f(s)$ is the Mellin transform of f with respect to the parameter s, and s is a complex number. The inverse Mellin transform is given by the inverse integral

$$f(x) = \frac{1}{2\pi\iota} \int_{c-\iota\infty}^{c+\iota\infty} M_f(s)\, x^{-s} ds$$

where $\iota = \sqrt{-1}$, and c is a real number in the strip of analyticity of $M_f(s)$. For more detail the reader is referred to Mathai (1993, Definition 1.8, pp 23) [25].

Definition A.8.2. Meijer's G-function, with the parameters $a_1, \ldots, a_p$ and $b_1, \ldots, b_q$ is defined as

$$G_{p,q}^{m,n}\left(z \left| \begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix} \right. \right) = \frac{1}{2\pi\iota} \int_L g(s)\, z^{-s} ds$$

where $\iota = \sqrt{-1}$, L is a suitable contour, $z \neq 0$,

$$g(s) = \frac{\prod_{\alpha=1}^{m} \Gamma(b_\alpha + s) \prod_{\alpha=1}^{n} \Gamma(1 - a_\alpha - s)}{\prod_{\alpha=m+1}^{q} \Gamma(1 - b_\alpha - s) \prod_{\alpha=n+1}^{p} \Gamma(a_\alpha + s)},$$

where m, n, p and q are integers with $0 \leq n \leq p$ and $0 \leq m \leq q$. The parameters $a_1, \ldots, a_p$ and $b_1, \ldots, b_q$ are complex numbers such that no pole of $\Gamma(b_\alpha + s)$, $\alpha = 1, \ldots, m$ coincides with any pole of $\Gamma(1 - a_\beta - s)$, $\beta = 1, \ldots, n$. The empty product is interpreted as 1. For more detail, the reader is referred to Mathai (1993, Definition 2.1, pp 60) [25].

A.9 Summary

In this appendix it was seen how the elements of an abstract division algebra, in particular the complex numbers and quaternions, may be replaced by matrices. General results that are found in the literature on quaternions, when utilising this representation approach, were given. Finally, this Appendix concluded with various functions and polynomials that are required in the development of the quaternion distribution theory.

Appendix B

Acronyms

This appendix contains a list of the most common acronyms and abbreviations used throughout this work.m They are listed alphabetically and typeset in bold, with the meaning of the acronym or abreviation alongside:

cf characteristic function

pdf probability density function

Appendix C

Symbols

The following list of symbols and notational conventions are used throughout this work.

C.1 Spaces and operators

$\mathbb{R}$	The set of all real numbers.
$\mathbb{C}$	The set of all complex numbers.
$\mathbb{Q}$	Set of all quaternions.
$\mathcal{R}$	Ring.
D	Devision algebra.
$\mathbb{P}$	Field.
G	Arbitrary division algebra.
$B(\mathbb{Q}) = B$	Borel set defined on the elements of $\mathbb{Q}$.
$B(\mathbb{R}) = B_0$	Borel set defined on the elements of $\mathbb{R}$.
$\mathbb{R}^4$	4-dimensional real space.
$\mathbb{R}^{4p}$	$4p$-dimensional real space.
$M_n(\mathbb{R})$	Collection of $n \times n$ matrices with real entries.
$M_{n \times p}(\mathbb{R})$	Collection of $n \times p$ matrices with real entries.
$M_n(\mathbb{Q})$	Collection of $n \times n$ matrices with quaternion entries.
$M_{n \times p}(\mathbb{Q})$	Collection of $n \times p$ matrices with quaternion entries.
$\underset{n \times p}{\mathbf{R}} = [r_{st}]$	A $n \times p$ real matrix with r_{st} as the st^{th} element.
$\underset{n \times p}{\mathbf{Q}} = [q_{st}]$	A $n \times p$ quaternion matrix with q_{st} as the st^{th} element.
$\mathbf{A}^{-1}$	The inverse of the real, complex or quaternion matrix $\underset{p \times p}{\mathbf{A}}$.
$\mathbf{A}'$	The transpose of the real, complex or quaternion matrix $\underset{p \times p}{\mathbf{A}}$.
$\det(\mathbf{A})$	The determinant of the real, complex or quaternion matrix $\underset{p \times p}{\mathbf{A}}$.
$\lvert\det(\mathbf{A})\rvert$	The absolute value of the determinant of the real, complex or quaternion matrix $\underset{p \times p}{\mathbf{A}}$.
$\operatorname{tr}\mathbf{A}$	The trace of the real, complex or quaternion matrix $\underset{p \times p}{\mathbf{A}}$.
$\operatorname{Retr}\mathbf{A}$	The trace of the real component of the real, complex or quaternion matrix $\underset{p \times p}{\mathbf{A}}$.
$\bar{a}$	The conjugate of a complex or quaternion number.
$\bar{\mathbf{A}}$	Conjugate matrix.
$\simeq$	Isomorphic to.
$i,\ j,\ k,\ \iota$	Imaginary root of -1.
$\otimes$	Kronecker product.
vec	Vec operator.

C.2 Vectors and matrices

$\mathbf{D}_\lambda$	Diagonal matrix with diagonal entries the characteristic roots of the quaternion Hermitian matrix $\mathbf{Q}$.
$\mathbf{R}^{\mathbf{D}}$	Diagonal matrix such that $\mathbf{R}^{\mathbf{D}} \simeq \mathbf{D}_\lambda$.
$\bar{\mathbf{H}}' = \mathbf{H}^{-1}$	Symplectic matrix in the quaternion case.

C.3 Distributions and functions

$\mathbb{Q}N(\mu, \sigma^2)$	Univariate quaternion normal distribution with mean μ and variance σ^2.
$\mathbb{Q}N(p; \underline{\mu}, \mathbf{\Sigma})$	Multivariate quaternion normal distribution with mean $\underline{\mu}$ and covariance matrix $\mathbf{\Sigma}$.
$\mathbb{Q}N\left(n \times p; \mu, \mathbf{\Sigma} \otimes \mathbf{R}\right)$	Matrix-variate quaternion normal distribution with mean μ, with $\mathbf{\Sigma}$ the covariance matrix of the columns, and $\mathbf{R}$ a real covariance matrix of the rows.
χ_p	Chi-squared distribution with p degrees of freedom.
$\mathbb{Q}W_p\left(\mathbf{\Sigma}, n\right)$	The quaternion Wishart distribution, with n degrees of freedom and covariance matrix $\mathbf{\Sigma}$
$\mathbb{Q}W_p\left(\mathbf{\Sigma}, n, \mathbf{\Omega}\right)$	The non-central quaternion Wishart distribution, with n degrees of freedom, covariance matrix $\mathbf{\Sigma}$ and non-centrality parameter $\mathbf{\Omega}$.
$\mathbb{Q}B1(p, n_1, n_2)$	The quaternion matrix-variate beta type I distribution with n_1 and n_2 degrees of freedom
$\mathbb{Q}B1(p, n_1, n_2, \mathbf{\Omega})$	The non-central quaternion matrix-variate beta type I distribution with n_1 and n_2 degrees of freedom and non-centrality parameter $\mathbf{\Omega}$.
$\Gamma\left(\cdot\right)$	The real gamma function.
$\mathbb{Q}\Gamma_p\left(\cdot\right)$	The quaternion multivariate gamma function.
$\mathbb{Q}B_p\left(\cdot\right)$	The quaternion multivariate beta function.
${}_r\mathbb{Q}F_s\left(\cdot\right)$	The quaternion generalised hypergeometric function.
$\mathbb{Q}C_\kappa(\mathbf{A})$	The zonal polynomial of the quaternion Hermitian matrix $\mathbf{A}$.
$M_f\left(\cdot\right)$	The Mellin transform.
$G_{p,q}^{m,n}\left(\cdot\right)$	Meijer's G-function.

C.4 Variables and observations

Z	A quaternion stochastic variable.
z	A value of Z.
$\underset{p\times 1}{\underline{Z}} = [Z_1, \ldots, Z_p]'$	A quaternion stochastic vector (p-dimensional vector with quaternions as elements.
$\underset{p\times 1}{\underline{z}} = [z_1, \ldots, z_p]'$	An observed value of $\underset{p\times 1}{\underline{z}}$.
$X_1,\ X_2,\ X_3,\ X_4$	Real stochastic variables.
$x_1,\ x_2,\ x_3,\ x_4$	Observed values of X_1, X_2, X_3, X_4.
$\underset{p\times 1}{\underline{X}_1},\ \underset{p\times 1}{\underline{X}_2},\ \underset{p\times 1}{\underline{X}_3},\ \underset{p\times 1}{\underline{X}_4},\ \underset{4p\times 1}{\underline{Z}_0}$	Real stochastic vectors (vectors with real stochastic variables).
$\underset{p\times 1}{\underline{x}_1},\ \underset{p\times 1}{\underline{x}_2},\ \underset{p\times 1}{\underline{x}_3},\ \underset{p\times 1}{\underline{x}_4},\ \underset{4p\times 1}{\underline{z}_0}$	Observed values of $\underset{p\times 1}{\underline{X}_1}$, $\underset{p\times 1}{\underline{X}_2}$, $\underset{p\times 1}{\underline{X}_3}$, $\underset{p\times 1}{\underline{X}_4}$, $\underset{4p\times 1}{\underline{Z}_0}$.
$\underset{4\times 1}{Z_0}$	Associated real variable values of Z.
$\underset{4\times 1}{Z_0^*}$	Associated real variable values of Z (rearranged).
$\underset{4p\times 1}{\underline{Z}_0}$	Associated real variable values of $\underline{Z}$.
$\underset{4p\times 1}{\underline{Z}_0^*}$	Associated real variable values of $\underline{Z}$ (rearranged).
$\underset{n\times p}{\mathbf{Z}}$	A quaternion stochastic matrix ($n \times p$-dimensional matrix with quaternions as elements.
$\underset{n\times 4p}{\mathbf{Z}_0^*}$	Associated real matrix of $\mathbf{Z}$.

C.5 Average values

avg X	The average value of $X_1, \ldots, X_n$, i.e. $\frac{1}{n}\sum_{s=1}^{n} X_s$ with avg x the observed value of avg X.
avg Z	The average value of $Z_1, \ldots, Z_n$, i.e. $\frac{1}{n}\sum_{s=1}^{n} Z_s$ with avg z the observed value of avg Z.
avg $\underline{Z}$	The average value of $\underline{Z}_1, \ldots, \underline{Z}_n$, i.e. $\frac{1}{n}\sum_{s=1}^{n} \underline{Z}_s$ with avg $\underline{z}$ the observed value of avg $\underline{Z}$.
avg Z_0^*	The average value of $Z_{01}^*, \ldots, Z_{0n}^*$, i.e. $\frac{1}{n}\sum_{s=1}^{n} Z_{0s}^*$ with avg z_0^* the observed value of avg Z_0^*.
avg $\underline{Z}_0^*$	The average value of $\underline{Z}_{01}^*, \ldots, \underline{Z}_{0n}^*$, i.e. $\frac{1}{n}\sum_{s=1}^{n} \underline{Z}_{0s}^*$ with avg $\underline{z}_0^*$ the observed value of avg $\underline{Z}_0^*$.

Index