NOTATION FOR NORMAL FORMS

JAMES MURDOCK

In the book, I used a number of operations that are anti-homomorphisms (either of groups or of Lie algebras). The purpose of this note is to investigate the question: To what extent is it possible to have a notation that covers the entire range of operations used in normal form theory without introducing antihomomorphisms? For the convenience of the reader (and myself), every notation used here that also occurs in the book is defined the same way in both places. Whenever we require a new notation, a new symbol is used.

1. Let $\mathfrak{G}$ be a matrix Lie group (such as $\mathrm{GL}(n)$) and $\mathfrak{g}$ its Lie algebra (such as $\mathrm{gl}(n)$). The only consistent way to define operations that are (true) representations is as follows:

$$\mathrm{In}(T)S = TST^{-1}, \tag{0.1}$$

$$\mathrm{Ad}(T)A = TAT^{-1}, \tag{0.2}$$

$$\mathrm{ad}(A)B = [A, B] = AB - BA. \tag{0.3}$$

   Here $S, T \in \mathfrak{G}$ and $A, B \in \mathfrak{g}$. These are to be contrasted with $\mathbb{C}_T(S) = T^{-1}ST$, $\mathbb{S}_T(A) = T^{-1}AT$, and $\mathbb{L}_A B = [B, A] = BA - AB$ in the book. The operation In (for *inner automorphism*) is a group action (not anti-action) of $\mathfrak{G}$ on itself, because $\mathrm{In}(T_1T_2) = \mathrm{In}(T_1) \circ \mathrm{In}(T_2)$. Ad (for *adjoint*) is a group representation (not anti-representation) of $\mathfrak{G}$ on $\mathfrak{g}$ (regarded as a vector space). The operation ad is a Lie algebra representation of $\mathfrak{g}$ (with commutator bracket) into the Lie algebra of linear endomorphisms of $\mathfrak{g}$ (with commutator bracket), that is, $\mathrm{ad}([P, Q]) = [\mathrm{ad}(P), \mathrm{ad}(Q)]$. Each of the operations $\mathbb{C}$, $\mathbb{S}$, and $\mathbb{L}$ is an anti-object of the same kind. The reasons that I adopted the anti-objects in the book will be described later.
2. Similarly, if $\mathfrak{G}$ is a Lie group of diffeomorphisms (such as the group of germs of diffeomorphisms of $\mathbb{R}^n$ that fix the origin) and $\mathfrak{g}$ is the associated Lie algebra of vector fields (such as the germs of smooth vector fields on $\mathbb{R}^n$ with a rest point at the origin),

Date: May 12, 2003.

the only system of operations that are all representations is

$$\mathrm{In}(\psi)\varphi = \psi \circ \varphi \circ \psi^{-1}, \tag{0.4}$$

$$(\mathrm{Ad}(\psi)a)(y) = \psi'(\psi^{-1}(y))a(\psi^{-1}(y)), \tag{0.5}$$

$$(\mathrm{ad}(a)b)(x) = a'(x)b(x) - b'(x)a(x). \tag{0.6}$$

The reason for using y as the variable in the definition of Ad is that it suits the common situation in which a system $\dot{x} = a(x)$ is transformed to $\dot{y} = b(y)$ by the change of variables $y = \psi(x)$; in this case $b = \mathrm{Ad}(\psi)a$. Notice that the transformation $y = \psi(x)$ is *from the old coordinates to the new*. The book's system (with C, S, and L in place of In, Ad, and ad) was adopted in part because S matches the applied mathematicians' convention of writing coordinate transformations the other way, from the new coordinates to the old.

3. Everyone is agreed that if $a = (a_1, \dots, a_n)$ is a vector field (remember that this denotes a column vector; see page xvi of the book), then the associated differential operator $\mathcal{D}_a$ should be defined as

$$(\mathcal{D}_a f)(x) = f'(x)a(x) = a_1 \frac{\partial f}{\partial x_1} + \cdots + a_n \frac{\partial f}{\partial x_n}. \tag{0.7}$$

Likewise, everyone agrees that the Lie derivative L_a should be defined as

$$(\mathsf{L}_a b) = b'(x)a(x) - a'(x)b(x). \tag{0.8}$$

These merely state that the derivative of a function or a vector along the curve should be the rate of change as you move along the curve in the direction of its parameterization, not in the reverse direction. But it is at this point that a conflict arises that cannot be avoided. If we agree (with 1 and 2 above) that *all brackets of linear operators should be ordinary commutator brackets*, so that $[\mathcal{D}_a, \mathcal{D}_b] = \mathcal{D}_a\mathcal{D}_b - \mathcal{D}_b\mathcal{D}_a$ and similarly for L, then we have

$$\mathcal{D}_{[a,b]} = [\mathcal{D}_b, \mathcal{D}_a] \tag{0.9}$$

and

$$\mathsf{L}_{[a,b]} = [\mathsf{L}_b, \mathsf{L}_a]. \tag{0.10}$$

Thus $\mathcal{D}$ and L are fundamentally Lie algebra anti-homomorphisms, and this cannot be corrected.

The way that this conflict is commonly corrected (in that part of the dynamical systems literature that is dominated by differential geometry) is to identify each vector field a with its differential operator $\mathcal{D}_a$ and use the commutator bracket of these operators as the bracket of the vector fields. Then equation (0.9) disappears (because $[a, b]$ is never used), and equation (0.10) is replaced by an equation showing that the map $\mathcal{D}_a \mapsto \mathsf{L}_a$ is a Lie algebra homomorphism. It is common to denote $\mathcal{D}_a$ as X and L_a as $\mathrm{ad}(X)$. While this works fine for the Lie algebras, it must be noted that ad defined in this way does not belong to any family of operators In, Ad, ad in which In and Ad are group homomorphisms. One has merely pushed the conflict under the rug.

In the book, I took Lie differentiation L as fundamental. This is one reason that I took C, S, and L as the fundamental system of operations even though C and S are anti-homomorphisms.

4. What I propose now is taking In, Ad, and ad (defined as in items 1 and 2 above) as fundamental, but retaining $\mathcal{D}$ and L (defined in (0.7) and (0.8) as antihomomorphisms because of their importance. We of course have

$$\mathsf{L} = -\mathrm{ad}, \tag{0.11}$$

so it is always possible to replace L by a true homomorphism at the cost of a sign. Similarly, we can introduce a new operation δ_a by

$$(\delta_a f)(x) = -f'(x)a(x) \tag{0.12}$$

so that

$$\mathcal{D} = -\delta. \tag{0.13}$$

With this operation in hand, the important formula

$$\mathsf{L}_a(fv) = (\mathcal{D}_a f)v + f\mathsf{L}_a v \tag{0.14}$$

(through which the space of equivariants is seen to be a module over the invariants) can be written as

$$\mathrm{ad}(a)(fv) = (\delta_a f)v + f(\mathrm{ad}(v)). \tag{0.15}$$

It is then possible to do the entire theory of invariants, equivariants, and normal forms without the use of any antihomomorphisms, using only In, Ad, ad, and δ. Of course one will not always want to write the formulas this way, because of the applied significance of $\mathcal{D}$ and L. But it is easy to switch between

$\mathcal{D}$ and δ, and between L and ad, according to what works best in a given situation.

5. We illustrate by discussing the form of the homological equation, using format 2a. If $a(x)$ is normalized already through grade $j-1$, the generator u_j for the next transformation satisfies the homological equation
$$\mathsf{L}_A u_j = a_j - b_j$$
according to the book. Here $a_0(x) = Ax$, and u_j is the generator of the transformation *from the new coordinates to the old.* To write this without using antihomomorphisms, we have two options: either write
$$(\mathrm{ad}A)u_j = b_j - a_j,$$
reversing the familiar order on the right hand side, or better (because it fits with the definition of Ad, as discussed above), change u_j to be the generator of the transformation from old coordinates to new and write
$$(\mathrm{ad}A)u_j = a_j - b_j.$$
When using format 2a or 2b, there is no great concern about whether the transformation goes from new coordinates to old or vice versa, because the generator simply changes sign.
6. Finally, we discuss the effect of these notations on the triads of maps $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\}$ and $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$. Considering a nilpotent linear term $A = N$, we form the triad $\{N, M, H\}$ as usual and set $\{X, Y, Z\} = \{N, M, H\}$ for the sl(2) normal form style. (For the inner product style we would use $\{X, Y, Z\} = \{M^*, N^*, H^*\}$, which of course already introduces an antihomomorphism.) In the book we put $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\} = \{\mathcal{D}_Y, \mathcal{D}_X, \mathcal{D}_Z\}$ and $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\} = \{\mathsf{L}_Y, \mathsf{L}_X, \mathsf{L}_Z\}$; the "switch" reflects the anti-representation nature of L. Instead we could use $\{\mathcal{X}, \mathcal{Y}, \mathcal{Z}\} = \{\delta_X, \delta_Y, \delta_Z\}$ and $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\} = \{\mathrm{ad}(X), \mathrm{ad}(Y), \mathrm{ad}(Z)\}$, with no switch. Of course this is equivalent to using $\{-\mathcal{D}_X, -\mathcal{D}_Y, -\mathcal{D}_Z\}$ and $\{-\mathsf{L}_X, -\mathsf{L}_Y, -\mathsf{L}_Z\}$. But notice that the sl(2) normal form is $\ker \mathsf{L}_M = \ker \mathsf{L}_Y = \ker \mathrm{ad}(Y)$, which is $\ker \mathsf{X}$ in the book but $\ker \mathsf{Y}$ in the "new" system. So (to have the correct kernel at the top of the chains) Y should map up and X should map down. The whole system of notation using $\{X, Y, Z\}$ starting in section 2.5 was arranged to fit the way things would come out later using the book's definition of $\{\mathsf{X}, \mathsf{Y}, \mathsf{Z}\}$, so it is rather awkward to change it.

Department of Mathematics, Iowa State University