

Useless Notes #2: Sign-B-Gone

1 Grassmann's Curse

Often in physics, and sometimes in mathematics, things anticommute. Generally, every scalar, function, operator, everything that *multiplies*, has a label denoting whether it is Grassmann even or Grassmann odd, dictating whether it is supposed to act in a commutative or an anti-commutative way. I will often interchangeably use ‘bosonic’ for Grassmann even and ‘fermionic’ for Grassmann odd, but some people may complain about this.

Fermionic scalars anticommute: if a and b are Grassmann-odd, then

$$ab = -ba. \tag{1}$$

Bosonic scalars commute with everything.

The problem is that whenever you do anything more complicated, massive snowdrifts of signs start piling up as symbols start moving past each other. As just a taste, consider the graded commutator of two operators: given operators A and B , this is defined as

$$[A, B] = AB - (-)^{|A||B|}BA, \tag{2}$$

where $(-)^{|A|}$ is $+$ if A is Grassmann even and $-$ if A is Grassmann odd. Then the Jacobi relation is

$$\pm[[A, B], C] \pm [[B, C], A] \pm [[C, A], B] = 0 \tag{3}$$

for *some* choice of the three signs, but which ones are the correct ones? This calculation, specifically, is merely rather annoying. If you try anything actually involved, however, the signs will progress beyond mere annoyance and you will find you might as well be trying to swim to Cockaigne.

The actual instigator for me developing the two pieces of notation in this writeup was dealing with the signs in the Batalin–Vilkovisky (or BV) formalism, so I will be using it as a running example of the application of my notation.

In the BV formalism, there is some list $\chi^n, n = 1, \dots, N$ of ‘fields’ and a corresponding list χ_n^\ddagger of ‘anti-fields’ of the opposite Grassmann parity. Given functions $F(\chi^n, \chi_n^\ddagger)$ and $G(\chi^n, \chi_n^\ddagger)$, the BV bracket is defined as

$$(F, G) = \sum_n \left(\frac{\partial_R F}{\partial \chi^n} \frac{\partial_L G}{\partial \chi_n^\ddagger} - \frac{\partial_R F}{\partial \chi_n^\ddagger} \frac{\partial_L G}{\partial \chi^n} \right). \tag{4}$$

The ∂_L and ∂_R denote differentiation from the left and differentiation from the right, respectively. These differ because if θ_1 and θ_2 are separate fermionic variables, then

$$\frac{\partial_L(\theta_1\theta_2)}{\partial\theta_1} = \theta_2, \tag{5}$$

while

$$\frac{\partial_R(\theta_1\theta_2)}{\partial\theta_1} = -\frac{\partial_R(\theta_2\theta_1)}{\partial\theta_1} = -\theta_2. \tag{6}$$

Like the commutator, the BV bracket has a fixed behavior under swapping F and G and satisfies a Jacobi relation. Attempting to work these out directly, however, is going to be too painful to handle, which is why we are going to need to improve the notation.

2 On the Origin of Signs

Before I can lift the sword of notation to strike at this infestation of signs, I must first devote some time to the study of the life-cycle of these creatures. In particular, what principle determines the signs in expressions like the graded commutator?

It turns out that all signs are controlled by the behavior of the scalars. For example, the Grassmann parity can be diagnosed by whether it commutes or anticommutes with a generic fermionic scalar. Then, you can fix the signs in the graded commutator of operators by demanding for a generic fermionic scalar η ,

$$\eta[A, B] = [\eta A, B] = (-)^{|A|}[A\eta, B] = (-)^{|A|}[A, \eta B] = (-)^{|A|+|B|}[A, B]\eta \tag{7}$$

There is a minor problem here. There is not in fact any such thing as ‘a generic fermionic scalar’. Were this a bosonic scalar, one could check that the formula works for each real number but there’s no value you can assign to an anticommuting scalar other than zero. The resolution to this puzzle comes from the algebraic geometers, who have had to deal with the same problem.

In algebraic geometry, one focuses on the algebras of functions rather than on the points themselves. Given a commutative ring (i.e. a set of functions with addition and multiplication), there is a way to assign a set of points to it and a notion of the value of a function at a point but a function is not necessarily determined by its value at all of the points. This happens whenever there are (non-zero) functions f such that f^n vanishes for some positive integer n . This is exactly the situation that we’re dealing with here: anything fermionic squares to zero.

When Grothendieck proposed functions which aren't determined by their value at all points, the algebraic geometers of the time found it abhorrent, but like many abhorrent things, it came to be accepted since it was simply too useful for proving theorems.

The result is that we will need to do everything in families. Let one have a 'supermanifold' parameterized by the anti-commuting coordinates $\theta^1, \dots, \theta^n$. (A supermanifold is a space that can be locally parameterized by coordinates which can be either bosonic or fermionic, just as an ordinary manifold is locally parameterized by bosonic coordinates.) This produces a ring of functions R with elements of the form

$$f(\theta^i) = f^{(0)} + f_i^{(1)}\theta^i + f_{ij}^{(2)}\theta^i\theta^j + \dots, \quad (8)$$

where $f_{ijk\dots}^{(n)}$ is antisymmetric in its indices. Then, we will demand that any construction works for an arbitrary ring of scalars and is compatible with extension to a larger ring of scalars (i.e. replacing each coordinate θ^i by a function $\theta^i(\eta^a)$ for some other system of coordinates η^a).

So now given a graded vector space $V = V_0 \oplus V_1$, where V_0 denotes the bosonic elements and V_1 denotes the fermionic elements, you can construct the R -module $R \otimes V$. (A R -module is a set which possesses the structures of addition and multiplication by scalars in R .)

Given a general R -module M with an action by scalars on the left, this gives an action on the right by

$$m \cdot r = (-)^{|r||m|} r \cdot m. \quad (9)$$

This sign is forced since without it,

$$(r_1 r_2) \cdot m = m \cdot (r_1 r_2) = (m \cdot r_1) \cdot r_2 = r_2 \cdot (r_1 \cdot m) = (r_2 r_1) \cdot m, \quad (10)$$

which is wrong.

Now consider a homomorphism $\phi : N \rightarrow M$ of R -modules. In the bosonic case, such a function is required to satisfy $\phi(rm) = r\phi(m)$. A morphism can also be multiplied by a scalar giving $(r \cdot \phi)(m) = r \cdot (\phi(m))$. In the super case, these two prescriptions are contradictory. There are two resolutions: the function 'acting from the left' and the function 'acting from the right'. In the former case, one has

$$\phi(m \cdot r) = \phi(m) \cdot r \quad (r \cdot \phi)(m) = r \cdot \phi(m). \quad (11)$$

In latter case, let the action of ϕ on m be $m \circ \phi$. Then, we have

$$(r \cdot m) \circ \phi = r \cdot (m \circ \phi) \quad m \circ (\phi \cdot r) = (m \circ \phi) \cdot r \quad (12)$$

I have deliberately written these relations so that there are no explicit signs. Note that when written this way, the symbols do not change their order. There is a broader lesson here: the

general rule is that to all there is a natural order and signs always correspond precisely to the changes in order of the symbols.

What counts as a ‘symbol’ can sometimes be a bit funny: in the case of the BV bracket, the bracket swaps the Grassmann parity relative to the total parity of its arguments. One can check that the fermionic scalars act as

$$\eta(F, G) = (\eta F, G) = (-)^{|F|}(F\eta, G) = (-)^{|F|+1}(F, \eta G) = (-)^{|F|+1+|G|}(F, G)\eta, \quad (13)$$

so really, that comma is fermionic.

3 Abstract Indices

The first of our notations will be to deal with the fact that the different fields χ^n have different Grassmann parity. The problem is now that the traditional form of index notation will no longer be able to handle the signs correctly. For example, suppose that the standard Einstein summation prescription of just summing over all values of the indices worked correctly for the contractions $v_1^n w_{1,n}$ and $v_2^m w_{2,m}$. Then, we will be forced to have

$$\sum_{n,m} v_1^n w_{1,n} v_2^m w_{2,m} = \sum_{n,m} (-)^{|w_{1,n}||v_2^m|} v_1^n v_2^m w_{1,n} w_{2,m}. \quad (14)$$

So the moment you start shuffling symbols around, the signs wreck the index notation.

What we could do is to manually work out the rules for all of the signs by consistency. This would be rather difficult, so I will instead turn to using the category theory of the situation. Now, I have generally found that pure category theory is always trivial, if sometimes in a non-trivial way. Category theory can help organize calculations and help you guess what sort of things you should and shouldn’t write down, but it will not help you directly with the actual business of extracting numbers. What we’re trying to do here, however, is designing notation, which is precisely organizing calculations and determining what the allowed expressions are, so abstract nonsense will be of direct use to us here.

This list of fields is supposed to come from a basis of some super vector space V . More generally, one can let V be a (free) R -module, for some ring of scalars R . Given multiple vector spaces, V_1, V_2, \dots, V_n , one can form a tensor product

$$V_1 \otimes_R V_2 \otimes_R \cdots \otimes_R V_n. \quad (15)$$

The relative tensor product \otimes_R is like the usual one, but with the additional condition that $(v_1 \cdot r) \otimes v_2 = v_1 \otimes (r \cdot v_2)$ for all $r \in R$, not just the real/complex numbers.

There is a canonical map $V_1 \otimes_R V_2 \rightarrow V_2 \otimes_R V_1$ which preserves the R -module structure. This is the one which has the correct sign:

$$v_1 \otimes v_2 \mapsto (-)^{|v_1||v_2|} v_2 \otimes v_1. \quad (16)$$

This swap operation gives the structure of a symmetric monoidal category, which just means that this swap gives rise to consistent permutations of larger tensor products and is compatible with composition with other functions.

The essence of contraction of indices is the existence of a dual space V^* (in this case another R -module) and a contraction map $V^* \otimes_R V \rightarrow R$. Given a vector $v \in V$, let one attach an abstract index v^A to it. Then, given a dual vector $w \in V^*$, one can form the scalar contraction $v^A w_A$. Given a larger tensor product whose factors are V and V^* , the maps one can make correspond exactly to the possible contractions of the indices. The signs are forced by the symmetry rules interchanging the factors.

Now, being canonical is nice and all, but you might want to work with a basis at some point. Suppose one has a basis $e_1, \dots, e_n \in V$. Restoring the abstract index gives e_i^A , where i is a literal index that varies from 1 to n and A is the abstract index. One can also construct a left dual basis f_A^i satisfying $f_A^i e_j^A = (-)^{|i|} e_j^A f_A^i = \delta_j^i$. Here $|i| = |e_i| = |f^i|$. The right dual basis differs by multiplication by $(-)^{|i|}$, which is related to the difference between right and left differentiation.

Now, for example, one can expand the identity map in terms of basis elements

$$\delta_B^A = \sum_i e_i^A f_B^i \quad (17)$$

and check that $\delta_B^A e_i^B = e_i^A$. This explains why the supertrace of a matrix is the trace of the bosonic part minus the trace of the fermionic part: the trace of the identity is

$$\delta_A^A = \sum_i e_i^A f_A^i = \sum_i (-)^{|i|} f_A^i e_i^A = \sum_i (-)^{|i|}. \quad (18)$$

We can now deal with coordinates and taking derivatives without having to deal with the annoying sign problems. Given a function f on a patch of super-manifold parameterized by a super-vector space V , one can now take the coordinates to be x^A , where A is an abstract index for V . The advantage of this is that x is now a purely bosonic symbol and the contraction rules are now clear. One can now construct a derivative ∂_A such that

$$\partial_A(w_B x^B) = w_A. \quad (19)$$

This ∂_A is also a purely bosonic symbol.

As an application, let us return to our BV bracket. We want to assemble our list of fields χ^n and antifields χ_n^\dagger into one object. (In the more general version of the BV formalism, there is no inherent distinction between the fields and the antifields and one only keeps track of the odd symplectic structure – but what's the definition of an odd symplectic structure?)

A coordinate is a function that takes in the point and outputs a number, so it lives in the dual space. So we have dual basis elements $[\chi^n]_A$ and $[\chi_n^\dagger]_A$ along with basis elements $[\chi_n^\vee]^A$ and $[\chi^{\dagger\vee n}]^A$. Let us pick the signs such that

$$[\chi_n^\vee]^A [\chi^m]_A = \delta_n^m, \quad [\chi^{\dagger\vee n}]^A [\chi_m^\dagger]_A = \delta_m^n, \quad (20)$$

and the cross terms vanish.

Now, we can construct the combined field

$$\Psi^A = \sum_n (\chi^n [\chi_n^\vee]^A + \chi_n^\dagger [\chi^{\dagger\vee n}]^A) \quad (21)$$

so that

$$\chi^n = [\chi^n]_A \Psi^A, \quad \chi_n^\dagger = [\chi_n^\dagger]_A \Psi^A \quad (22)$$

Note that the above equations are independent of order since Ψ^A is now bosonic.

The derivative $\partial_A \chi^n = [\chi^n]_A$, so

$$\frac{\partial_L F(\Psi)}{\partial \chi^n} = [\chi_n^\vee]^A \partial_A F, \quad \frac{\partial_R F(\Psi)}{\partial \chi_n^\dagger} = (-)^{|\chi^n|} \partial_A F [\chi_n^\vee]^A \quad (23)$$

So now we can re-express our BV bracket as

$$(F, G) = \partial_A F \varepsilon^{AB} \partial_B G, \quad (24)$$

where

$$\varepsilon^{AB} = \sum_n (-)^{|\chi^n|} ([\chi_n^\vee]^A [\chi^{\dagger\vee n}]^B + [\chi^{\dagger\vee n}]^A [\chi_n^\vee]^B). \quad (25)$$

We now see that our inverse symplectic form ε^{AB} is actually symmetric under swapping its indices. (One of χ^n and χ_n^\dagger is bosonic for each n , so they commute.)

The odd symplectic form itself is

$$\eta_{AB} = \sum_n ([\chi_n^\dagger]_A [\chi^n]_B - [\chi^n]_A [\chi_n^\dagger]_B) \quad (26)$$

and satisfies

$$\eta_{AB} \varepsilon^{BC} = \delta_A^C. \quad (27)$$

This time, η_{AB} is antisymmetric in its indices. This antisymmetry is forced by

$$\eta_{AB}\varepsilon^{BC} = \delta_A^C = \varepsilon^{CB}\eta_{BA} = -\eta_{BA}\varepsilon^{BC}. \quad (28)$$

It's now pretty easy to guess the appropriate conditions on the general version of the BV formalism: you want a fermionic $\eta_{AB}(\Psi)$, antisymmetric in its indices, satisfying

$$\partial_A\eta_{BC} + \partial_B\eta_{CA} + \partial_C\eta_{AB} = 0 \quad (29)$$

and with inverse ε^{AB} giving the bracket

$$(F, G) = \partial_A F \varepsilon^{AB} \partial_B G. \quad (30)$$

Just like in the usual symplectic case, to any function F there is a vector field

$$\partial_A F \varepsilon^{AB} \hat{\partial}_B, \quad (31)$$

where I've put a hat on $\hat{\partial}_B$ to denote that this is a vector field rather than an actual derivative. There is now the relation

$$[\partial_* F \varepsilon^{**} \hat{\partial}_*, \partial_* G \varepsilon^{**} \hat{\partial}_*] = \partial_*(F, G) \varepsilon^{**} \hat{\partial}_*, \quad (32)$$

where I am now eliding neighboring indices contracted like matrix multiplication, so $w_A M^A_B \varepsilon^{BC} \eta_{CD} T_{EF}^D v^F$ becomes $w_* M^*_{**} \varepsilon^{**} \eta_{**} T_{E*}^* v^*$.

Therefore, the Jacobi relation on vector fields gives rise to the Jacobi relation of the BV bracket. (This is potentially up to a constant term, but that can be killed by choosing F and G well.) Working this out is still a bit annoying, since the interchanges of the order of the functions F and G still cause signs.

4 Sign Indicators

The way to deal with these last signs is by separating the order that the Grassmann signs see from the order of the symbols on the page. Therefore, let us introduce sign indicators of the form $\llbracket abc \rrbracket$. These indicators mean either plus or minus depending on the relative order of the symbols in the expression in the brackets to the expression outside. For example, for scalars a and b ,

$$ab = \llbracket ab \rrbracket ab = \llbracket ab \rrbracket ba, \quad (33)$$

while

$$\llbracket ba \rrbracket ab = \llbracket ba \rrbracket ba = ba = (-)^{|a||b|} ab. \quad (34)$$

The use of the sign indicators comes when the order of the symbols in the expression is fixed by other considerations. The graded commutator of two operators A and B is now

$$[A, B] = AB - (-)^{|A||B|}BA = \llbracket AB \rrbracket AB - \llbracket AB \rrbracket BA. \quad (35)$$

I will often abuse my own notation and treat the sign indicator as a separate term that can be factored out:

$$[A, B] = \llbracket AB \rrbracket (AB - BA). \quad (36)$$

When two of the same fermionic symbol come up, you will need to disambiguate them (here a subscript in double brackets) so that the relative order between the sign indicator and the expression is well defined. So if ψ is a fermionic operator, then its anticommutator with itself is

$$\{\psi, \psi\} = \llbracket \psi_{[1]} \psi_{[2]} \rrbracket \{\psi_{[1]}, \psi_{[2]}\} = \llbracket \psi_{[1]} \psi_{[2]} \rrbracket (\psi_{[1]} \psi_{[2]} - \psi_{[2]} \psi_{[1]}) = 2 \llbracket \psi_{[1]} \psi_{[2]} \rrbracket \psi_{[1]} \psi_{[2]} = 2\psi^2. \quad (37)$$

The second to last equality comes from swapping the labels on the second term and then swapping the order of symbols in the sign indicator to bring it back to $\llbracket \psi_{[1]} \psi_{[2]} \rrbracket$.

In a larger expression $\llbracket a_1 a_2 \dots \rrbracket$, swapping any two neighboring symbols a_i and a_{i+1} gives a sign of $(-)^{|a_i||a_{i+1}|}$. The effect of swapping symbols farther apart in general depends on the Grassmann parity of the symbols between them. However, if they are known to be bosonic, then they obviously have no effect under changing places and swapping two fermionic symbols always gives a minus sign no matter where they are placed.

Finally, in the course of manipulations, symbols can disappear and be created. To account for this, I will denote $\llbracket \hat{A}\hat{B}/CDEF \rrbracket$ to denote pulling $CDEF$ out from the expression (in that order) and replacing them with AB . In order for this operation to be consistent, the total Grassmann parity of AB must be the same as the total Grassmann parity of $CDEF$. Thus, if one has fermionic operators ψ and $\bar{\psi}$ satisfying $\{\psi, \bar{\psi}\} = 1$, then

$$\llbracket \dots \psi \bar{\psi} \dots \rrbracket \dots \psi \bar{\psi} \dots = \llbracket \dots \psi \bar{\psi} \dots \rrbracket \dots \left(\bar{\psi} \psi + \llbracket \hat{\psi} \hat{\bar{\psi}} / \cdot \rrbracket 1 \right) \dots \quad (38)$$

Let's apply our notation to the case of the BV bracket. First, the symmetry of the BV bracket is such that

$$\llbracket F \varepsilon G \rrbracket (F, G) = \llbracket F \varepsilon G \rrbracket \partial_A F \varepsilon^{AB} \partial_B G = \llbracket F \varepsilon G \rrbracket \partial_B G \varepsilon^{BA} \partial_A F = \llbracket F \varepsilon G \rrbracket (G, F) \quad (39)$$

The ‘comma’ of the BV bracket carries Grassmann parity, so it needs to be accounted for in the sign indicator. On the other hand, the derivatives ∂_A and the miscellaneous indices all carry no Grassmann parity due to the magic of abstract index notation and can be dropped.

Unwinding the sign indicator gives

$$(F, G) = (-)^{|F|+|G|+|F||G|}(G, F). \quad (40)$$

(The reversal of three symbols can be created by three transpositions of neighboring elements.)

Now we can check the commutator of the associated vector fields without introducing more signs than necessary. In gory detail,

$$\begin{aligned}
& \left[\partial_* F \varepsilon_{[1]}^{**} \hat{\partial}_*, \partial_* G \varepsilon_{[2]}^{**} \hat{\partial}_* \right] \\
&= \llbracket F \varepsilon_{[1]} G \varepsilon_{[2]} \rrbracket \left[\partial_* F \varepsilon_{[1]}^{**} \hat{\partial}_*, \partial_* G \varepsilon_{[2]}^{**} \hat{\partial}_* \right] \\
&= \llbracket \dots \rrbracket \left(\partial_* F \varepsilon_{[1]}^{**} \partial_* (\partial_A G \varepsilon_{[2]}^{AB}) - \partial_* G \varepsilon_{[2]}^{**} \partial_* (\partial_A F \varepsilon_{[1]}^{AB}) \right) \hat{\partial}_B \\
&= \llbracket \dots \rrbracket \left(\partial_* F \varepsilon_{[1]}^{**} \partial_A \partial_* G \varepsilon_{[2]}^{**} + \partial_* F \varepsilon_{[1]}^{**} \partial_A G \partial_* \varepsilon_{[2]}^{AB} \right. \\
&\quad \left. + \partial_* G \varepsilon_{[1]}^{**} \partial_A \partial_* F \varepsilon_{[2]}^{AB} + \partial_* G \varepsilon_{[1]}^{**} \partial_A F \partial_* \varepsilon_{[2]}^{AB} \right) \hat{\partial}_B \\
&= \llbracket \dots \rrbracket \left(\partial_A (\partial_* F \varepsilon_{[1]}^{**} \partial_* G) \varepsilon_{[2]}^{AB} \right. \\
&\quad \left. - \partial_C F \partial_D G \varepsilon_{[2]}^{AB} \partial_A \varepsilon_{[1]}^{CD} + \partial_C F \partial_A G \varepsilon_{[1]}^{CD} \partial_D \varepsilon_{[2]}^{AB} + \partial_A F \partial_C G \varepsilon_{[1]}^{CD} \partial_D \varepsilon_{[2]}^{AB} \right) \hat{\partial}_B \\
&= \llbracket \dots \rrbracket \left(\partial_A (F,_{[1]} G) \varepsilon_{[2]}^{AB} + \partial_C F \partial_D G (\varepsilon_{[1]}^{BA} \partial_A \varepsilon_{[2]}^{CD} + \varepsilon_{[1]}^{CA} \partial_A \varepsilon_{[2]}^{DB} + \varepsilon_{[1]}^{DA} \partial_A \varepsilon_{[2]}^{BC}) \right) \hat{\partial}_B \\
&= \llbracket F \varepsilon_{[1]} G \varepsilon_{[2]} \rrbracket \partial_* (F,_{[1]} G) \varepsilon_{[2]}^{**} \hat{\partial}_* \\
&= \partial_* (F, G) \varepsilon_{**} \hat{\partial}_*.
\end{aligned} \quad (41)$$

This looks a bit gnarly, but the derivation of the corresponding identity for the ordinary symplectic case has precisely the same steps. I have freely used the fact that swapping $\llbracket 1 \rrbracket$ and $\llbracket 2 \rrbracket$ in the expression (but not the sign indicator) contributes a minus sign. I have also used the identity

$$\varepsilon^{AD} \partial_D \varepsilon^{BC} + \varepsilon^{BD} \partial_D \varepsilon^{CA} + \varepsilon^{CD} \partial_D \varepsilon^{AB} = 0 \quad (42)$$

which is a consequence of the closure condition on the odd symplectic form η and the identity $\partial_A \varepsilon^{BC} = -\varepsilon^{B*} \partial_A \eta_{**} \varepsilon^{*C}$ which is a general fact about the variation of the inverse of a matrix.

We can now finally write down and derive that Jacobi relation.

$$((F, G), H) = \llbracket F \varepsilon_{[1]} G \varepsilon_{[2]} H \rrbracket \partial_* (F,_{[1]} G) \varepsilon_{[2]}^{**} \partial_* H \quad (43)$$

$$= \llbracket \dots \rrbracket \left(\partial_* F \varepsilon_{[1]}^{**} \partial_* (\partial_* G \varepsilon_{[2]}^{**} \partial_* H) - \partial_* G \varepsilon_{[2]}^{**} \partial_* (\partial_* F \varepsilon_{[1]}^{**} \partial_* H) \right) \quad (44)$$

$$= \llbracket \dots \rrbracket \left((F,_{[1]} (G,_{[2]} H)) + (G,_{[1]} (F,_{[2]} H)) \right), \quad (45)$$

so

$$\llbracket F \varepsilon_{[1]} G \varepsilon_{[2]} H \rrbracket \left(((F,_{[1]} G),_{[2]} H) + ((G,_{[1]} H),_{[2]} F) + ((H,_{[1]} F),_{[2]} G) \right) = 0. \quad (46)$$

Easy peasy. (For the appropriate definition of ‘easy’ and the appropriate definition of ‘peasy’.)