October 5, 2013
Posted at http://socrates.berkeley.edu/~schwrtz/Linear.pdf

Linear Equations for Noncommutative Algebras

Charles Schwartz*

Department of Physics, University of California
Berkeley, California 94720

Abstract

Starting from the known solution of the 2-term linear equation for a quaternionic variable, aX+Xb = r, we show how to solve the 3-term equation, aX + Xb + cXd = r. Then we show how to reduce any longer linear equation to a solvable form. This technique is extended to any system of linear equations in several unknowns over the quaternions. Finally, we suggest how this approach may be extended to linear equations over other hypercomplex algebras.

*E-mail: schwartz@physics.berkeley.edu

1 Introduction

The solution of linear equations $ax = b$, in one or many variables, is an elementary school subject, so long as one stays within a commutative algebraic system, such as the real or complex numbers. However, when one starts looking at noncommutative numbers and variables, then things get more difficult, since ax is not the same as xa, yet they are both seen as linear expressions.

Earlier work [3] with noncommuting algebras has focused on linear equations such as,

$$\sum_{n=1}^{N} A_{mn}\, X_n = R_m \quad , \tag{1.1}$$

where all the given variables A_{mn} stand on one side of the unknown variables X_n. This is solved with appropriate modifications of familiar rules for matrix manipulations. Further studies of that type continue to this day.[7]

However, the more difficult problem involves multiplications on both sides of the unknown X.

In this paper we focus mostly on quaternionic variables, starting in Section 2 with the simplest nontrivial equation $aX + Xb = c$, having two distinct terms on the left hand side. This has been solved before and we focus on a particular method for reaching that solution. That method is then extended in Section 3 to solve a three term quaternionic equation, $aX + Xb + cXd = r$. Then, in Section 4, we add a new twist and achieve a general solution of the four term and larger equations of this type.

These results for equations with a single unknown X lead us to note, in Section 6, how one may handle many-variable linear equations over the quaternions. Finally, in Section 7, we speculate on how this analytical technique may be extended to other noncommutative algebras.

It is well known that any such equations can be replaced by a system of real equations, whose solution may be calculated by standard means; but the approach here is to work, as much as possible, within the hypercomplex algebra.

2 Two Term Equation for Quaternions

We deal here with quaternionic variables (over the real numbers),

$$a = a_0 + ia_1 + ja_2 + ka_3, \qquad X = X_0 + iX_1 + jX_2 + kX_3, \qquad etc., \tag{2.1}$$

and because of the noncommutativity of the basic quaternions i, j, k we must be careful of the order of factors in any multiplication.

The simplest nontrivial linear equation has two terms and is written, in canonical form,

$$aX + Xb = r, \tag{2.2}$$

where the quaternionic variables a, b, r are given and we seek to solve for the unknown quaternionic variable X.

The solution to Eq. (2.2) has been found by various people. I do not know the earliest published record of this but will cite my own [6]. Some authors refer to this as Sylvester's equation; but that may apply to the larger case when the quantities are all $n \times n$ matrices. The line of present investigation is to work directly within the algebra of quaternionic variables. Here is the solution.

$$X = D^{-1}(ar + rb^*), \qquad D = a^2 + a(b + b^*) + bb^*. \tag{2.3}$$

The symbol (*) stands for complex conjugation and it means changing the sign of all imaginary parts of any quaternionic variable; it also reverses the order of factors in a product, as $(ab)^* = b^*a^*$. We recognize that $(b + b^*)$ and bb^* are each real numbers and therefore commute with everything else. The special case when $D = 0$ will often attract the attention of mathematicians [4] looking for theorems about existence and uniqueness of solutions; I will just note this as a singular situation.

It will be helpful to introduce a notation [1] [5] for a linear multiplication operator $(a|b)$ that allows us to specify one multiplication from the left and also one from the right, acting upon any variable:

$$(a|b)X = aXb, \qquad (a|b)(c|d) = (ac|db). \tag{2.4}$$

With this notation, we can describe the process of solving Eq. (2.2) as follows.

$$[(a|1) + (1|b)]X = r, \tag{2.5}$$

$$[(a|1) + (1|b^*)][(a|1) + (1|b)]X = [(a|1) + (1|b^*)]r, \tag{2.6}$$

$$[a^2X + aXb^* + aXb + Xbb^*] = ar + rb^*, \tag{2.7}$$

and this leads us to Eq. (2.3).

3 Three Term Equation for Quaternions

Now we start with a canonical form of the three term linear equation.

$$aX + Xb + cXd = r. \tag{3.1}$$

Following the method of Section 2, we apply the operator

$$[(a|1) + (1|b^*) + (c|d^*)] \tag{3.2}$$

to Eq. (3.1). This results in,

$$[a^2 + a(b + b^*) + bb^* + c(db^* + bd^*) + c^2dd^*]X + acXd + caXd^* = ar + rb^* + crd^*, \tag{3.3}$$

where we have noted that $(db^* + bd^*)$ is real. This equation (3.3) still looks like a three term equation; but we notice that the last two terms on the left hand side can be rearranged as,

$$\frac{1}{2}(ac + ca)X(d + d^*) + \frac{1}{2}(ac - ca)X(d - d^*). \tag{3.4}$$

In the first expression here we can move $(d + d^*)$ to the left of X; and this leaves us with a two term linear equation:

$$AX + BXC = R, \tag{3.5}$$

where

$$A = a^2 + a(b + b^*) + bb^* + c(db^* + bd^*) + c^2dd^* + \frac{1}{2}(ac + ca)(d + d^*), \tag{3.6}$$

$$B = \frac{1}{2}(ac - ca), \qquad C = d - d^*, \qquad R = ar + rb^* + crd^*. \tag{3.7}$$

Finally, we multiply Eq. (3.5) on the left by B^{-1} and we have the canonical form (2.2) of the two term linear equation, whose solution was given above. In the case $B = 0$, Eq. (3.5) is even simpler to solve; and there may also be singular situations as noted for the two term equation.

In a 2008 paper Janovská and Opfer [2] say that for the linear quaternionic equation with three or more terms it "seems not to be possible" to obtain a solution "just by applying quaternionic algebra". The result given here contradicts their conclusion.

4 Four and More Terms for Quaternions

The first attempt to continue this line of attack was to write a four term equation as,

$$aX + Xb + cXd + eXf = r; \tag{4.1}$$

and then apply a four term linear multiplier, $[(a|1) + (1|b^*) + (c|d^*) + (e|f^*)]$. The result was not a reduction to a three term equation: it remained a (more complicated) four term equation. It became clear that going on to even more terms by following that method would be a failure.

Something different was needed; and here is what I discovered.

A four term linear equation is the most general one can have with quaternionic variables. This may be seen by writing down the following version,

$$aX + bXi + cXj + dXk = r. \tag{4.2}$$

Now, suppose I want to add something else, say gXh. I can expand $h = h_0 + h_1 i + h_2 j + h_3 k$ and then simply add each of those individual terms to what is already seen in Eq. (4.2).

Another way of seeing this is to start with a general linear equation of any number of terms as follows.

$$\lambda(X) = \sum_p g_p \; X \; h_p = r. \tag{4.3}$$

Then we write, for each term, $h_p = h_{p,0} + h_{p,1} i + h_{p,2} j + h_{p,3} k$; and then collect the terms as follows so that we arrive at Equation (4.2).

$$a = \sum_p g_p \; h_{p,0}, \qquad b = \sum_p g_p \; h_{p,1}, \tag{4.4}$$

$$c = \sum_p g_p \; h_{p,2}, \qquad d = \sum_p g_p \; h_{p,3}. \tag{4.5}$$

Let me give this approach (4.2) a new name: the Revised Starting Form.

Now operate on Eq. (4.2) with a linear operator L composed as follows.

$$L = [(\alpha|1) + (\beta|i) + (\gamma|j) + (\delta|k)]. \tag{4.6}$$

where the four Greek letters stand for as-yet-undetermined quaternionic variables. The result is,

$$W_0 X + W_1 Xi + W_2 Xj + W_3 Xk = R = Lr, \tag{4.7}$$

$$W_0 = \alpha a - \beta b - \gamma c - \delta d, \qquad W_1 = \alpha b + \beta a - \gamma d + \delta c, \tag{4.8}$$

$$W_2 = \alpha c + \beta d + \gamma a - \delta b, \qquad W_3 = \alpha d - \beta c + \gamma b + \delta a. \tag{4.9}$$

Now, if I can make any one of the W terms vanish, then we are left with a three term equation, which can be solved as described in the previous Section. For example we can elminate W_1 by choosing $\beta = -\alpha b a^{-1}, \gamma = \delta = 0$. Many other choices are also available.

Can one make the larger leap and ask to have all three expressions, W_1, W_2, W_3 equal to zero? Then the solution would be simply, $X = W_0^{-1}\ R$. The answer is Yes. Start by setting $\alpha = 1$. Then we see that Equations (4.8, 4.9) give us three simultaneous linear equations in three unknowns β, γ, δ; and the essential point is that these unknowns all stand on the left side of the equations. Thus we can use standard techniques - such as Gaussian elimination - for solving linear equations, being careful to apply all multiplications from the right side. We have come back to the easy problem of Equation (1.1).

5 Discussion for One Quaternionic Unknown

This completes the solution of the general linear equation for one unknown quaternionic variable - a problem that has long stymied others. I do not try to write a single formula for the answer (as one sees in Eq. (2.3)); but the calculational process is now completely defined in terms of steps to be taken, all of it dealing with the given quaternionic variables. The previous Section, starting with the Revised Starting Form, can be read as providing an algorithm that can be readily programmed into a computer.

Of course, there is always the numerical (computer aided) approach, since these equations can be reduced to four simultaneous real linear equations in the unknowns X_0, X_1, X_2, X_3; but here we are trying to work within the algebra of quaternionic variables.

6 Extension to N Quaternionic Unknowns

The above equations involve only a single unknown variable X. We can readily extend this to a system of N linear equations in N unknowns. Then, the same method shown in Section 4 can be applied, with the quantities a, b, c, d replaced by NxN matrices and X, r replaced by N-component column vectors. The elements of all those are quaternionic variables. The system of equations (4.8, 4.9) is again all "one-sided" in the unknown matrices shown as Greek letters. It is all computable.

Here are the details of going from one unknown to many. Start by writing the most general system of linear equations in N unknowns as,

$$\sum_{n=1}^{N} \Lambda_{mn}(X_n) = R_m, \qquad m = 1, N. \tag{6.1}$$

Here, each $\Lambda_{mn}(X_n)$ is just like what we wrote as $\lambda(X)$ for a single unknown in (4.3) ; and so we can revise this to read,

$$\Lambda_{mn}(X_n) = a_{mn}\ X_n + b_{mn}\ X_n\ i + c_{mn}\ X_n\ j + d_{mn}\ X_n\ k. \tag{6.2}$$

This lets us write the entire system of equations as

$$\mathcal{A}X + \mathcal{B}X\ i + \mathcal{C}X\ j + \mathcal{D}X\ k = R, \tag{6.3}$$

where each one of those script letters is an NxN matrix of (given) quaternionic variables that multiply the vector X from the left. We proceed to reduce this to solvable form exactly as we did in Section 4 for a single unknown: Multiply from the left by the two-sided multiplier L from (4.6), with the four Greek letters now standing for (as yet undetermined) NxN matrices of quaternionic variables. We proceed, exactly as in Section 4, to eliminate the three W terms with imaginary quaternions standing to the right of X. The resulting equations now are seen as a 3x3 matrix of NxN matrices, with all of the entries being quaternions that multiply in only one direction; all of this is calculable.

7 Extension to Other Hypercomplex Algebras

Can we take these results for quaternionic variables and extend them for other noncommutative (but associative) algebras? Let's try to do this, sketchily.

Let the algebra be based upon a set of unit elements e_i, $i = 1, n$. Then we have our variables, known and unknown, written as

$$a = a_0 + \sum_{i=1}^{n} e_i\ a_i, \qquad X = X_0 + \sum_{i=1}^{n} e_i\ X_i, \quad etc. \tag{7.1}$$

We also assume we are given the multiplication table for those elements:

$$e_i\, e_j = \sum_k f_{i,j}^k\, e_k, \tag{7.2}$$

where the constants f are numbers that commute with everything else. With this, it seems that the analysis of Section 4 can be completely reproduced, starting with n+1 terms instead of four in Equation (4.2).

Note that the particular solutions for quaternions, found in Sections 2 and 3, may not be applicable in general. However, the general technique presented in the latter part of Section 4 should work.

Acknowledgment

I am grateful to Gerhard Opfer, University of Hamburg, Germany, for some stimulating exchanges.

References

[1] S. De Leo and P. Rotelli, *The quaternionic Dirac Lagrangian*, Mod. Phys. Lett. **A 11** (1996) 357.

[2] D. Janovská and G. Opfer, *Linear equations in quaternionic variables*, Mitt. Math. Ges. Hamburg **27** (2008) 223-234.

[3] O. Ore, *Linear equations in non-commutative fields*, Annals of Mathematics, Second Series **Vol. 32 No. 3** (1931) 463.

[4] Results of various online searches.

[5] C. Schwartz, *Relativistic quaternionic wave equation*, J. Math. Phys. **47** (2006) 122301.

[6] C. Schwartz, *Analytic functions of a quaternionic variable*, arXiv:0803.3782v1 [math.FA] (2008).

[7] G. J. Song, *Characterization of the W-weighted Drazin inverse over the quaternion skew field with applications*, Elec. J. Linear Algebra **26** (2013) 1.