

Grassmann Mechanics, Multivector Derivatives and Geometric Algebra

AUTHORS

Chris Doran

Anthony Lasenby

Stephen Gull

In Z. Oziewicz, A. Borowiec and B. Jancewicz, editors
Spinors, Twistors, Clifford Algebras and Quantum Deformations
(Kluwer Academic, Dordrecht, 1993), p. 215-226

Abstract

A method of incorporating the results of Grassmann calculus within the framework of geometric algebra is presented, and shown to lead to a new concept, the multivector Lagrangian. A general theory for multivector Lagrangians is outlined, and the crucial role of the multivector derivative is emphasised. A generalisation of Noether's theorem is derived, from which conserved quantities can be found conjugate to discrete symmetries.

1 Introduction

Grassmann variables enjoy a key role in many areas of theoretical physics, second quantization of spinor fields and supersymmetry being two of the most significant examples. However, not long after introducing his anticommuting algebra, Grassmann himself [1] introduced an inner product which he unified with his exterior product to give the familiar Clifford multiplication rule

$$ab = a \cdot b + a \wedge b. \quad (1)$$

What is surprising is that this idea has been lost to future generations of mathematical physicists, none of whom (to our knowledge) have investigated the possibility of recovering this unification, and thus viewing the results of Grassmann algebra as being special cases of the far wider mathematics that can be carried out with geometric (Clifford) algebra [2].

There are a number of benefits to be had from this shift of view. For example it becomes possible to “geometrize” Grassmann algebra, that is, give the results a significance in real geometry, often in space or spacetime. Also by making available the associative Clifford product, the possibility of generating new mathematics is opened up, by taking Grassmann systems further than previously possible. It is an example of this second possibility that we will illustrate in this paper.

A detailed introduction to these ideas is contained in [3], which is the first of a series of papers [4, 5, 6, 7] in which we aim to show that many of concepts of modern physics, including 2-spinors, twistors, Grassmann dimensions, supersymmetry and internal symmetry groups, can be expressed purely in terms of the real geometric algebras of space and spacetime. This, coupled with David Hestenes’ demonstration that the Dirac and Pauli equations can also be expressed in the same real algebras [8], has led us to believe that these algebras (with multiple copies for many particles) are all that are required for fundamental physics.

This paper starts with a brief survey of the translation between Grassmann and geometric algebra, which is used to motivate the concept of a multivector Lagrangian. The rest of the paper develops this concept, making full use of the multivector derivative [2]. The point to stress is that as a result of the translation we have gained something new, which can then only be fully developed outside Grassmann algebra, within the framework of geometric algebra. This is possible because geometric algebra provides a richer algebraic structure than pure Grassmann algebra.

Throughout we have used most of the conventions of [2], so that vectors are written in lower case, and multivectors in upper case. The Clifford product of the multivectors A and B is written as AB . The subject of Clifford algebra suffers from a nearly stifling plethora of conventions and notations, and we have settled on the one that, if it is not already the most popular, we believe should be. A full introduction to our conventions is provided in [3].

2 Translating Grassmann Algebra into Geometric Algebra

Given a set of n Grassmann generators $\{\zeta_i\}$, satisfying

$$\{\zeta_i, \zeta_j\} = 0, \quad (2)$$

we can map these into geometric algebra by introducing a set of n independent vectors $\{e_i\}$, and replacing the product of Grassmann variables by the exterior product,

$$\zeta_i \zeta_j \leftrightarrow e_i \wedge e_j. \quad (3)$$

In this way any combination of Grassmann variables can be replaced by a multivector. Note that nothing is said about the interior product of the e_i vectors, so the $\{e_i\}$ frame is completely arbitrary.

In order for the above scheme to have computational power, we need a translation for the second ingredient that is crucial to modern uses of Grassmann algebra, namely Berezin calculus [9]. Looking at differentiation first, this is defined by the rules,

$$\frac{\partial \zeta_j}{\partial \zeta_i} = \delta_{ij} \quad (4)$$

$$\zeta_j \overleftarrow{\partial}_{\zeta_i} = \delta_{ij}, \quad (5)$$

(together with the graded Leibnitz rule). This can be handled entirely within the algebra generated by the $\{e_i\}$ frame by introducing the reciprocal frame $\{e^i\}$, defined by

$$e_i \cdot e^j = \delta_i^j. \quad (6)$$

Berezin differentiation is then translated to

$$\frac{\partial}{\partial \zeta_i} \leftrightarrow e^i \cdot (\quad) \quad (7)$$

so that

$$\frac{\partial \zeta_j}{\partial \zeta_i} \leftrightarrow e^i \cdot e_j = \delta_j^i. \quad (8)$$

Note that we are using lower and upper indices to distinguish a frame from its reciprocal, rather than to simply distinguish metric signature.

Integration is defined to be equivalent to right differentiation, i.e.

$$\int f(\zeta) d\zeta_n d\zeta_{n-1} \dots d\zeta_1 = f(\zeta) \overleftarrow{\frac{\partial}{\partial \zeta_n}} \overleftarrow{\frac{\partial}{\partial \zeta_{n-1}}} \dots \overleftarrow{\frac{\partial}{\partial \zeta_1}}. \quad (9)$$

In this expression $f(\zeta)$ translates to a multivector F , so the whole expression becomes

$$(\dots ((F \cdot e^n) \cdot e^{n-1}) \dots) \cdot e^1 = \langle FE^n \rangle, \quad (10)$$

where E^n is the pseudoscalar for the reciprocal frame,

$$E^n = e^n \wedge e^{n-1} \dots \wedge e^1, \quad (11)$$

and $\langle FE^n \rangle$ denotes the scalar part of the multivector FE^n .

Thus we see that Grassmann calculus amounts to no more than Clifford contraction, and the results of “Grassmann analysis” [10, 9] can all be expressed as simple algebraic identities for multivectors. Furthermore these results are now given a firm geometric significance through the identification of Clifford elements with directed line, plane segments *etc.* Further details and examples of this are given in [3].

It is our opinion that this translation shows that the introduction of Grassmann variables to physics is completely unnecessary, and that instead genuine Clifford entities should be employed. This view results not from a mathematical prejudice that Clifford algebras are in some sense “more fundamental” than Grassmann algebras (such statements are meaningless), but is motivated by the fact that physics clearly does involve Clifford algebras at its most fundamental level (the electron). Furthermore, we believe that a systematic use of the above translation would be of great benefit to areas currently utilising Grassmann variables, both in geometrizing known results, and, more importantly, opening up possibilities for new mathematics. Indeed, if new results cannot be generated, the above exercise

would be of very limited interest.

It is one of the possibilities for new mathematics that we wish to illustrate in the rest of this paper. The idea has its origin in pseudoclassical mechanics, and is illustrated with one of the simplest Grassmann Lagrangians,

$$L = \frac{1}{2}\dot{\zeta}_i\dot{\zeta}_i - \frac{1}{2}\epsilon_{ijk}\omega_i\dot{\zeta}_j\dot{\zeta}_k, \quad (12)$$

where ω_i are a set of three scalar constants. This Lagrangian is supposed to represent the “pseudoclassical mechanics of spin” [11, 12]. Following the above procedure we translate this to

$$L = \frac{1}{2}e_i\wedge\dot{e}_i - \omega, \quad (13)$$

where

$$\omega = \omega_1(e_2\wedge e_3) + \omega_2(e_3\wedge e_1) + \omega_3(e_1\wedge e_2), \quad (14)$$

which gives a *bivector* valued Lagrangian. This is typical of Grassmann Lagrangians, and can be easily extended to supersymmetric Lagrangians, which become mixed grade multivectors. This raises a number of interesting questions; what does it mean when a Lagrangian is multivector-valued, and do all the usual results for scalar Lagrangians still apply? In the next section we will provide answers to some of these, illustrating the results with the Lagrangian of (13). In doing so we will have thrown away the origin of the Lagrangian in Grassmann algebra, and will work entirely within the framework geometric algebra, where we hope it is evident that the possibilities are far greater.

3 The Variational Principle for Multivector Lagrangians

Before proceeding to derive the Euler-Lagrange equations for a multivector Lagrangian, it is necessary to first recall the definition of the multivector derivative ∂_X , as introduced in [13, 2]. Let X be a mixed-grade multivector

$$X = \sum_r X_r, \quad (15)$$

and let $F(X)$ be a general multivector valued function of X . The A derivative of F is defined by

$$A*\partial_X F(X) = \left. \frac{\partial}{\partial\tau} F(X + \tau A) \right|_{\tau=0}, \quad (16)$$

where $*$ denotes the scalar product

$$A*B = \langle AB \rangle. \quad (17)$$

We now introduce an arbitrary vector basis $\{e_j\}$, which is extended to a basis for the entire algebra $\{e_J\}$, where J is a general index. The multivector derivative is defined by

$$\partial_X = \sum_J e^J e_J * \partial_X. \quad (18)$$

∂_X thus inherits the multivector properties of its argument X , so that in particular it contains the same grades. A simple example of a multivector derivative is when X is just a position vector x , in which case ∂_x is the usual vector derivative (sometimes referred to as the Dirac operator). A special case is provided when the argument is a scalar, α , when we continue to write ∂_α .

A useful result of general applicability is

$$\partial_X \langle XA \rangle = P_X(A) \quad (19)$$

where $P_X(A)$ is the projection of A onto the terms containing the same grades as X . More complicated results can be derived by expanding in a basis, and repeatedly applying (19).

Now consider an initially scalar-valued function $L = L(X_i, \dot{X}_i)$ where X_i are general multivectors, and \dot{X}_i denotes differentiation with respect to time. We wish to extremise the action

$$S = \int_{t_1}^{t_2} dt L(X_i, \dot{X}_i). \quad (20)$$

Following e.g. [14], we write,

$$X_i(t) = X_i^0(t) + \epsilon Y_i(t) \quad (21)$$

where Y_i is a multivector containing the same grades as X_i , ϵ is a scalar, and X_i^0 represents the extremal path. With this we find

$$\partial_\epsilon S = \int_{t_1}^{t_2} dt \left(Y_i * \partial_{X_i} L + \dot{Y}_i * \partial_{\dot{X}_i} L \right) \quad (22)$$

$$= \int_{t_1}^{t_2} dt Y_i * \left(\partial_{X_i} L - \partial_t (\partial_{\dot{X}_i} L) \right) \quad (23)$$

(summation convention implied), and from the usual argument about stationary

paths, we can read off the Euler-Lagrange equations

$$\partial_{X_i} L - \partial_t(\partial_{\dot{X}_i} L) = 0. \quad (24)$$

We now wish to extend this argument to a multivector-valued L . In this case taking the scalar product of L with an arbitrary constant multivector A produces a scalar Lagrangian $\langle LA \rangle$, which generates its own Euler-Lagrange equations,

$$\partial_{X_i} \langle LA \rangle - \partial_t(\partial_{\dot{X}_i} \langle LA \rangle) = 0. \quad (25)$$

An ‘allowed’ multivector Lagrangian is one for which the equations from each A are mutually consistent. This has the consequence that if L is expanded in a basis, each component is capable of simultaneous extremisation.

From (25), a necessary condition on the dynamical variables is

$$\partial_{X_i} L - \partial_t(\partial_{\dot{X}_i} L) = 0. \quad (26)$$

For an allowed multivector Lagrangian this equation is also *sufficient* to ensure that (25) is satisfied for all A . We will take this as part of the definition of a multivector Lagrangian. To see how this can work, consider the bivector-valued Lagrangian of (13). From this we can construct the scalar Lagrangian $\langle LB \rangle$, where B is a bivector, and we can derive the equations of motion

$$\partial_{e_i} \langle LB \rangle - \partial_t(\partial_{\dot{e}_i} \langle LB \rangle) = 0 \quad (27)$$

$$\Rightarrow (\dot{e}_i + \epsilon_{ijk} \omega_j e_k) \cdot B = 0. \quad (28)$$

For this to be satisfied for all B , we simply require that the bracket vanishes. If instead we use (26), together with the 3- d result

$$\partial_a a \wedge b = 2a, \quad (29)$$

we find the equations of motion

$$\dot{e}_i + \epsilon_{ijk} \omega_j e_k = 0. \quad (30)$$

Thus, for the Lagrangian of (13), equation (26) is indeed sufficient to ensure that (27) is satisfied for all B .

Recalling (14), equations (30) can be written compactly as [3]

$$\dot{e}_i = e^i \cdot \omega, \quad (31)$$

which are a set of three coupled vector equations — nine scalar equations for nine unknowns. This illustrates how multivector Lagrangians have the potential to package up large numbers of equations into a single entity, in a highly compact manner. Equations (31) are studied and solved in [3].

This example also illustrates a second point, which is that, for a fixed A , (25) does not always lead to the full equations of motion. It is only by allowing A to vary that we arrive at (26). Thus it is crucial to the formalism that L is a multivector, and that (25) holds for all A , as we shall see in the following section, where we consider symmetries.

4 Noether's Theorem for Multivector Lagrangians

One of the most powerful ways of analysing the equations of motion resulting from a Lagrangian is via the symmetry properties of the Lagrangian itself. The general tool for doing this is Noether's theorem, and it is important that an analogue of this can be found for the case of multivector Lagrangians. There turn out to be two types of symmetry to be considered, depending on whether the transformation of variables is governed by a scalar or by a multivector parameter. We will look at these separately.

It should be noted that as all our results are expressed in the language of geometric algebra, we are explicitly working in a *coordinate-free* way, and thus all the symmetry transformations considered are *active*. Passive transformations have no place in this scheme, as the introduction of an arbitrary coordinate system is an unnecessary distraction.

4.1 Scalar Controlled Transformations

Given an allowed multivector Lagrangian of the type $L = L(X_i, \dot{X}_i)$, we wish to consider variations of the variables X_i controlled by a single scalar parameter, α . We thus write $X'_i = X'_i(X_i, \alpha)$, and define $L' = L(X'_i, \dot{X}'_i)$, so that L' has the same functional dependence as L . Making use of the identity $L' = \langle L' A \rangle \partial_A$, we proceed

as follows:

$$\partial_\alpha L' = (\partial_\alpha X'_i) * \partial_{X'_i} \langle L' A \rangle \partial_A + (\partial_\alpha \dot{X}'_i) * \partial_{\dot{X}'_i} \langle L' A \rangle \partial_A \quad (32)$$

$$= (\partial_\alpha X'_i) * \left(\partial_{X'_i} \langle L' A \rangle - \partial_t (\partial_{\dot{X}'_i} \langle L' A \rangle) \right) \partial_A + \partial_t \left((\partial_\alpha X'_i) * \partial_{\dot{X}'_i} L' \right). \quad (33)$$

If we now assume that the equations of motion are satisfied for the X'_i (which must be checked for any given case), we have

$$\partial_\alpha L' = \partial_t \left((\partial_\alpha X'_i) * \partial_{\dot{X}'_i} L' \right), \quad (34)$$

and if L' is independent of α , the corresponding conserved current is $(\partial_\alpha X'_i) * \partial_{\dot{X}'_i} L'$. Note how important it was in deriving this that (25) be satisfied for all A . Equation (34) is valid whatever the grades of X_i and L , and in (34) there is no need for α to be infinitesimal. If L' is not independent of α , we can still derive useful consequences from,

$$\partial_\alpha L'|_{\alpha=0} = \partial_t \left((\partial_\alpha X'_i) * \partial_{\dot{X}'_i} L' \right) \Big|_{\alpha=0}. \quad (35)$$

As a first application of (35), consider time translation,

$$X'_i(t, \alpha) = X_i(t + \alpha) \quad (36)$$

$$\Rightarrow \partial_\alpha X'_i|_{\alpha=0} = \dot{X}_i, \quad (37)$$

so (35) gives (assuming there is no explicit time-dependence in L)

$$\partial_t L = \partial_t (\dot{X}_i * \partial_{\dot{X}_i} L). \quad (38)$$

Hence we can define the conserved Hamiltonian by

$$H = \dot{X}_i * \partial_{\dot{X}_i} L - L. \quad (39)$$

Applying this to (13), we find

$$H = \dot{e}_i * \partial_{\dot{e}_i} L - L \quad (40)$$

$$= \frac{1}{2} e_i \wedge \dot{e}_i - L \quad (41)$$

$$= \omega, \quad (42)$$

so the Hamiltonian is, of course, a bivector, and conservation implies that $\dot{\omega} = 0$,

which is easily checked from the equations of motion.

There are two further applications of (35) that are worth detailing here. First, consider dilations

$$X'_i = e^\alpha X_i, \quad (43)$$

so (35) gives

$$\partial_\alpha L'|_{\alpha=0} = \partial_t(X_i * \partial_{\dot{X}_i} L). \quad (44)$$

For the Lagrangian of (13), $L' = e^{2\alpha} L$, and we find that

$$2L = \partial_t(\tfrac{1}{2}e_i \wedge e_i) \quad (45)$$

$$= 0, \quad (46)$$

so when the equations of motion are satisfied, the Lagrangian vanishes. This is quite typical of first order Lagrangians.

Second, consider rotations

$$X'_i = e^{\alpha B/2} X_i e^{-\alpha B/2}, \quad (47)$$

where B is an arbitrary constant bivector specifying the plane(s) in which the rotation takes place. Equation (35) now gives

$$\partial_\alpha L'|_{\alpha=0} = \partial_t \left((B \times X_i) * \partial_{\dot{X}_i} L \right), \quad (48)$$

where $B \times X_i$ is one half the commutator $[B, X_i]$. Applying this to (13), we find

$$B \times L = \partial_t(\tfrac{1}{2}e_i \wedge (B \cdot e_i)). \quad (49)$$

However, since $L = 0$ when the equations of motion are satisfied, we see that

$$e_i \wedge (B \cdot e_i) \quad (50)$$

must be constant for all B . In [3] it is shown that this is equivalent to conservation of the metric tensor g , defined by

$$g(e^i) = e_i. \quad (51)$$

4.2 Multivector Controlled Transformations

The most general transformation we can write down for the variables X_i governed by a single multivector M is

$$X'_i = f(X_i, M), \quad (52)$$

where f and M are time-independent functions and multivectors respectively. In general f need not be grade preserving, which opens up a route to considering analogues of supersymmetric transformations.

In order to write down the equivalent equation to (34), it is useful to introduce the differential notation of [2],

$$A * \partial_M f(X_i, M) = \underline{f}_A(X_i, M). \quad (53)$$

We can now proceed in a similar manner to the preceding section, and derive,

$$A * \partial_M L' = \underline{f}_A(X_i, M) * \partial_{X'_i} L' + \underline{f}_A(\dot{X}_i, M) * \partial_{\dot{X}'_i} L' \quad (54)$$

$$\begin{aligned} &= \underline{f}_A(X_i, M) * \left(\partial_{X'_i} \langle L' B \rangle - \partial_t (\partial_{\dot{X}'_i} \langle L' B \rangle) \right) \partial_B + \partial_t \left(\underline{f}_A(X_i, M) * \partial_{\dot{X}'_i} L' \right) \\ &= \partial_t \left(\underline{f}_A(X_i, M) * \partial_{\dot{X}'_i} L' \right), \end{aligned} \quad (56)$$

where again we have assumed that the equations of motion are satisfied for the transformed variables. We can remove the A dependence from this by differentiating, to yield

$$\partial_M L' = \partial_t \left(\partial_A \underline{f}_A(X_i, M) * \partial_{\dot{X}'_i} L' \right), \quad (57)$$

and if L' is independent of M , the corresponding conserved quantity is

$$\partial_A \underline{f}_A(X_i, M) * \partial_{\dot{X}'_i} L' = \hat{\partial}_M f(X_i, \hat{M}) * \partial_{\dot{X}'_i} L', \quad (58)$$

where the hat on \hat{M} denotes that this is the M acted on by ∂_M . Which form of (58) is appropriate to any given problem will depend on the context. Nothing much is gained by setting $M = 0$ in (57), as usually multivector controlled transformations are not simply connected to the identity.

In order to illustrate (57), consider reflection symmetry applied to the Lagrangian of (13), that is

$$f(e_i, n) = -n e_i n^{-1} \quad (59)$$

$$\Rightarrow L' = n L n^{-1}. \quad (60)$$

Since $L = 0$ when the equations of motion are satisfied, the left hand side of (57) vanishes, and we find that

$$\frac{1}{2}\partial_a \underline{f}_a(e_i, n) \wedge (ne_i n^{-1}) \quad (61)$$

is conserved. Now

$$\underline{f}_a(e_i, n) = -ae_i n^{-1} + ne_i n^{-1} a n^{-1}, \quad (62)$$

so (61) becomes

$$\frac{1}{2}\partial_a \langle -e_i^2 a n^{-1} + ne_i n^{-1} a e_i n^{-1} \rangle_2 = -e_i^2 n^{-1} - e_i \cdot n^{-1} n e_i n^{-1} \quad (63)$$

$$= -n(e_i^2 n^{-1} + e_i \cdot n^{-1} e_i) n^{-1}. \quad (64)$$

This is basically the same as was found for rotations, and again the conserved quantity is the metric tensor g . This is no surprise since rotations can be built out of reflections, so it is natural to expect the same conserved quantities for both.

Equation (57) is equally valid for scalar Lagrangians, and for the case of reflections will again lead to conserved quantities which are those that are usually associated with rotations. For example considering

$$L = \dot{x}^2 - \omega^2 x^2, \quad (65)$$

it is not hard to show from (57) that the angular momentum $x \wedge \dot{x}$ is conserved. This shows that many standard treatments of Lagrangian symmetries [14] are unnecessarily restrictive in only considering infinitesimal transformations. The subject is richer than this suggests, but without the powerful multivector calculus the necessary formulae are simply not available.

5 Conclusions

Grassmann calculus finds a natural setting within geometric algebra, where the additional mathematical structure allows for a number of generalisations. This is illustrated by Grassmann (pseudoclassical) mechanics, which opens up a new field — that of the multivector Lagrangian. In order to carry out such generalisations, it is necessary to have available the most powerful techniques of geometric algebra. For Lagrangian mechanics it turns out that the multivector derivative fulfills this role, allowing for tremendous compactness and clarity. Elsewhere [5] the multivector derivative is developed and presented as the natural tool for the study of Lagrangian

field theory.

It is our opinion that the translation of Berezin calculus into geometric algebra will be of great benefit in other fields where Grassmann variables are routinely employed. A start on this has been made in [3, 4], but clearly the potential subject matter is vast, and much work remains.

Chris Doran would like to acknowledge his gratitude to Sidney Sussex College for financial support in attending this conference.

References

- [1] H. Grassmann. Der ort der Hamilton'schen quaternionen in der Ausdehnungslehre. *Math. Ann.*, 12:375, 1877.
- [2] D. Hestenes and G. Sobczyk. *Clifford Algebra to Geometric Calculus*. Reidel, Dordrecht, 1984.
- [3] A.N. Lasenby, C.J.L. Doran, and S.F. Gull. Grassmann calculus, pseudoclassical mechanics and geometric algebra. *J. Math. Phys.*, 34(8):3683, 1993.
- [4] A.N. Lasenby, C.J.L. Doran, and S.F. Gull. 2-spinors, twistors and supersymmetry in the spacetime algebra. In Z. Oziewicz, B. Jancewicz, and A. Borowiec, editors, *Spinors, Twistors, Clifford Algebras and Quantum Deformations*, page 233. Kluwer Academic, Dordrecht, 1993.
- [5] A.N. Lasenby, C.J.L. Doran, and S.F. Gull. A multivector derivative approach to Lagrangian field theory. *Found. Phys.*, 23(10):1295, 1993.
- [6] C.J.L. Doran, A.N. Lasenby, S.F. Gull, S.S. Somaroo, and A.D. Challinor. Spacetime algebra and electron physics. *Adv. Imag. & Elect. Phys.*, 95:271, 1996.
- [7] C.J.L. Doran, A.N. Lasenby, and S.F. Gull. States and operators in the spacetime algebra. *Found. Phys.*, 23(9):1239, 1993.
- [8] D. Hestenes. Observables, operators, and complex numbers in the Dirac theory. *J. Math. Phys.*, 16(3):556, 1975.
- [9] F.A. Berezin. *The Method of Second Quantization*. Academic Press Ltd., London, 1966.

- [10] B. de Witt. *Supermanifolds*. Cambridge University Press, 1984.
- [11] F.A. Berezin and M.S. Marinov. Particle spin dynamics as the Grassmann variant of classical mechanics. *Ann. Phys.*, 104:336, 1977.
- [12] P.G.O. Freund. *Supersymmetry*. Cambridge University Press, 1986.
- [13] D. Hestenes. Multivector calculus. *J. Math. Anal. Appl.*, 24:313, 1968.
- [14] H. Goldstein. *Classical Mechanics*. Addison–Wesley, Reading MA, 1950.