Costas J. Papachristou

TRANSFORMATIONS

AND

SYMMETRIES

This volume consists of 5 articles that deal with various kinds of transformations. Specifically,

– coordinate transformations in a linear space or manifold;

– Bäcklund transformations (BTs) relating solutions of different partial differential equations (PDEs) or different solutions of the same PDE;

– symmetry transformations of PDEs, producing new solutions from old ones by continuously varying a set of parameters;

– recursion operators as BTs relating different symmetries of a PDE; and

– transformations relating recursion operators of BT-related PDEs.

The articles may be viewed as “chapters” of a single book and it might thus be useful to be read in sequence, given that, to some extent, each article utilizes concepts and ideas introduced in the preceding articles.

On active and passive transformations

Costas J. Papachristou

Department of Physical Sciences, Hellenic Naval Academy, Piraeus, Greece

papachristou@hna.gr

The concepts of active and passive transformations on a vector space are discussed. Orthogonal coordinate transformations and matrix representations of linear operators are considered in particular.

1. Introduction

A physical situation may *appear* changing for two reasons: the physical system itself may pass from one state to another, or, the *same* state of the system may be viewed from two different points of view (e.g., by two different observers, using different frames of reference). The former case corresponds to an *"active"* view of the situation, while the latter one to a *"passive"* view.

Given that many physical quantities are vectors, of particular interest in Physics are linear transformations on vector spaces. Starting with the prototype transformation of rotation on a plane, we study both the active and the passive view of these transformations. In the case of a Euclidean space with Cartesian coordinates, a passive transformation corresponding to a change of basis is an orthogonal transformation. On the other hand, an active transformation on a vector space is produced by a linear operator, which is represented by a matrix in a given basis. A change of basis, leading to a different representation, is a passive transformation on this space.

2. Active view of transformations

Consider the *xy*-plane with Cartesian coordinates (x, y) and basis unit vectors $\{\hat{u}_x, \hat{u}_y\}$. We call $\mathbf{R}(\theta)$ the rotation operator on this plane, i.e., the operator which rotates any vector $\vec{A}$ on the plane by an angle θ (see Fig. 2.1; by convention, $\theta>0$ for counterclockwise rotation while $\theta<0$ for clockwise rotation). This operator is linear, given that adding two vectors and then rotating the sum is the same as first rotating the vectors and then adding them.

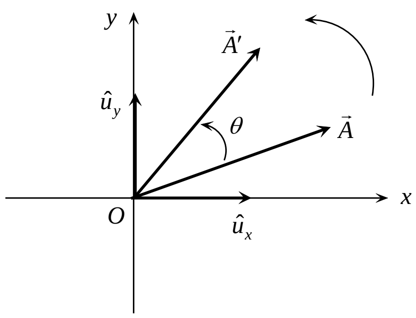

Figure 2.1

Imagine, in particular, that we rotate each vector in the basis $\{\hat{u}_x , \hat{u}_y\}$ by an angle θ to obtain a new set of vectors $\{\hat{u}_x{}' , \hat{u}_y{}'\}$ (Fig. 2.2). The transformation equations describing these rotations are

$$\begin{aligned} \hat{u}_x{}' &= \mathbf{R}(\theta)\hat{u}_x = \cos\theta\,\hat{u}_x + \sin\theta\,\hat{u}_y \\ \hat{u}_y{}' &= \mathbf{R}(\theta)\hat{u}_y = -\sin\theta\,\hat{u}_x + \cos\theta\,\hat{u}_y \end{aligned} \qquad (2.1)$$

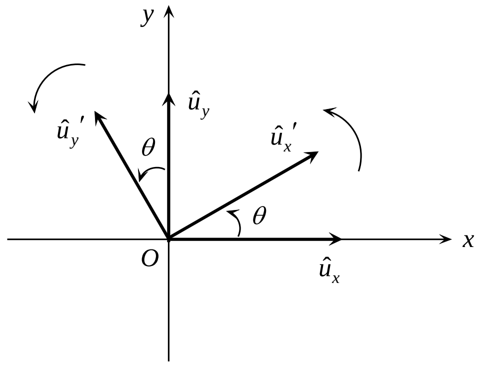

Figure 2.2

Now, let $\vec{A} = A_x\,\hat{u}_x + A_y\,\hat{u}_y$ be a vector on the *xy*-plane (see Fig. 2.1). The rotation operator $\mathbf{R}(\theta)$ will transform it into a new vector

$$\vec{A}' = \mathbf{R}(\theta)\vec{A} = A_x{}'\hat{u}_x + A_y{}'\hat{u}_y \qquad (2.2)$$

We want to express the components $A_x{}'$ and $A_y{}'$ in terms of A_x, A_y and θ. By the linearity of $\mathbf{R}(\theta)$ and by using (2.1), we have:

$$\begin{aligned} \vec{A}' &= \mathbf{R}(\theta)\left(A_x\hat{u}_x + A_y\hat{u}_y\right) = A_x\mathbf{R}(\theta)\hat{u}_x + A_y\mathbf{R}(\theta)\hat{u}_y \\ &= \left(A_x\cos\theta - A_y\sin\theta\right)\hat{u}_x + \left(A_x\sin\theta + A_y\cos\theta\right)\hat{u}_y \end{aligned}$$

By comparing this with (2.2), we get:

$$\begin{aligned} A_x{}' &= A_x\cos\theta - A_y\sin\theta \\ A_y{}' &= A_x\sin\theta + A_y\cos\theta \end{aligned} \qquad (2.3)$$

We define the matrix

$$M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \qquad (2.4)$$

The systems (2.1) and (2.3) are then rewritten in the form of matrix equations as

$$\begin{bmatrix} \hat{u}_x{}' \\ \hat{u}_y{}' \end{bmatrix} = M^T \begin{bmatrix} \hat{u}_x \\ \hat{u}_y \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} A_x{}' \\ A_y{}' \end{bmatrix} = M \begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{2.5}$$

respectively, where M^T is the transpose of M.

We note that the vectors $\vec{A}$ and $\vec{A}' = \mathbf{R}(\theta)\vec{A}$ are *different* geometrical objects, the latter one being a transformation of the former. On the other hand, the components of these vectors, connected by (2.3), are referred to the *same* basis $\{\hat{u}_x, \hat{u}_y\}$. This is the general idea of the *active view* of a linear transformation.

In a more abstract sense, we consider an n-dimensional vector space Ω with basis vectors $\{\hat{e}_1, \hat{e}_2, ..., \hat{e}_n\} \equiv \{\hat{e}_k\}$, and we let $\mathbf{R}$ be a linear operator on Ω. We assume that the basis vectors transform under $\mathbf{R}$ as follows:

$$\hat{e}_i{}' = \mathbf{R}\,\hat{e}_i = \hat{e}_j R^j{}_i \quad \text{(sum on } j) \tag{2.6}$$

where the familiar summation convention for repeated upper and lower indices has been used. Thus, for each value of i, the right-hand side of (2.6) is actually a sum over all values of j, i.e., from j=1 to j=n. Explicitly,

$$\begin{aligned} \hat{e}_1{}' &= \hat{e}_1 R^1{}_1 + \hat{e}_2 R^2{}_1 + \cdots + \hat{e}_n R^n{}_1 \\ \hat{e}_2{}' &= \hat{e}_1 R^1{}_2 + \hat{e}_2 R^2{}_2 + \cdots + \hat{e}_n R^n{}_2 \\ &\vdots \\ \hat{e}_n{}' &= \hat{e}_1 R^1{}_n + \hat{e}_2 R^2{}_n + \cdots + \hat{e}_n R^n{}_n \end{aligned} \tag{2.7}$$

Now, let

$$\vec{V} = V^1 \hat{e}_1 + V^2 \hat{e}_2 + \cdots + V^n \hat{e}_n \equiv V^i \hat{e}_i \tag{2.8}$$

be a vector in Ω, and let $\vec{V}' = \mathbf{R}\vec{V}$. We have:

$$\vec{V}' = \mathbf{R}(V^j \hat{e}_j) = V^j \mathbf{R}\,\hat{e}_j = V^j \hat{e}_i R^i{}_j \equiv V^{i\,\prime} \hat{e}_i \ .$$

Therefore the components of the original and the transformed vector are related by

$$V^{i\,\prime} = R^i{}_j V^j \tag{2.9}$$

or, explicitly,

$$\begin{aligned} V^{1\prime} &= R^1{}_1 V^1 + R^1{}_2 V^2 + \cdots + R^1{}_n V^n \\ V^{2\prime} &= R^2{}_1 V^1 + R^2{}_2 V^2 + \cdots + R^2{}_n V^n \\ &\vdots \\ V^{n\prime} &= R^n{}_1 V^1 + R^n{}_2 V^2 + \cdots + R^n{}_n V^n \end{aligned} \qquad (2.10)$$

Define the $n\times n$ matrix

$$M = \left[R^i{}_j \right] \quad \text{with} \quad M_{ij} = R^i{}_j \qquad (2.11)$$

The basis transformations (2.6) are then written as

$$\begin{bmatrix} \hat{e}_1{}' \\ \vdots \\ \hat{e}_n{}' \end{bmatrix} = M^T \begin{bmatrix} \hat{e}_1 \\ \vdots \\ \hat{e}_n \end{bmatrix} \qquad (2.12)$$

while the component transformations (2.9) become

$$\begin{bmatrix} V^{1\prime} \\ \vdots \\ V^{n\prime} \end{bmatrix} = M \begin{bmatrix} V^1 \\ \vdots \\ V^n \end{bmatrix} \qquad (2.13)$$

3. Passive view of transformations

Imagine that our previous x-y system of axes on the plane, with basis unit vectors $\{\hat{u}_x , \hat{u}_y\}$, is rotated counterclockwise by an angle θ to obtain a new system of axes x' and y' with corresponding basis $\{\hat{u}_x{}' , \hat{u}_y{}'\}$ (Fig. 3.1). As before, the two bases are related by the system of equations

$$\begin{aligned} \hat{u}_x{}' &= \cos\theta\, \hat{u}_x + \sin\theta\, \hat{u}_y \\ \hat{u}_y{}' &= -\sin\theta\, \hat{u}_x + \cos\theta\, \hat{u}_y \end{aligned} \qquad (3.1)$$

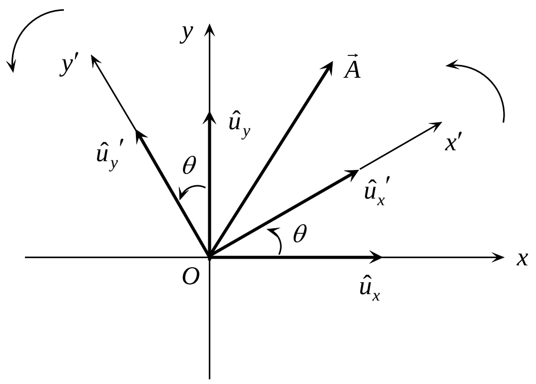

Figure 3.1

A vector $\vec{A}$ on the plane can be expressed in both these bases, as follows:

$$\vec{A} = A_x \hat{u}_x + A_y \hat{u}_y = A_x' \hat{u}_x' + A_y' \hat{u}_y' \tag{3.2}$$

Substituting the basis transformations (3.1) into the right-hand side of (3.2), and equating coefficients of similar unprimed basis vectors, we find:

$$\begin{aligned} A_x &= A_x' \cos\theta - A_y' \sin\theta \\ A_y &= A_x' \sin\theta + A_y' \cos\theta \end{aligned} \tag{3.3}$$

Solving this for the primed components, we get:

$$\begin{aligned} A_x' &= A_x \cos\theta + A_y \sin\theta \\ A_y' &= -A_x \sin\theta + A_y \cos\theta \end{aligned} \tag{3.4}$$

Notice that, in contrast to what we did in the previous section, here we keep the geometrical object $\vec{A}$ *fixed* and simply expand it in two *different* bases. This is the adopted practice in the *passive view* of a transformation.

Introducing the matrix

$$M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

we rewrite our previous equations in the matrix forms

$$\begin{bmatrix} \hat{u}_x' \\ \hat{u}_y' \end{bmatrix} = M^T \begin{bmatrix} \hat{u}_x \\ \hat{u}_y \end{bmatrix} \tag{3.5}$$

and

$$\begin{bmatrix} A_x \\ A_y \end{bmatrix} = M \begin{bmatrix} A_x' \\ A_y' \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} A_x' \\ A_y' \end{bmatrix} = M^{-1} \begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{3.6}$$

where

$$M^{-1} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = M^T \tag{3.7}$$

Notice that the transformation matrix M is *orthogonal*. As will be shown below, this is related to the fact that the transformation (rotation of axes) relates two Cartesian bases in a Euclidean space.

By comparing (2.3) and (3.4) it follows that the transformation equations of the passive view reduce to those of the active view upon replacing θ with $-\theta$. Physically this means that a passive transformation in which the vector $\vec{A}$ is fixed and the basis of our space is rotated *counterclockwise* is equivalent to an active transformation in which the basis is fixed and the vector $\vec{A}$ is rotated *clockwise*.

Let us generalize to the case of an *n*-dimensional vector space Ω with basis $\{\hat{e}_1, \hat{e}_2, ..., \hat{e}_n\} \equiv \{\hat{e}_k\}$. Let $\{\hat{e}_k{}'\}$ be another basis related to the former one by

$$\hat{e}_i{}' = \hat{e}_j \Lambda^j{}_{i'} \tag{3.8}$$

(note sum on j). A vector $\vec{V}$ in Ω may be expressed in both these bases, as follows:

$$\vec{V} = V^i \hat{e}_i = V^{j\prime} \hat{e}_j{}' = V^{j\prime} \hat{e}_i \Lambda^i{}_{j'}$$

where use has been made of (3.8). This yields

$$V^i = \Lambda^i{}_{j'} V^{j\prime} \tag{3.9}$$

Introducing the $n \times n$ matrix

$$M = \left[\Lambda^i{}_{j'} \right] \quad \text{with} \quad M_{ij} = \Lambda^i{}_{j'} \tag{3.10}$$

we write

$$\begin{bmatrix} \hat{e}_1{}' \\ \vdots \\ \hat{e}_n{}' \end{bmatrix} = M^T \begin{bmatrix} \hat{e}_1 \\ \vdots \\ \hat{e}_n \end{bmatrix} \tag{3.11}$$

and

$$\begin{bmatrix} V^1 \\ \vdots \\ V^n \end{bmatrix} = M \begin{bmatrix} V^{1\prime} \\ \vdots \\ V^{n\prime} \end{bmatrix} \quad \Rightarrow \quad \begin{bmatrix} V^{1\prime} \\ \vdots \\ V^{n\prime} \end{bmatrix} = M^{-1} \begin{bmatrix} V^1 \\ \vdots \\ V^n \end{bmatrix} \tag{3.12}$$

4. Orthogonal transformations in a Euclidean space

In this section the *passive* view of transformations will be adopted. Let Ω be an *n*-dimensional Euclidean space with Cartesian[1] coordinates $(x^1, x^2, ..., x^n) \equiv (x^k)$ and corresponding Cartesian basis $\{\hat{e}_k\}$. Let $(x^{k\prime})$ be another Cartesian coordinate system for

[1] Cartesian systems of coordinates exist only in Euclidean spaces. For example, you can define a system of Cartesian coordinates on a plane but you *cannot* define such coordinates on the surface of a sphere, which is a *non-Euclidean* space.

Ω, with corresponding basis $\{\hat{e}_k{}'\}$. We assume that the two coordinate systems have a common origin *O*≡(0,0,...,0). Both Cartesian bases are *orthonormal*, in the sense that

$$\hat{e}_i \cdot \hat{e}_j = \hat{e}_i{}' \cdot \hat{e}_j{}' = \delta_{ij} \qquad (4.1)$$

Assuming that the *handedness* of the two coordinate systems is the same (e.g., for $n=3$, both coordinate systems are right-handed) it is apparent that a linear transformation from one basis to the other is a "rotation" in *Ω*. Let us explore this in more detail.

Definition: A linear transformation from a Cartesian basis to another is said to be an *orthogonal transformation*.

Proposition 4.1: An orthogonal transformation is represented by an *orthogonal* matrix *M*:

$$M^{-1} = M^T \quad \Leftrightarrow \quad M^T M = M M^T = \mathbf{1} \qquad (4.2)$$

Proof: Assume a linear basis transformation of the form (3.8): $\hat{e}_i{}' = \hat{e}_j \Lambda^j{}_{i'}$. Also, let *M* be the transformation matrix defined in (3.10). We have:

$$\hat{e}_i{}' \cdot \hat{e}_j{}' = \left(\hat{e}_k \Lambda^k{}_{i'}\right) \cdot \left(\hat{e}_l \Lambda^l{}_{j'}\right) = \delta_{kl} \Lambda^k{}_{i'} \Lambda^l{}_{j'} = \sum_k \Lambda^k{}_{i'} \Lambda^k{}_{j'}$$
$$= \sum_k M_{ki} M_{kj} = \sum_k \left(M^T\right)_{ik} M_{kj} = \left(M^T M\right)_{ij}$$

where we have taken into account that the original (unprimed) basis is orthonormal. Given that the same is true for the transformed (primed) basis, we have:

$$\left(M^T M\right)_{ij} = \delta_{ij} \quad \Leftrightarrow \quad M^T M = \mathbf{1} .$$

The *magnitude* of a vector $\vec{V}$ is a non-negative quantity whose square is expressed in a Cartesian basis in terms of the scalar (dot) product, as follows:

$$\left|\vec{V}\right|^2 = \vec{V} \cdot \vec{V} = \left(V^i \hat{e}_i\right) \cdot \left(V^j \hat{e}_j\right) = V^i V^j \hat{e}_i \cdot \hat{e}_j = \delta_{ij} V^i V^j \qquad (4.3)$$

[Obviously, the last term in (4.3) is the sum of the squares of the components of $\vec{V}$.]

Proposition 4.2: An orthogonal transformation preserves the Cartesian form (4.3) of the magnitude of a vector.

Proof: By using the transformation formula (3.9) for components of vectors, derived in the previous section, we have:

$$\delta_{ij} V^i V^j = \delta_{ij} \left(\Lambda^i{}_{k'} V^{k'} \right) \left(\Lambda^j{}_{l'} V^{l'} \right) = \left(\sum_i \Lambda^i{}_{k'} \Lambda^i{}_{l'} \right) V^{k'} V^{l'}$$

$$= \left(\sum_i M_{ik} M_{il} \right) V^{k'} V^{l'} = \left(\sum_i \left(M^T \right)_{ki} M_{il} \right) V^{k'} V^{l'}$$

$$= \left(M^T M \right)_{kl} V^{k'} V^{l'} = \delta_{kl} V^{k'} V^{l'}$$

For a more compact proof, define the matrices

$$\left[V^k \right] \equiv \begin{bmatrix} V^1 \\ \vdots \\ V^n \end{bmatrix} \quad \text{and} \quad \left[V^k \right]^T \equiv \begin{bmatrix} V^1 & \cdots & V^n \end{bmatrix}$$

and similarly for the corresponding primed quantities. Then, in the unprimed basis,

$$\left| \vec{V} \right|^2 = \left[V^k \right]^T \left[V^k \right].$$

Using the fact that, by (3.12), $\left[V^k \right] = M \left[V^{k'} \right]$, we have:

$$\left[V^k \right]^T \left[V^k \right] = \left(M \left[V^{k'} \right] \right)^T M \left[V^{k'} \right] = \left[V^{k'} \right]^T M^T M \left[V^{k'} \right]$$

$$= \left[V^{k'} \right]^T \left[V^{k'} \right]$$

Comment: The above proof suggests an alternate definition of an orthogonal transformation as a linear transformation in a Euclidean space that preserves the Cartesian form of the magnitude of vectors. In fact, this is the way orthogonal transformations are usually defined in textbooks.

Now, let P be a point in Ω, with Cartesian coordinates $(x^1, x^2, \ldots, x^n) \equiv (x^k)$. In this system of coordinates the position vector of P can be written as $\vec{r} = x^i \hat{e}_i$. Since this vector is a geometrical object independent of the system of coordinates, we can write:

$$\vec{r} = x^i \hat{e}_i = x^{j\,\prime} \hat{e}_j{}' .$$

By using (3.8) we find, as in Sec. 3,

$$x^i = \Lambda^i{}_{j'} x^{j\,\prime} \tag{4.4}$$

which is the analog of (3.9). If M is the matrix defined in (3.10), and if $[x^k]$ is the column vector of the x^k, then by the general matrix relation (3.12) we have:

$$\left[x^k\right] = M\left[x^{k'}\right] \quad \Rightarrow \quad \left[x^{k'}\right] = M^{-1}\left[x^k\right] = M^T\left[x^k\right] \tag{4.5}$$

where the orthogonality condition (4.2) has been used. Let us call

$$M^T \equiv L \quad \text{with} \quad L_{ij} = M_{ji} = \Lambda^j{}_{i'} \tag{4.6}$$

Then the matrix relation (4.5) can be written as a system of n linear equations of the form

$$\begin{aligned} x^{1\prime} &= L_{11}\, x^1 + L_{12}\, x^2 + \cdots + L_{1n}\, x^n \\ x^{2\prime} &= L_{21}\, x^1 + L_{22}\, x^2 + \cdots + L_{2n}\, x^n \\ &\vdots \\ x^{n\prime} &= L_{n1}\, x^1 + L_{n2}\, x^2 + \cdots + L_{nn}\, x^n \end{aligned} \tag{4.7}$$

which equations represent an orthogonal coordinate transformation in Ω.

As an example for n=2, let Ω be a plane with Cartesian coordinates $(x^1, x^2) \equiv (x, y)$. A position vector in Ω is written: $\vec{r} = x\hat{u}_x + y\hat{u}_y$. As seen in Sec. 3, the transformation matrix M for a rotation of the basis vectors by an angle θ is

$$M = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \Rightarrow \quad L = M^T = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}.$$

The coordinate transformation equations (4.7) are written here as

$$\begin{aligned} x' &= x\cos\theta + y\sin\theta \\ y' &= -x\sin\theta + y\cos\theta \end{aligned}$$

Exercise: By using the relations $\vec{V} = V^j \hat{e}_j$ and $\hat{e}_j{}' = \hat{e}_l \Lambda^l{}_{j'}$, together with (3.10) and (4.1), show the following:

$$V^i = \hat{e}_i \cdot \vec{V} \,,$$

$$M_{ij} = \hat{e}_i \cdot \hat{e}_j{}' \,.$$

Under an orthogonal transformation from one Cartesian system of coordinates to another, the components V^k of a vector transform like the coordinates x^k themselves. That is,

$$V^{i'} = L_{ij} V^j \,.$$

From (4.7) we have that

$$L_{ij} = \frac{\partial x^{i'}}{\partial x^j} .$$

Therefore,

$$V^{i'} = \frac{\partial x^{i'}}{\partial x^j} V^j \quad \text{and, conversely,} \quad V^i = \frac{\partial x^i}{\partial x^{j'}} V^{j'} \tag{4.8}$$

5. Active and passive view combined

Let *Ω* be an *n*-dimensional vector space with basis $\{\hat{e}_k\}$ $(k = 1, 2, \ldots, n)$. Let **A** be a linear operator on *Ω*. The action of **A** on the basis vectors is given by

$$\mathbf{A}\,\hat{e}_j = \sum_i \hat{e}_i A_{ij} \equiv \hat{e}_i A_{ij} \tag{5.1}$$

(Note a slight change in the summation convention; in this section subscripts only will be used.) The $n \times n$ matrix $A=[A_{ij}]$ is the *matrix representation of the operator* **A** in the basis $\{\hat{e}_k\}$.

A vector in *Ω* is written:

$$\vec{x} = \sum_i x_i\, \hat{e}_i \equiv x_i\, \hat{e}_i \tag{5.2}$$

Let $\vec{y} = \mathbf{A}\,\vec{x}$. If $\vec{y} = y_i\, \hat{e}_i$, then, by the linearity of **A** and by using (5.1) and (5.2) we find that

$$y_i = A_{ij}\, x_j \quad (\text{sum on } j) \tag{5.3}$$

which represents a system of *n* linear equations for $i=1,...,n$. In matrix form,

$$[y_k] = A\,[x_k] \tag{5.4}$$

where $[x_k]$ and $[y_k]$ are column vectors.

Now, let **A** and **B** be linear operators on *Ω*. We define their product **C**=**AB** by

$$\mathbf{C}\,\vec{x} = (\mathbf{AB})\,\vec{x} \equiv \mathbf{A}\,(\mathbf{B}\vec{x}) , \quad \forall\, \vec{x} \in \Omega \tag{5.5}$$

Then, in the basis $\{\hat{e}_k\}$,

$$\mathbf{C}\,\hat{e}_j = \mathbf{A}\,(\mathbf{B}\,\hat{e}_j) = \mathbf{A}\,(\hat{e}_l\, B_{lj}) = B_{lj}\,(\mathbf{A}\,\hat{e}_l) = A_{il}\, B_{lj}\, \hat{e}_i \equiv \hat{e}_i\, C_{ij}$$

where

$$C_{ij} = A_{il} B_{lj} \quad \text{or, in matrix form,} \quad C = AB \tag{5.6}$$

That is, in any basis of *Ω*,

> *the matrix of the product of two operators is the product of the matrices of these operators.*

Consider now a change of basis (passive transformation) with transformation matrix $T=[T_{ij}]$:

$$\hat{e}_j{}' = \hat{e}_i T_{ij} \tag{5.7}$$

The inverse transformation is

$$\hat{e}_j = \hat{e}_i{}' \left(T^{-1}\right)_{ij} \tag{5.8}$$

The same vector may be expressed in both these bases as $\vec{x} = x_i \hat{e}_i = x_j{}' \hat{e}_j{}'$, from which we get, by using (5.7) and (5.8),

$$x_i = T_{ij}\, x_j{}' \quad \text{and} \quad x_i{}' = \left(T^{-1}\right)_{ij} x_j \tag{5.9}$$

In matrix form,

$$[x_k] = T\,[x_k{}'] \quad \text{and} \quad [x_k{}'] = T^{-1}[x_k] \tag{5.10}$$

How do the matrix elements of a linear operator **A** transform under a change of basis of the form (5.7)? In other words, how does the matrix of an active transformation transform under a passive transformation? Let $\vec{y} = \mathbf{A}\,\vec{x}$. By combining (5.10) with (5.4), we have:

$$[y_k{}'] = T^{-1}[y_k] = T^{-1}A[x_k] = T^{-1}AT\,[x_k{}'] \equiv A'\,[x_k{}'] \;\Rightarrow$$

$$A' = T^{-1}A\,T \tag{5.11}$$

For an alternative proof, note that

$$\mathbf{A}\,\hat{e}_j{}' = \mathbf{A}\,(\hat{e}_i T_{ij}) = T_{ij}\,\mathbf{A}\,\hat{e}_i = T_{ij}\,\hat{e}_l A_{li} = A_{li} T_{ij}\,\hat{e}_k{}' \left(T^{-1}\right)_{kl}$$

$$= \left(T^{-1}AT\right)_{kj} \hat{e}_k{}' \equiv \hat{e}_k{}' A'_{kj} \;\Rightarrow\; A' = T^{-1}AT$$

as before. A transformation of the form (5.11) is called a *similarity transformation*.

By applying the properties of the trace and the determinant of a matrix to (5.11) it is not hard to show that, under basis transformations, *the trace and the determinant of the matrix representation of an operator remain unchanged:* $\mathrm{tr}A=\mathrm{tr}A'$, $\det A=\det A'$. This means that the trace and the determinant are basis-independent quantities that are properties of the operator itself, rather than properties of its representation.

Definition: A vector $\vec{x} \neq 0$ is said to be an *eigenvector* of the linear operator **A** if a constant λ exists such that

$$\mathbf{A}\vec{x} = \lambda\vec{x} \tag{5.12}$$

The constant λ is an *eigenvalue* of **A**, to which eigenvalue this eigenvector belongs. Note that, in general, more than one eigenvector may belong to the same eigenvalue.

In a given basis $\{\hat{e}_k\}$, the linear system (5.3) corresponding to the eigenvalue equation (5.12) takes on the form

$$A_{ij}\, x_j = \lambda\, x_i \quad \text{or} \quad (A_{ij} - \lambda\,\delta_{ij})\, x_j = 0 \tag{5.13}$$

where $[A_{ij}]=A$ is the matrix of the operator **A** in the given basis. This is a homogeneous linear system of equations, which has a nontrivial solution for the eigenvector components iff

$$\det[A_{ij} - \lambda\,\delta_{ij}] = 0 \quad \text{or} \quad \det(A - \lambda 1) = 0 \tag{5.14}$$

where 1 here is the n-dimensional unit matrix. This polynomial equation determines the eigenvalues λ_i (not necessarily all different from each other) of the operator **A**.

Now, in general, for any value of the constant λ the matrix $(A–\lambda 1)$ is the representation of the operator $(\mathbf{A}–\lambda\mathbf{1})$ in the considered basis $\{\hat{e}_k\}$. Under a basis transformation to $\{\hat{e}_k{}'\}$ this matrix transforms according to (5.11):

$$(A–\lambda 1)' = T^{-1}(A–\lambda 1)\, T = T^{-1} A\, T - \lambda 1 \equiv A' - \lambda 1 \ .$$

On the other hand, by the invariance of the determinant under this transformation,

$$\det(A' - \lambda 1) = \det(A - \lambda 1) \ .$$

In particular, if λ is an eigenvalue of the operator **A**, the right-hand side of the above equation vanishes in view of (5.14) and, therefore, the same must be true for the left-hand side *for the same value of* λ. That is, the polynomial equation (5.14) determines the eigenvalues of **A** uniquely, regardless of the chosen representation. We conclude that

> *the eigenvalues of an operator are a property of the operator itself and do not depend on the choice of basis of the space* Ω.

If we can find n linearly independent eigenvectors $\{\vec{x}_k\}$ of **A**, belonging to the corresponding eigenvalues λ_k (not necessarily all different) we can use these vectors to define a basis of Ω. The matrix representation of **A** in this basis is given by (5.1): $\mathbf{A}\vec{x}_j = \vec{x}_i A_{ij}$. On the other hand, if $\lambda_j \equiv \lambda'$, then $\mathbf{A}\vec{x}_j = \lambda'\vec{x}_j = \lambda'\delta_{ij}\vec{x}_i$. Therefore, since the $\vec{x}_k$ are linearly independent, we must have $A_{ij}=\lambda'\delta_{ij}$. We conclude that, in the eigenvector basis the matrix representation of the operator **A** has the *diagonal* form

$$A = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) \ .$$

Moreover, by the above formula and by the fact that the quantities tr$A$, det$A$ and λ_k are basis-independent (i.e., invariant under basis transformations) it follows that, in *any* basis of Ω,

$$\mathrm{tr}\,A = \lambda_1 + \lambda_2 + \ldots + \lambda_n \quad , \qquad \det A = \lambda_1\,\lambda_2 \ldots \lambda_n \tag{5.15}$$

Proposition 5.1: Let **A** and **B** be two linear operator on Ω. We assume that **A** and **B** have a common set of n linearly independent eigenvectors $\{\vec{x}_k\}$. Then the operators **A** and **B** *commute*:

$$\mathbf{AB} = \mathbf{BA} \quad \Leftrightarrow \quad [\mathbf{A}, \mathbf{B}] \equiv \mathbf{AB} - \mathbf{BA} = 0$$

where $[\mathbf{A}, \mathbf{B}]$ denotes the *commutator* of **A** and **B**.

Proof: Since the n vectors $\{\vec{x}_k\}$ are linearly independent, they define a basis of Ω. By assumption, for each value of k the vector $\vec{x}_k$ is an eigenvector of both **A** and **B**, with corresponding eigenvalues, say, α and β. Then,

$$(\mathbf{AB})\,\vec{x}_k \equiv \mathbf{A}\,(\mathbf{B}\,\vec{x}_k) = \mathbf{A}\,(\beta\vec{x}_k) = \beta\,(\mathbf{A}\,\vec{x}_k) = \beta\alpha\,\vec{x}_k$$

and similarly, $(\mathbf{BA})\,\vec{x}_k = \alpha\beta\vec{x}_k$. Thus,

$$(\mathbf{AB})\,\vec{x}_k = (\mathbf{BA})\,\vec{x}_k \quad \Leftrightarrow \quad [\mathbf{A},\mathbf{B}]\,\vec{x}_k = 0\,,$$

for all k=1,...,n. Now, let $\vec{\Psi} = \xi_i\,\vec{x}_i$ be an arbitrary vector in Ω. Then,

$$[\mathbf{A},\mathbf{B}]\,\vec{\Psi} = [\mathbf{A},\mathbf{B}](\xi_i\,\vec{x}_i) = \xi_i\,[\mathbf{A},\mathbf{B}]\vec{x}_i = 0, \quad \forall\,\vec{\Psi} \in \Omega\,.$$

This means that $[\mathbf{A}, \mathbf{B}]$=0.

Definition: An operator **A** is said to be *nonsingular* if det$A$≠0 (note that this is a *basis-independent* property). A nonsingular operator is *invertible*, in the sense that an inverse linear operator $\mathbf{A}^{-1}$ on Ω exists such that $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{1}_{\mathrm{op}}$, where $\mathbf{1}_{\mathrm{op}}$ is the unit operator. This allows us to write

$$\vec{y} = \mathbf{A}\,\vec{x} \quad \Leftrightarrow \quad \vec{x} = \mathbf{A}^{-1}\,\vec{y}\,.$$

By (5.4) it follows that, if A is the matrix representation of the nonsingular operator **A** in some basis, then *the matrix of the inverse operator* $\mathbf{A}^{-1}$ *is the inverse* A^{-1} *of* A. As is well known, the matrix A may have an inverse iff det$A$≠0, whence the definition of a nonsingular operator. In view of the second relation in (5.15),

all eigenvalues of a nonsingular operator are nonzero.

Indeed, if even one eigenvalue vanishes, then det$A$=0 in *any* representation.

6. Comments

Both the active and the passive view are of importance in Physics. Let us see some examples:

1. The *Galilean transformation* of Classical Mechanics and the *Lorentz transformation* of Relativity[2] are *passive* transformations connecting different inertial frames of reference. When expressed in terms of mathematical equations, all physical laws are required to be invariant in form upon passing from one inertial frame to another.

2. The operators of Quantum Mechanics[3] are *active* transformations from a quantum state to a new state. On the other hand, both states and operators may be represented by matrices in different bases, the transformation from one basis to another being a *passive* transformation. Typically, the basis vectors of the quantum-mechanical space are chosen to be eigenvectors of linear operators representing physical quantities such as energy, angular momentum, etc. In such a basis the related operator is represented by a *diagonal* matrix, the diagonal elements being the *eigenvalues* of the operator. Physically, these eigenvalues give the possible values that a measurement of the associated physical quantity may yield in an experiment.

[2] H. Goldstein, *Classical Mechanics*, 2nd Ed. (Addison-Wesley, 1980).
[3] E. Merzbacher, *Quantum Mechanics*, 3rd Ed. (Wiley, 1998).

Transformation Lie groups and operator representations

Costas J. Papachristou[1]

Hellenic Naval Academy

A note on continuous groups of transformations on linear spaces and manifolds and the operator representations of transformation Lie groups and algebras.

1. Lie groups and Lie algebras: An overview

We review some basic definitions concerning Lie groups and Lie algebras.

A *group* is a set $G=\{a,b,c,...\}$ equipped with an internal "multiplication" operation with the following properties:

1. Closure: $ab \in G,\ \forall\, a, b \in G$.
2. Associativity: $a(bc) = (ab)c$.
3. Identity element: $\exists\, e \in G$: $ae=ea,\ \forall\, a \in G$.
4. Inverse element: $\forall\, a \in G,\ \exists\, a^{-1} \in G$: $aa^{-1} = a^{-1}a = e$.

A group is *abelian* (or commutative) if $ab=ba,\ \forall\, a, b \in G$.

A *subgroup* of G is a subset $H \subseteq G$ that is itself a group under the group operation of G. Obviously, H must contain the identity element e of G as well as the inverse of any element of H.

A map $\varphi : G \rightarrow G'$ from a group G to a group G' is called a *homomorphism* if it preserves group multiplication. That is, for any a, $b \in G$, the images $\varphi(a) \in G'$ and $\varphi(b) \in G'$ satisfy the relation

$$\varphi(a)\,\varphi(b) = \varphi(ab) .$$

If the homomorphism φ is 1-1, it is called an *isomorphism*. An isomorphic relation of G with a group of matrices or operators is called a matrix or operator *representation* of G, accordingly.

A real *Lie algebra* $\mathcal{L}$ of dimension n is an n-dimensional real vector space equipped with an internal *Lie bracket* operation [,] that satisfies the following properties:

1. Closure: $[a,b] \in \mathcal{L},\ \forall\, a, b \in \mathcal{L}$.
2. Linearity: $[\kappa a+\lambda b,\, c] = \kappa[a,c]+\lambda[b,c]$ $(\kappa, \lambda \in R)$.
3. Antisymmetry: $[a,b] = -[b,a]$. Corollary: $[a,a]=0$.
4. Jacobi identity: $\big[a,[b,c]\big] + \big[b,[c,a]\big] + \big[c,[a,b]\big] = 0$.

A Lie algebra is *abelian* (or commutative) if $[a,b]=0$, $\forall\, a, b \in \mathcal{L}$.

---

[1] cjpapachristou@gmail.com

A *subalgebra* S of $\mathcal{L}$ is a subspace of $\mathcal{L}$ that itself is a Lie algebra. The algebra S is an *invariant subalgebra* or *ideal* of $\mathcal{L}$ if $[a,b]\in S,\ \forall\, a\in S,\ b\in\mathcal{L}$. A Lie algebra $\mathcal{L}$ is said to be *simple* if it contains no ideals other than itself; $\mathcal{L}$ is *semisimple* if it contains no *abelian* ideals.

Examples of Lie algebras:

1. The algebra of ($m\times m$) matrices, with $[A,B]=AB-BA$ (*commutator*). Diagonal matrices constitute an abelian subalgebra of this algebra.

2. The algebra of all vectors in 3-dimensional space, with $[\vec{V},\vec{W}]=\vec{V}\times\vec{W}$ (vector product). Vectors parallel to a given axis form an abelian subalgebra of this algebra.

A map $\psi: \mathcal{L}\to\mathcal{L}'$ from a Lie algebra $\mathcal{L}$ to a Lie algebra $\mathcal{L}'$ is a *homomorphism* if it satisfies the following properties:

$$\psi(\kappa a+\lambda b) = \kappa\psi(a)+\lambda\psi(b)\quad (\kappa,\lambda\in R)\ ;$$
$$\psi([a,b]) = [\psi(a),\psi(b)]\ .$$

If the map ψ is 1-1, it is called an *isomorphism*. Isomorphic Lie algebras $\mathcal{L}$ and $\mathcal{L}'$ have equal dimensions [1]: $dim\mathcal{L}=dim\mathcal{L}'$.

Let $\{\tau_i\ /\ i=1,2,\dots,n\}$ be a basis of an n-dimensional Lie algebra $\mathcal{L}$. Since the Lie bracket of any two basis elements τ_i and τ_j is an element of $\mathcal{L}$, it must be a linear combination of the $\{\tau_k\}$. That is,

$$[\tau_i,\tau_j]=C_{ij}^k\,\tau_k \tag{1}$$

(sum on k from 1 to n). By the antisymmetry of the Lie bracket, $C_{ij}^k=-C_{ji}^k$. The real constants C_{ij}^k are called *structure constants* of the Lie algebra $\mathcal{L}$.

Proposition 1: Let $\psi: \mathcal{L}\to\mathcal{L}'$ be a Lie algebra isomorphism. If $\{\tau_k\}$ ($k=1,2,\dots,n$) is a basis of $\mathcal{L}$, then $\{\psi(\tau_k)\}$ is a basis of $\mathcal{L}'$.

Proof: Being a basis of $\mathcal{L}$, the $\{\tau_k\}$ are linearly independent; hence no linear combination of them can be zero (unless, of course, all coefficients are trivially zero). Now, by the properties of ψ, a linear combination of the $\{\tau_k\}$ is mapped onto a linear combination of the $\{\psi(\tau_k)\}$ with the same coefficients. This means that the latter combination cannot vanish, since it can only be zero if the former one is zero as well; that is, if all coefficients in the combination are zero. We conclude that the $\{\psi(\tau_k)\}$ are linearly independent and may serve as a basis for $\mathcal{L}'$.

Proposition 2: Isomorphic Lie algebras share common structure constants.

Proof: Let $\psi: \mathcal{L}\to\mathcal{L}'$ be a Lie algebra isomorphism and let τ_i, τ_j be any two basis elements of $\mathcal{L}$. Then, $\psi(\tau_i)$ and $\psi(\tau_j)$ are basis elements of $\mathcal{L}'$. By (1) and by the properties of ψ,

$$\psi([\tau_i, \tau_j]) = \psi(C_{ij}^k \tau_k) \Rightarrow [\psi(\tau_i), \psi(\tau_j)] = C_{ij}^k \psi(\tau_k)\,; \text{ q.e.d.}$$

Roughly speaking, a *Lie group* is a group G whose elements depend on a number of parameters that can be varied in a continuous way. The *dimension n* of G is the number of real parameters parametrizing the elements of G. We assume that $dimG=n$ and we let $\{\lambda^1, \lambda^2, \dots, \lambda^n\}$ be the set of n parameters of G. We arrange the parameterization of G so that the identity element of G corresponds to $\lambda^k=0$ for all $k=1,2,...,n$.

An important class of Lie groups consists of groups of $(m\times m)$ matrices parametrized by n parameters λ^k ($k=1,2,...,n$). Since an $(m\times m)$ matrix produces a *linear transformation* on an m-dimensional Euclidean space, matrix groups are called *linear groups*.

Lie groups are closely related to Lie algebras. Let G be an n-dimensional Lie group of $(m\times m)$ matrices $A(\lambda^1, \lambda^2, \dots, \lambda^n) \equiv A(\lambda)$ (where by λ we collectively denote the set of the n parameters λ^k). We define the n $(m\times m)$ matrices τ_k by

$$\tau_k = \frac{\partial A(\lambda)}{\partial \lambda^k}\Big|_{\lambda^1=\lambda^2=\cdots=\lambda^n=0} \qquad (2)$$

or, in terms of matrix elements,

$$(\tau_k)_{pq} = \frac{\partial A_{pq}}{\partial \lambda^k}\Big|_{\lambda^1=\lambda^2=\cdots=\lambda^n=0}$$

($k=1,2,...,n$; $p, q=1,2,...,m$). The n matrices τ_k are called *infinitesimal operators* (or generators) of the Lie group G and form the basis of an n-dimensional real Lie algebra $\mathcal{L}$ [1]. Thus $[\tau_i, \tau_j] = C_{ij}^k \tau_k$, where the C_{ij}^k are real constants. A general element a of $\mathcal{L}$ is written as a linear combination of the τ_k: $a=\xi^k\tau_k$ (sum on k), for real coefficients ξ^k. [Note carefully that the matrix elements $(\tau_k)_{pq}$ *themselves* are *not* required to be real numbers!]

Now, let $a=\lambda^k\tau_k$ be the general element of $\mathcal{L}$. The general element $A(\lambda)$ of the Lie group G parametrized by the λ^k can then be written as [1,2]

$$A(\lambda) = e^a = \exp(\lambda^k \tau_k) \qquad (3)$$

where e^a is the matrix exponential function

$$e^a \equiv \exp a = \sum_{l=0}^{\infty} \frac{a^l}{l!} = 1 + a + \frac{a^2}{2} + \cdots$$

For infinitesimal values of the parameters λ^k we may use the approximate expression

$$e^a \simeq 1 + a$$

so that

$$A(\lambda) \simeq 1 + \lambda^k \tau_k \,.$$

The simplest example of a Lie group is a one-parameter continuous group, such as the group $SO(2)$ of rotations on a plane. A rotation of a vector by an angle λ is represented by the (2×2) orthogonal matrix

$$A(\lambda)=\begin{bmatrix}\cos\lambda & -\sin\lambda\\ \sin\lambda & \cos\lambda\end{bmatrix}\qquad(\lambda\in R)\,.$$

(Notice that $A^tA=1$ and $\det A=1$.) Then

$$\frac{dA}{d\lambda}=\begin{bmatrix}-\sin\lambda & -\cos\lambda\\ \cos\lambda & -\sin\lambda\end{bmatrix}$$

and, by Eq. (2), the single basis element τ of the associated Lie algebra is

$$\tau=\frac{dA}{d\lambda}\Big|_{\lambda=0}=\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}\,.$$

According to (3), $A(\lambda)=e^{\lambda\tau}$ and, for infinitesimal λ, $A(\lambda)\simeq 1+\lambda\tau$. Indeed, by setting $\sin\lambda=\lambda$ and $\cos\lambda=1$, we have:

$$A(\lambda)\simeq\begin{bmatrix}1 & -\lambda\\ \lambda & 1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}+\lambda\begin{bmatrix}0 & -1\\ 1 & 0\end{bmatrix}=1+\lambda\tau\,.$$

Another single-parameter Lie group is the unitary group $U(1)$ with elements $\{e^{i\lambda}\}$ $(\lambda\in R)$, which may be regarded as (1×1) matrices. Consider the map φ: $U(1)\to SO(2)$ defined by

$$\varphi\left(e^{i\lambda}\right)=\begin{bmatrix}\cos\lambda & -\sin\lambda\\ \sin\lambda & \cos\lambda\end{bmatrix}\,.$$

This map is a homomorphism, since

$$\begin{aligned}\varphi\left(e^{i\lambda}\cdot e^{i\lambda'}\right)=\varphi\left(e^{i(\lambda+\lambda')}\right)&=\begin{bmatrix}\cos(\lambda+\lambda') & -\sin(\lambda+\lambda')\\ \sin(\lambda+\lambda') & \cos(\lambda+\lambda')\end{bmatrix}\\ &=\begin{bmatrix}\cos\lambda & -\sin\lambda\\ \sin\lambda & \cos\lambda\end{bmatrix}\begin{bmatrix}\cos\lambda' & -\sin\lambda'\\ \sin\lambda' & \cos\lambda'\end{bmatrix}\\ &=\varphi\left(e^{i\lambda}\right)\cdot\varphi\left(e^{i\lambda'}\right).\end{aligned}$$

Moreover, it can be shown [1] that the map φ is 1-1. Therefore, φ is a Lie-group isomorphism.

We finally remark that isomorphic Lie groups have isomorphic Lie algebras [1]. More generally, under certain restrictions, homomorphic Lie groups may have isomorphic Lie algebras, as the case is with the groups $SU(2)$ and $SO(3)$.

2. Group operators for a linear Lie group

Let $(x^1, \dots, x^m)$ be a system of coordinates on R^m. Consider an n-dimensional Lie group G represented by $(m\times m)$ matrices g the elements of which depend on n real parameters $(a^1, \dots, a^n)$. We call $\vec{x}$ the column vector with components $(x^1, \dots, x^m)$. The action of G on this vector is expressed as

$$\vec{x} \to g\vec{x}\ , \quad g \in G \tag{4}$$

In effect, Eq. (4) describes a linear coordinate transformation[2] on R^m.

Let $\{L_1, \dots, L_n\}$ be a basis of the Lie algebra of G, where the $\{L_\gamma\}$ are $(m\times m)$ matrices.[3] Then there exist real structure constants $C_{\alpha\beta}^{\gamma}$ such that the following commutation relations are satisfied:

$$[L_\alpha\, , L_\beta] = C_{\alpha\beta}^{\gamma} L_\gamma \quad (\text{sum on } \gamma) \tag{5}$$

An element $g\in G$ can then be put in the form $g=\exp(a^\lambda L_\lambda)$ [1,2] so that (4) is written: $\vec{x} \to \exp(a^\lambda L_\lambda)\,\vec{x}$. For infinitesimal values δa^λ of the group parameters,

$$\exp(\delta a^\lambda L_\lambda) \simeq 1 + \delta a^\lambda L_\lambda$$

so that $\vec{x} \to (1+\delta a^\lambda L_\lambda)\,\vec{x} \equiv \vec{x} + \delta\vec{x}$, where

$$\delta\vec{x} = \delta a^\lambda L_\lambda \vec{x} \;\;\Leftrightarrow\;\; \delta x^i = \delta a^\lambda (L_\lambda)^i{}_k\, x^k \tag{6}$$

The expression $g=\exp(a^\lambda L_\lambda)$ is a representation of G in terms of linear coordinate transformations (4) on R^m. We now seek a different realization of G in terms of transformations of functions $F(\vec{x})$, $\vec{x} \in R^m$. We define the operators

$$T(g):\; F \to T(g)F, \quad g\in G$$

by

$$[T(g)\,F](\vec{x}) = F\left(g^{-1}\vec{x}\right) \tag{7}$$

Proposition 1: The operators $T(g)$ constitute an operator representation of G.

Proof: Let $g_1 , g_2 \in G$. Then, for an arbitrary function F on R^m,

$$\begin{aligned}[T(g_1 g_2)F](\vec{x}) &= F\left(g_2^{-1} g_1^{-1}\vec{x}\right) = [T(g_2)F]\left(g_1^{-1}\vec{x}\right) = \left\{T(g_1)[T(g_2)F]\right\}(\vec{x}) \\ &\equiv \left\{[T(g_1)T(g_2)]F\right\}(\vec{x}) \;\;\Rightarrow\end{aligned}$$

$$T(g_1 g_2) = T(g_1)\, T(g_2)\ , \quad \text{q.e.d.}$$

[2] For definiteness we regard this as an *active* transformation from a point $x\in R^m$ with coordinates x^k to a point x' with coordinates $x^{k\prime} = (gx)^k$.
[3] Greek indices run from 1 to n while Latin indices run from 1 to m. The summation convention will be used throughout.

For $g = \exp(a^\lambda L_\lambda) \simeq 1 + \delta a^\lambda L_\lambda$ we have that $g^{-1} \simeq 1 - \delta a^\lambda L_\lambda$, and so (7) yields, by using Eq. (A.1) in the Appendix:

$$[T(g)F](\vec{x}) \simeq F\left(\vec{x} - \delta a^\lambda L_\lambda \vec{x}\right) \simeq F(\vec{x}) - \delta a^\lambda L_\lambda \vec{x} \cdot \vec{\nabla} F(\vec{x})$$
$$= \left(1 - \delta a^\lambda L_\lambda \vec{x} \cdot \vec{\nabla}\right) F(\vec{x}) \;\Rightarrow$$

$$T(g) \simeq 1 - \delta a^\lambda L_\lambda \vec{x} \cdot \vec{\nabla} \equiv 1 + \delta a^\lambda P_\lambda \tag{8}$$

where $P_\lambda = -L_\lambda \vec{x} \cdot \vec{\nabla} = -(L_\lambda \vec{x})^i \dfrac{\partial}{\partial x^i} \;\Rightarrow$

$$P_\lambda = -(L_\lambda)^i{}_k\, x^k \frac{\partial}{\partial x^i} \equiv -(L_\lambda)^i{}_k\, x^k \partial_i \tag{9}$$

where we have introduced the notation $\partial_i \equiv \partial/\partial x^i$. For finite values of the group parameters a^λ, Eq. (8) generalizes to $T(g) = \exp(a^\lambda P_\lambda)$ [3,4].

Proposition 2: The operators $\{P_\lambda\}$ are the basis of a Lie algebra isomorphic to the Lie algebra of the matrices $\{L_\gamma\}$. Thus, if the commutation relations (5) are valid, then also

$$[P_\alpha\,,\, P_\beta] = C^\gamma_{\alpha\beta} P_\gamma \tag{10}$$

Proof: Consider the linear mapping

$$\Psi:\; L \to P = \Psi(L) = -L^i{}_k\, x^k \partial_i \tag{11}$$

where the matrix L is an element of the Lie algebra of G. Let L_1 , L_2 be two such elements. Then,

$$P_1 = \Psi(L_1) = -(L_1)^i{}_k\, x^k \partial_i \;, \quad P_2 = \Psi(L_2) = -(L_2)^i{}_k\, x^k \partial_i \;.$$

We have:

$$\Psi([L_1\,,L_2]) = \Psi(L_1L_2 - L_2L_1) = \Psi(L_1L_2) - \Psi(L_2L_1) \quad \text{(since } \Psi \text{ is linear)}$$
$$= -(L_1L_2)^i{}_k\, x^k \partial_i + (L_2L_1)^i{}_k\, x^k \partial_i$$
$$= -(L_1)^i{}_j (L_2)^j{}_k\, x^k \partial_i + (L_2)^i{}_j (L_1)^j{}_k\, x^k \partial_i \,.$$

On the other hand,

$$[\Psi(L_1)\,,\Psi(L_2)] = [P_1\,,P_2] = P_1P_2 - P_2P_1$$
$$= (L_1)^i{}_j\, x^j \partial_i [(L_2)^k{}_l\, x^l \partial_k] - (L_2)^k{}_l\, x^l \partial_k [(L_1)^i{}_j\, x^j \partial_i] \,.$$

After a lengthy but straightforward calculation, and by canceling out second-order derivatives, we find:

$$[\Psi(L_1)\,,\Psi(L_2)] = -(L_1)^i{}_j (L_2)^j{}_k\, x^k \partial_i + (L_2)^i{}_j (L_1)^j{}_k\, x^k \partial_i \,.$$

We thus conclude that

$$\Psi([L_1, L_2]) = [\Psi(L_1), \Psi(L_2)]$$

which is what we needed to prove. Moreover,

$$[P_\alpha , P_\beta] = [\Psi(L_\alpha), \Psi(L_\beta)] = \Psi\left([L_\alpha , L_\beta]\right) = \Psi\left(C^\gamma_{\alpha\beta} L_\gamma\right)$$
$$= C^\gamma_{\alpha\beta}\, \Psi(L_\gamma) = C^\gamma_{\alpha\beta} P_\gamma$$

which verifies (10).

Example: Let $G=SO(3)$, the group of (3×3) real orthogonal matrices with unit determinant. It is a 3-parameter Lie group [1,5] and thus the associated Lie algebra *so*(3) is 3-dimensional. The basis of *so*(3) consists of the (3×3) antisymmetric matrices

$$L_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} , \quad L_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} , \quad L_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

with commutation relations

$$[\, L_i \,,\, L_j \,] = \varepsilon_{ijk} L_k \quad \text{(sum on } k)$$

where ε_{ijk} is antisymmetric in all pairs of indices, with $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = 1$. [We use Latin instead of Greek indices for the basis elements of *so*(3) since the number of these elements matches the dimensions of R^3, on which space both the *SO*(3) and *so*(3) matrices act.] We notice that

$$(L_i)^j{}_k = -\,\varepsilon_{ijk} \;.$$

The operator representation of the basis of *so*(3) is, according to (9),

$$P_i = -\,(L_i)^j{}_k\, x^k \partial_j \;=\; \varepsilon_{ijk}\, x^k \partial_j$$

or, analytically,

$$P_1 = x^3 \partial_2 - x^2 \partial_3 \;, \quad P_2 = x^1 \partial_3 - x^3 \partial_1 \;, \quad P_3 = x^2 \partial_1 - x^1 \partial_2 \;.$$

The reader may check that $[P_i \,,\, P_j] = \varepsilon_{ijk} P_k$.

3. Group operators for general coordinate transformations

The previous results are valid for *linear* (matrix) groups, in which case $g\vec{x}$ represents the action of an $(m\times m)$ matrix on a vector of R^m. More generally, consider an m-dimensional manifold M with coordinates $(x^1, \dots, x^m)$ and let G be an n-dimensional local Lie group of coordinate transformations on M (see [6] for rigorous definitions and examples). The elements g of G depend on n real parameters $(a^1, \dots, a^n)$. We call $x \equiv (x^1, \dots, x^m)$ a point on M and we denote by gx a (possibly nonlinear) coordinate transformation on this manifold. To the first order in the group parameters a^λ, i.e., for infinitesimal δa^λ, such a transformation is approximately linear in the δa^λ. We write:

$$(gx)^i \simeq x^i + \delta x^i \quad \text{where} \quad \delta x^i = \delta a^\lambda U_\lambda^i(x^k) \tag{12}$$

$(i = 1, \dots, m;\ \lambda = 1, \dots, n)$.

Let $F(x)$ be an arbitrary function on M. As before, we define the operators

$$T(g):\ F \to T(g)F, \quad g \in G$$

by

$$[T(g)F](x) = F\left(g^{-1}x\right) \tag{13}$$

Again, the $T(g)$ constitute an operator representation of G:

$$T(g_1 g_2) = T(g_1)\, T(g_2)\ .$$

[Careful: $g_1 g_2$ is no longer a matrix product but a succession of coordinate transformations! It is still true, however, that $(g_1 g_2)^{-1} = g_2^{-1} g_1^{-1}$.]

Given that, by (12),

$$(gx)^i \simeq x^i + \delta a^\lambda U_\lambda^i(x^k)$$

we have that

$$(g^{-1}x)^i \simeq x^i - \delta a^\lambda U_\lambda^i(x^k)\ .$$

Let us justify this statement:

$$\begin{aligned}(g^{-1}gx)^i &\simeq (gx)^i - \delta a^\lambda U_\lambda^i\left((gx)^k\right)\\ &\simeq x^i + \delta a^\lambda U_\lambda^i(x^k) - \delta a^\lambda U_\lambda^i\left(x^k + \delta a^\rho U_\rho^k\right)\end{aligned}$$

By using Eq. (A.1) in the Appendix we have that, to the first order in the δa^λ,

$$\begin{aligned}\delta a^\lambda U_\lambda^i\left(x^k + \delta a^\rho U_\rho^k\right) &\simeq \delta a^\lambda\left[U_\lambda^i(x^k) + \delta a^\rho U_\rho^j \partial_j U_\lambda^i(x^k)\right]\\ &\simeq \delta a^\lambda U_\lambda^i(x^k)\end{aligned}$$

Thus, finally, $(g^{-1}gx)^i = x^i \iff g^{-1}gx \equiv$ identity transformation.

By using (A.1) once more, the infinitesimal version of (13) is written:

$$[T(g)F](x) \simeq F\left(x^i - \delta a^\lambda U_\lambda^i\right) \simeq F(x) - \delta a^\lambda U_\lambda^i \partial_i F(x)$$
$$= \left(1 - \delta a^\lambda U_\lambda^i \partial_i\right) F(x) \;\Rightarrow$$

$$T(g) \simeq 1 - \delta a^\lambda U_\lambda^i \partial_i \equiv 1 + \delta a^\lambda P_\lambda \qquad (14)$$

where

$$P_\lambda = -U_\lambda^i (x^k)\, \partial_i \qquad (15)$$

It can be proven [3] that the operators P_λ ($\lambda = 1, \dots, n$) form the basis of an n-dimensional Lie algebra:

$$[P_\alpha, P_\beta] = C_{\alpha\beta}^\gamma P_\gamma \qquad (16)$$

Let us see what this implies: Let

$$P_\alpha = -U_\alpha^i (x^k)\, \partial_i \;, \quad P_\beta = -U_\beta^j (x^k)\, \partial_j \;.$$

Then,

$$[P_\alpha, P_\beta] = \left(U_\alpha^i \partial_i U_\beta^j - U_\beta^i \partial_i U_\alpha^j\right) \partial_j \,.$$

A set of real constants $C_{\alpha\beta}^\gamma$ must then exist such that

$$U_\alpha^i \partial_i U_\beta^j - U_\beta^i \partial_i U_\alpha^j = -C_{\alpha\beta}^\gamma U_\gamma^j \qquad (17)$$

Then,

$$[P_\alpha, P_\beta] = -C_{\alpha\beta}^\gamma U_\gamma^j \,\partial_j = C_{\alpha\beta}^\gamma P_\gamma \;.$$

Relations (17) are conditions for closure, under the Lie bracket, of the set of operators spanned by the basis $\{P_\lambda\}$; that is, conditions in order that this set constitute a Lie algebra.

Appendix: Multidimensional Taylor expansion

The Taylor series expansion of a function $f(x)$ about a point x can be written as

$$f(x+h)=\sum_{n=0}^{\infty}\frac{1}{n!}h^n\left(\frac{d}{dx}\right)^n f(x)=f(x)+h\frac{df(x)}{dx}+\cdots$$

We write:

$$f(x+h)=\left[\sum_{n=0}^{\infty}\frac{1}{n!}\left(h\frac{d}{dx}\right)^n\right]f(x)=\exp\left(h\frac{d}{dx}\right)f(x) \;\; .$$

For infinitesimal $h\equiv\delta x$ we may use the approximate expression

$$f(x+\delta x)\simeq f(x)+\delta x\frac{df(x)}{dx} \;\; .$$

More generally, consider a function $\Phi(x^1, x^2,\cdots)\equiv\Phi(\vec{r})$. Let $\vec{a}\equiv(a^1, a^2,\cdots)$ be a constant vector. Then,

$$\Phi(\vec{r}+\vec{a})=\sum_{n=0}^{\infty}\frac{1}{n!}(\vec{a}\cdot\vec{\nabla})^n\Phi(\vec{r})=\Phi(\vec{r})+\vec{a}\cdot\vec{\nabla}\Phi(\vec{r})+\cdots$$

where $\vec{\nabla}\Phi\equiv(\partial\Phi/\partial x^1, \partial\Phi/\partial x^2,\cdots)$. We write:

$$\Phi(\vec{r}+\vec{a})=\left[\sum_{n=0}^{\infty}\frac{1}{n!}(\vec{a}\cdot\vec{\nabla})^n\right]\Phi(\vec{r})=\exp(\vec{a}\cdot\vec{\nabla})\Phi(\vec{r}) \;\; .$$

For infinitesimal $\vec{a}\equiv\delta\vec{r}$,

$$\begin{gathered}\Phi(\vec{r}+\delta\vec{r})\simeq\Phi(\vec{r})+\delta\vec{r}\cdot\vec{\nabla}\Phi(\vec{r})\equiv\Phi(\vec{r})+\delta\Phi \quad \text{where}\\ \delta\Phi=\delta\vec{r}\cdot\vec{\nabla}\Phi(\vec{r})=\delta x^k\,\frac{\partial\Phi(\vec{r})}{\partial x^k} \quad (\text{sum on } k)\end{gathered} \tag{A.1}$$

Indeed, notice that, infinitesimally,

$$\delta\Phi\simeq d\Phi=\frac{\partial\Phi}{\partial x^k}dx^k \quad \text{where} \;\; dx^k\equiv\delta x^k \; .$$

References

1. J. F. Cornwell, *Group Theory in Physics*, Vol. I & II (Academic Press, 1984).
2. D. H. Sattinger, O. L. Weaver, *Lie Groups and Algebras with Applications to Physics, Geometry, and Mechanics* (Springer-Verlag, 1986).
3. R. Gilmore, *Lie Groups, Lie Algebras, and Some of their Applications* (Wiley, 1974).
4. M. Postnikov, *Lie Groups and Lie Algebras* (Mir Publishers, 1986).
5. W.-K. Tung, *Group Theory in Physics* (World Scientific, 1985).
6. P. J. Olver, *Applications of Lie Groups to Differential Equations*, 2nd Edition (Springer, 1993).

Bäcklund transformations: An introduction

Costas J. Papachristou[1]

Hellenic Naval Academy

The concept of a Bäcklund transformation (BT) is introduced. Certain applications of BTs – both older and more recent ones – are discussed.

1. Introduction

Given a difficult problem in mathematics we always look for some way to *transform* it to another problem that is easier to solve. Thus, for example, we seek an integrating factor that might transform a first-order ordinary differential equation into an exact one (or would reduce the order of a higher-order differential equation, in the more general case).

A notoriously difficult problem in the theory of partial differential equations (PDEs) is the case of *nonlinear* PDEs. In contrast to the case of linear PDEs, there is no general method for solving nonlinear ones. Thus, given a nonlinear PDE we look for ways to associate it with some other PDE (preferably a linear one!) whose solutions are already known. For example, the *Burgers equation* $u_t = u_{xx} + 2uu_x$ is a nonlinear PDE for the function $u(x,t)$ (subscripts denote partial derivatives with respect to the indicated variables). This PDE can be transformed into the linear *heat equation* $v_t = v_{xx}$ by using the so-called *Cole-Hopf transformation* $u = v_x / v$. As can be shown, if $v(x,t)$ is a solution of the heat equation then $u(x,t)$ is a solution of the Burgers equation (the converse is not true in general).

Bäcklund transformations (BTs) were originally devised mainly as a tool for obtaining solutions of nonlinear PDEs (see [1] and the references therein). They were later also proven useful as *recursion operators* for constructing infinite sequences of nonlocal symmetries and conservation laws of certain types of PDEs [2–6].

In simple terms, a BT is a system of PDEs connecting two fields that are required to independently satisfy two respective PDEs [call them (*a*) and (*b*)] in order for the system to be integrable for either field. We say that the PDEs (*a*) and (*b*) are *integrability conditions* for self-consistency of the BT. If a solution of PDE (*a*) is known, then a solution of PDE (*b*) is obtained simply by integrating the BT, without having to actually solve the latter PDE (which, presumably, would be a harder task). In the case where the two fields satisfy the same PDE, the *auto-BT* produces new solutions of this PDE from old ones.

As described above, a BT is an auxiliary tool for finding solutions of a given (usually nonlinear) PDE, using known solutions of the same or another PDE. Now, suppose the BT itself is the differential system whose solutions we are looking for. As will be seen, one possible way to solve this problem is to first seek parameter-dependent solutions of both integrability conditions of the BT. By properly matching the parameters (provided this is possible) a solution of the given differential system is obtained.

The above method is particularly effective in *linear* problems, given that parametric solutions of linear PDEs are generally easier to find. An important paradigm of a BT associated with a linear problem is offered by the Maxwell system of equations of

[1] cjpapachristou@gmail.com

electromagnetism [7,8]. As is well known, the consistency of this system demands that both the electric and the magnetic field independently satisfy a respective wave equation. The wave equations for the two fields have known, parameter-dependent solutions; namely, monochromatic plane waves with arbitrary amplitudes, frequencies and wave vectors (the "parameters" of the problem). By inserting these solutions into the Maxwell system, one may find the appropriate constraints for the parameters in order for the plane waves to also be solutions of Maxwell's equations.

In Section 2 we review the classical concept of a BT. The solution-generating process by using a BT is demonstrated in a number of examples.

In Sec. 3 a different perception of a BT is presented, according to which it is the BT itself whose solutions are sought. The concept of *parametric conjugate solutions* is introduced.

In Sec. 4 we examine the connection between BTs and recursion operators for generating infinite sequences of nonlocal symmetries of PDEs.

2. Bäcklund transformations and generation of solutions

Let $u(x,t)$ be a function of two variables. For the partial derivatives of u the following notation will be used:

$$\frac{\partial u}{\partial x} = \partial_x u = u_x \ , \quad \frac{\partial u}{\partial t} = \partial_t u = u_t \ , \quad \frac{\partial^2 u}{\partial x^2} = u_{xx} \ , \quad \frac{\partial^2 u}{\partial t^2} = u_{tt} \ , \quad \frac{\partial^2 u}{\partial x \partial t} = u_{xt} \ ,$$

etc. In general, a subscript will denote partial differentiation with respect to the indicated variable.

Let F be a function of x, t, u, as well as of a number of partial derivatives of u. We will denote this type of dependence by writing

$$F(x,t,u,u_x,u_t,u_{xx},u_{tt},u_{xt},\cdots) \equiv F[u] \ .$$

We also write

$$F_x = \partial_x F = \partial F / \partial x \ , \quad F_t = \partial_t F = \partial F / \partial t \ , \quad F_u = \partial_u F = \partial F / \partial u \ ,$$

etc. Note that in determining F_x and F_t we must take into account both the *explicit* and the *implicit* (through u and its partial derivatives) dependence of F on x and t. As an example, for $F[u] = 3xtu^2$ we have $F_x = 3tu^2 + 6xtuu_x$ and $F_t = 3xu^2 + 6xtuu_t$.

Consider now two partial differential equations (PDEs) $P[u]=0$ and $Q[v]=0$ for the unknown functions u and v, respectively, where the bracket notation introduced above is adopted. Both u and v are functions of two variables x, t. Independently, for the moment, consider also a pair of coupled PDEs for u and v:

$$B_1[u,v]=0 \quad (a) \qquad B_2[u,v]=0 \quad (b) \tag{1}$$

where the expressions $B_i\,[u,v]$ (i=1,2) may contain u, v as well as partial derivatives of u and v with respect to x and t. We note that u appears in both equations (a) and (b). The question then is: if we find an expression for u by integrating (a) for a given v, will it match the corresponding expression for u found by integrating (b) for the same v? The answer is that, in order that (a) and (b) be consistent with each other for solution for u, the function v must be properly chosen so as to satisfy a certain *consistency condition* (or *integrability condition* or *compatibility condition*).

By a similar reasoning, in order that (*a*) and (*b*) in (1) be mutually consistent for solution for *v*, for some given *u*, the function *u* must now itself satisfy a corresponding integrability condition.

If it happens that the two consistency conditions for integrability of the system (1) are precisely the PDEs $P[u]=0$ and $Q[v]=0$, we say that the above system constitutes a *Bäcklund transformation* (BT) connecting solutions of $P[u]=0$ with solutions of $Q[v]=0$. In the special case where $P \equiv Q$, i.e., when *u* and *v* satisfy *the same* PDE, the system (1) is called an *auto-Bäcklund* transformation (auto-BT) for this PDE.

Suppose now that we seek solutions of the PDE $P[u]=0$. Assume that we are able to find a BT connecting solutions *u* of this equation with solutions *v* of the PDE $Q[v]=0$ (if $P \equiv Q$, the auto-BT connects solutions *u* and *v* of the same PDE) and let $v=v_0(x,t)$ be some known solution of $Q[v]=0$. The BT is then a system of PDEs for the unknown *u*,

$$B_i\,[u, v_0] = 0 \;, \quad i = 1, 2 \tag{2}$$

The system (2) is integrable for *u*, given that the function v_0 satisfies *a priori* the required integrability condition $Q[v]=0$. The solution *u* then of the system satisfies the PDE $P[u]=0$. Thus a solution $u(x,t)$ of the latter PDE is found without actually solving the equation itself, simply by integrating the BT (2) with respect to *u*. Of course, this method will be useful provided that integrating the system (2) for *u* is simpler than integrating the PDE $P[u]=0$ itself. If the transformation (2) is an auto-BT for the PDE $P[u]=0$, then, starting with a known solution $v_0(x,t)$ of this equation and integrating the system (2), we find another solution $u(x,t)$ of the same equation.

Let us see some examples of the use of a BT to generate solutions of a PDE:

1. The *Cauchy-Riemann relations* of Complex Analysis,

$$u_x = v_y \quad (a) \qquad u_y = -v_x \quad (b) \tag{3}$$

(where the variable *t* has here been renamed *y*) constitute an auto-BT for the *Laplace equation*,

$$P\,[w] \equiv w_{xx} + w_{yy} = 0 \tag{4}$$

Let us explain this: Suppose we want to solve the system (3) for *u*, for a given choice of the function $v(x,y)$. To see if the PDEs (*a*) and (*b*) match for solution for *u*, we must compare them in some way. We thus differentiate (*a*) with respect to *y* and (*b*) with respect to *x*, and equate the mixed derivatives of *u*. That is, we apply the integrability condition $(u_x)_y = (u_y)_x$. In this way we eliminate the variable *u* and find the condition that must be obeyed by $v(x,y)$:

$$P\,[v] \equiv v_{xx} + v_{yy} = 0 \;\; .$$

Similarly, by using the integrability condition $(v_x)_y = (v_y)_x$ to eliminate *v* from the system (3), we find the necessary condition in order that this system be integrable for *v*, for a given function $u(x,y)$:

$$P\,[u] \equiv u_{xx} + u_{yy} = 0 \;\; .$$

In conclusion, the integrability of system (3) with respect to either variable u or v requires that the other variable must satisfy the Laplace equation (4).

Let now $v_0(x,y)$ be a known solution of the Laplace equation (4). Substituting $v=v_0$ in the system (3), we can integrate this system with respect to u. As can be shown by eliminating v_0 from the system, the solution u will also satisfy the Laplace equation (4). As an example, by choosing the solution $v_0(x,y)=xy$ we find a new solution $u(x,y)=(x^2-y^2)/2+C$.

2. The *Liouville equation* is written

$$P[u] \equiv u_{xt} - e^u = 0 \quad \Leftrightarrow \quad u_{xt} = e^u \tag{5}$$

Due to its nonlinearity, this PDE is hard to integrate directly. A solution is thus sought by means of a BT. We consider an auxiliary function $v(x,t)$ and an associated PDE,

$$Q[v] \equiv v_{xt} = 0 \tag{6}$$

We also consider the system of first-order PDEs,

$$u_x + v_x = \sqrt{2}\; e^{(u-v)/2} \quad (a) \qquad u_t - v_t = \sqrt{2}\; e^{(u+v)/2} \quad (b) \tag{7}$$

Differentiating the PDE (a) with respect to t and the PDE (b) with respect to x, and eliminating $(u_t - v_t)$ and (u_x+v_x) in the ensuing equations with the aid of (a) and (b), we find that u and v satisfy the PDEs (5) and (6), respectively. Thus, the system (7) is a BT connecting solutions of (5) and (6). Starting with the trivial solution v=0 of (6), and integrating the system (7), which reads

$$u_x = \sqrt{2}\; e^{u/2} \;, \qquad u_t = \sqrt{2}\; e^{u/2} \tag{7a}$$

we find a nontrivial solution of (5):

$$u(x,t) = -2\ln\left(C - \frac{x+t}{\sqrt{2}} \right)$$

(see Appendix).

3. The "*sine-Gordon" equation* has applications in various areas of Physics, e.g., in the study of crystalline solids, in the transmission of elastic waves, in magnetism, in elementary-particle models, etc. The equation (whose name is a pun on the related linear Klein-Gordon equation) is written

$$P[u] \equiv u_{xt} - \sin u = 0 \quad \Leftrightarrow \quad u_{xt} = \sin u \tag{8}$$

The following system of equations is an auto-BT for the nonlinear PDE (8):

$$\frac{1}{2}(u+v)_x = a\,\sin\left(\frac{u-v}{2}\right) \;, \qquad \frac{1}{2}(u-v)_t = \frac{1}{a}\,\sin\left(\frac{u+v}{2}\right) \tag{9}$$

where a ($\neq 0$) is an arbitrary real constant. [Because of the presence of a, the system (9) is called a *parametric* BT.] When u is a solution of (8) the BT (9) is integrable for v, which, in turn, also is a solution of (8): $P[v]=0$; and vice versa. Starting with the trivial solution $v=0$ of $v_{xt}=\sin v$, and integrating the system (9), which reads

$$u_x = 2a \sin\frac{u}{2} \quad , \quad u_t = \frac{2}{a}\sin\frac{u}{2} \tag{9a}$$

we obtain a new solution of (8):

$$u(x,t) = 4\arctan\left\{C\exp\left(ax+\frac{t}{a}\right)\right\}$$

(see Appendix).

3. Method of parametric conjugate solutions

As presented in the previous section, a BT is an auxiliary device for constructing solutions of a (usually nonlinear) PDE from known solutions of the same or another PDE. The related problem where solutions of the differential system representing the BT itself are sought is also of interest, however, and has been studied in connection with the Maxwell equations of electromagnetism [7,8].

To be specific, assume that we need to integrate a given system of PDEs connecting two unknown functions $u(x,y)$ and $v(x,y)$:

$$B_i[u,v]=0 \; , \quad i=1,2 \tag{10}$$

Suppose that the integrability of the above system for both functions requires that u and v separately satisfy the respective PDEs

$$P[u]=0 \quad (a) \qquad Q[v]=0 \quad (b) \tag{11}$$

That is, the system (10) is a BT connecting solutions of the PDEs (11). Assume, now, that these PDEs possess known *parameter-dependent solutions* of the form

$$u=f(x,y;\alpha,\beta,\ldots) \quad , \qquad v=g(x,y;\kappa,\lambda,\ldots) \tag{12}$$

where α, β, κ, λ, etc., are (real or complex) parameters. If values of these parameters can be determined for which u and v jointly satisfy the system (10), we say that the solutions u and v of the PDEs (11a) and (11b), respectively, are *conjugate through the BT* (10) (or *BT-conjugate*, for short). By finding a pair of BT-conjugate solutions (12) one thus automatically obtains a solution of the system (10).

Note that solutions of *both* integrability conditions (11) of the system (10) must now be known in advance! From the practical point of view the method is thus most applicable in *linear* problems, since it is much easier to find parameter-dependent solutions of the PDEs (11) in this case.

Let us see an example: Going back to the Cauchy-Riemann relations (3), which is an auto-BT connecting solutions of the Laplace equation (4), we try the following parametric solutions of the latter PDE:

$$u(x,y) = \alpha(x^2 - y^2) + \beta x + \gamma y \ ,$$
$$v(x,y) = \kappa xy + \lambda x + \mu y \ .$$

Substituting these expressions into the BT (3), we find that $\kappa=2\alpha$, $\mu=\beta$ and $\lambda=-\gamma$. Therefore, the solutions

$$u(x,y) = \alpha(x^2 - y^2) + \beta x + \gamma y \ ,$$
$$v(x,y) = 2\alpha xy - \gamma x + \beta y$$

of the Laplace equation are BT-conjugate through the Cauchy-Riemann relations.

As a counter-example, let us try a different combination of parametric solutions:

$$u(x,y) = \alpha xy \ , \quad v(x,y) = \beta xy \ .$$

Inserting these into the system (3) and taking into account the independence of x and y, we find that the only possible values of the parameters α and β are $\alpha=\beta=0$, so that $u(x,y)= v(x,y)=0$. Thus, no non-trivial BT-conjugate solutions exist in this case.

4. BTs as recursion operators for symmetries of PDEs

The concept of symmetries of PDEs has been extensively discussed in [1] and [9]. Let us review the main ideas:

Consider a PDE $F[u]=0$, where $u=u(x,t)$. A transformation $u(x,t) \to u'(x,t)$ from the function u to a new function u' represents a *symmetry* of this PDE if the following condition is satisfied: $u'(x,t)$ is a solution of $F[u]=0$ *if* $u(x,t)$ is a solution. That is,

$$F[u'] = 0 \quad \text{when} \quad F[u] = 0 \tag{13}$$

An *infinitesimal symmetry transformation* is written

$$u' = u + \delta u = u + \alpha Q[u] \tag{14}$$

where α is an infinitesimal parameter. The function $Q[u] \equiv Q(x, t, u, u_x, u_t, ...)$ is called the *symmetry characteristic* of the transformation (14).

In order that a function $Q[u]$ be a symmetry characteristic for the PDE $F[u]=0$, it must satisfy a certain PDE that expresses the *symmetry condition* for $F[u]=0$. We write, symbolically,

$$S(Q; u) = 0 \quad \text{when} \quad F[u] = 0 \tag{15}$$

where the expression S depends *linearly* on Q and its partial derivatives. Thus, (15) is a linear PDE for Q, in which equation the variable u enters as a sort of parametric function that is required to satisfy the PDE $F[u]=0$.

A *recursion operator* $\hat{R}$ [10] is a linear operator which, acting on any symmetry characteristic Q, produces a new symmetry characteristic $Q' = \hat{R} Q$. That is,

$$S(\hat{R} Q; u) = 0 \quad \text{when} \quad S(Q; u) = 0 \tag{16}$$

It is easy to show that *any power of a recursion operator also is a recursion operator*. This means that, starting with any symmetry characteristic Q, one may in principle obtain an infinite set of characteristics (thus, an infinite number of symmetries) by repeated application of the recursion operator.

A new approach to recursion operators was suggested in the early 1990s [2,3] (see also [4-6] and [11-13]). According to this view, a recursion operator for the PDE $F[u]=0$ is an auto-BT for the linear PDE (15) that expresses the symmetry condition of $F[u]=0$; that is, a BT producing new solutions Q' of (15) from old ones, Q. Typically, this type of BT produces *nonlocal* symmetries, i.e., symmetry characteristics depending on *integrals* (rather than derivatives) of u.

As an example, consider the *chiral field equation*

$$F[g] \equiv (g^{-1}g_x)_x + (g^{-1}g_t)_t = 0 \tag{17}$$

(as usual, subscripts denote partial differentiations) where g is a $GL(n,C)$-valued function of x and t (i.e., an invertible complex $n\times n$ matrix, differentiable for all x, t).

Let $Q[g]$ be a symmetry characteristic of the PDE (17). It is convenient to put

$$Q[g] = g\,\Phi[g]$$

and write the corresponding infinitesimal symmetry transformation in the form

$$g' = g + \delta g = g + \alpha g\,\Phi[g] \tag{18}$$

The symmetry condition that Q must satisfy will be a PDE linear in Q, thus in Φ also. As can be shown [9] this PDE is

$$S(\Phi;g) \equiv \Phi_{xx} + \Phi_{tt} + [g^{-1}g_x\,,\,\Phi_x] + [g^{-1}g_t\,,\,\Phi_t] = 0 \tag{19}$$

which must be valid when $F[g]=0$ (where, in general, $[A, B] \equiv AB-BA$ denotes the commutator of two matrices A and B).

For a given g satisfying $F[g]=0$, consider now the following system of PDEs for the matrix functions Φ and Φ':

$$\begin{aligned} \Phi'_x &= \Phi_t + [g^{-1}g_t\,,\,\Phi] \\ -\Phi'_t &= \Phi_x + [g^{-1}g_x\,,\,\Phi] \end{aligned} \tag{20}$$

The integrability condition $(\Phi'_x)_t = (\Phi'_t)_x$, together with the equation $F[g]=0$, require that Φ be a solution of (19): $S(\Phi\,;g) = 0$. Similarly, by the integrability condition $(\Phi_t)_x = (\Phi_x)_t$ one finds, after a lengthy calculation: $S(\Phi';g) = 0$.

In conclusion, for any g satisfying the PDE (17), the system (20) is a BT relating solutions Φ and Φ' of the symmetry condition (19) of this PDE; that is, relating different symmetries of the chiral field equation (17). Thus, if a symmetry characteristic $Q=g\Phi$ of (17) is known, a new characteristic $Q'=g\Phi'$ may be found by integrating the BT (20); the converse is also true. Since the BT (20) produces new symmetries from old ones, it may be regarded as a *recursion operator* for the PDE (17).

As an example, for any constant matrix M the choice $\Phi=M$ clearly satisfies the symmetry condition (19). This corresponds to the symmetry characteristic $Q=gM$. By integrating the BT (20) for Φ', we get $\Phi'=[X, M]$ and $Q'=g[X, M]$, where X is the "potential" of the PDE (17), defined by the system of PDEs

$$X_x = g^{-1} g_t \ , \quad -X_t = g^{-1} g_x \tag{21}$$

Note the *nonlocal* character of the BT-produced symmetry Q', due to the presence of the potential X. Indeed, as seen from (21), in order to find X one has to *integrate* the chiral field g with respect to the independent variables x and t. The above process can be continued indefinitely by repeated application of the recursion operator (20), leading to an infinite sequence of increasingly nonlocal symmetries.

Appendix

We describe the process of integrating the BTs (7*a*) and (9*a*) for the Liouville equation and the sine-Gordon equation, respectively.

1. The system (7*a*) reads

$$u_x = \sqrt{2}\; e^{u/2} \tag{A.1}$$

$$u_t = \sqrt{2}\; e^{u/2} \tag{A.2}$$

We integrate (A.1) for x, treating t as constant:

$$\frac{du}{dx} = \sqrt{2}\; e^{u/2} \;\Rightarrow\; \int e^{-u/2} du = \sqrt{2}\int dx \;\Rightarrow\; e^{-u/2} = -\frac{x}{\sqrt{2}} + h(t)$$

[where $h(t)$ is a function to be determined], from which we have that

$$u = -2\ln\left[-\frac{x}{\sqrt{2}} + h(t)\right] \quad \text{and therefore} \quad u_t = \frac{-2h'(t)}{-\dfrac{x}{\sqrt{2}} + h(t)} \ .$$

Substituting the above results into (A.2), we get:

$$h'(t) = -\frac{1}{\sqrt{2}} \quad \Rightarrow \quad h(t) = -\frac{t}{\sqrt{2}} + C \ .$$

Thus we finally have:

$$u(x,t) = -2\ln\left(C - \frac{x+t}{\sqrt{2}}\right) \ .$$

2. The system (9*a*) reads

$$u_x = 2a\sin\frac{u}{2} \tag{A.3}$$

$$u_t = \frac{2}{a}\sin\frac{u}{2} \tag{A.4}$$

Integrating (A.3) for x and using the integral formula

$$\int \frac{du}{\sin ku} = \frac{1}{k}\ln\left(\tan\frac{ku}{2}\right)$$

we have:

$$\frac{du}{dx} = 2a\sin\frac{u}{2} \;\Rightarrow\; \int \frac{du}{\sin(u/2)} = 2a\int dx \;\Rightarrow$$

$$\ln\left(\tan\frac{u}{4}\right) = ax + g(t) \tag{A.5}$$

Similarly, integrating (A.4) for t we find:

$$\ln\left(\tan\frac{u}{4}\right) = \frac{t}{a} + h(x) \tag{A.6}$$

By comparing (A.5) and (A.6) we have that

$$ax + g(t) = \frac{t}{a} + h(x) \;\Rightarrow\; h(x) - ax = g(t) - \frac{t}{a} \;.$$

But, a function of x cannot be identically equal to a function of t unless both are equal to the same constant C: $h(x) - ax = g(t) - t/a = C \;\Rightarrow$

$$h(x) = ax + C \;, \quad g(t) = \frac{t}{a} + C \;.$$

From (A.5) and (A.6) we then get

$$\ln\left(\tan\frac{u}{4}\right) = ax + \frac{t}{a} + C \;\Rightarrow\; \text{(by putting } C \text{ in place of } e^C \text{)}$$

$$u(x,t) = 4\arctan\left\{C\exp\left(ax + \frac{t}{a}\right)\right\} \;.$$

References

1. C. J. Papachristou, *Symmetry and integrability of classical field equations*, https://arxiv.org/abs/0803.3688.
2. C. J. Papachristou, *Potential symmetries for self-dual gauge fields*, Phys. Lett. A 145 (1990) 250.
3. C. J. Papachristou, *Lax pair, hidden symmetries, and infinite sequences of conserved currents for self-dual Yang-Mills fields*, J. Phys. A 24 (1991) L 1051.
4. C. J. Papachristou, *Symmetry, conserved charges, and Lax representations of nonlinear field equations: A unified approach*, Electron. J. Theor. Phys. 7, No. 23 (2010) 1.
5. C. J. Papachristou, B. K. Harrison, *Bäcklund-transformation-related recursion operators: Application to the self-dual Yang-Mills equation*, J. Nonlin. Math. Phys., Vol. 17, No. 1 (2010) 35.
6. C. J. Papachristou, *Symmetry and integrability of a reduced, 3-dimensional self-dual gauge field model*, Electron. J. Theor. Phys. 9, No. 26 (2012) 119.
7. C. J. Papachristou, *The Maxwell equations as a Bäcklund transformation*, Advanced Electromagnetics, Vol. 4, No. 1 (2015) 52, https://aemjournal.org/index.php/AEM/article/view/311.
8. C. J. Papachristou, A. N. Magoulas, *Bäcklund transformations: Some old and new perspectives*, Nausivios Chora, Vol. 6 (2016) C-3, https://nausivios.hna.gr/docs/2016C.pdf.
9. C. J. Papachristou, *Infinitesimal symmetry transformations of matrix-valued differential equations: An algebraic approach*, Nausivios Chora, Vol. 7 (2018) C-31, https://nausivios.hna.gr/docs/2018C3.pdf.
10. P. J. Olver, *Applications of Lie Groups to Differential Equations* (Springer-Verlag, 1993).
11. G. A. Guthrie, *Recursion operators and non-local symmetries*, Proc. R. Soc. Lond. A 446 (1994) 107.
12. M. Marvan, *Another look on recursion operators*, in *Differential Geometry and Applications* (Proc. Conf., Brno, 1995) 393.
13. M. Marvan, A. Sergyeyev, *Recursion operators for dispersionless integrable systems in any dimension*, Inverse Problems 28 (2012) 025011.

Symmetry operators and generation of symmetry transformations of partial differential equations

Costas J. Papachristou

Department of Physical Sciences, Hellenic Naval Academy, Piraeus, Greece
E-mail: cjpapachristou@gmail.com

Abstract. The study of symmetries of partial differential equations (PDEs) has been traditionally treated as a geometrical problem. Although geometrical methods have been proven effective with regard to finding infinitesimal symmetry transformations, they present certain conceptual difficulties in the case of matrix-valued PDEs; for example, the usual differential-operator representation of the symmetry-generating vector fields is not possible in this case. In this article an algebraic approach to the symmetry problem of PDEs – both scalar and matrix-valued – is described, based on abstract operators (characteristic derivatives) that admit a standard differential-operator representation in the case of scalar-valued PDEs. A number of examples are given.

Keywords: Matrix-valued differential equations, symmetry transformations, Lie algebras, recursion operators

1. Introduction

The problem of symmetries of a system of partial differential equations (PDEs) has been traditionally treated as a geometrical problem in the jet space of the independent and the dependent variables (including a sufficient number of partial derivatives of the latter variables with respect to the former ones). Two more or less equivalent approaches have been adopted: (*a*) invariance of the system of PDEs itself, under infinitesimal transformations generated by corresponding vector fields in the jet space [1]; (*b*) invariance of a differential ideal of differential forms representing the system of PDEs, under the Lie derivative with respect to the vector fields representing the symmetry transformations [2-6].

Although effective with regard to calculating symmetries, these geometrical approaches suffer from a certain drawback of conceptual nature when it comes to matrix-valued PDEs. The problem is related to the inevitably mixed nature of the coordinates in the jet space (scalar independent variables versus matrix-valued dependent ones) and the need for a differential-operator representation of the symmetry vector fields. How does one define differentiation with respect to matrix-valued variables? Moreover, how does one calculate the Lie bracket of two differential operators in which some (or all) of the variables, as well as the coefficients of partial derivatives with respect to these variables, are matrices?

Although these difficulties were handled in some way in [4-6], it was eventually realized that an alternative, purely algebraic approach to the symmetry problem would be more appropriate in the case of matrix-valued PDEs. Elements of this approach were presented in [7] and later applied in particular problems [8-10]; no formal theoretical framework was fully developed, however.

An attempt to develop such a framework is made in this article. In Sections 2 and 3 we introduce the concept of *characteristic derivatives* – an abstract generalization of vector fields in differential-operator form – and we demonstrate the Lie-algebraic character of the set of these derivatives.

The general symmetry problem for both scalar and matrix-valued PDEs is presented in Sec. 4, and the Lie-algebraic property of symmetries of a PDE is proven in Sec. 5. In Sec. 6 we discuss the concept of a *recursion operator* [1,8-14] by which an infinite set of symmetries may in principle be produced from any known symmetry. An application of these ideas is made in Sec. 7 by using the chiral field equation as an example.

A symmetry of a PDE amounts to the invariance of this equation under the action of a corresponding characteristic derivative. Given the latter operator an *infinitesimal* symmetry of the PDE may be defined. Section 8 discusses the use of symmetry operators to construct *finite* one-parameter symmetry transformations of PDEs. As a pedagogical example, a number of point-symmetry transformations for the two-dimensional Laplace equation are derived in Sec. 9.

To simplify our formalism we will restrict our analysis to the case of a single PDE in one dependent variable. For systems of scalar-valued PDEs in several dependent variables see, e.g., [1].

2. The fundamental operators

A PDE for the unknown function $u=u(x^1, x^2, \dots) \equiv u(x^k)$ [where by (x^k) we collectively denote the independent variables $x^1, x^2, \dots$] is an expression of the form $F[u]=0$, where $F[u] \equiv F(x^k, u, u_k, u_{kl}, \dots)$ is a function in the *jet space* [1] of the independent variables (x^k), the dependent variable u, and the partial derivatives of various orders of u with respect to the x^k, which derivatives will be denoted by using subscripts: u_k, u_{kl}, u_{klm}, etc. A *solution* of the PDE is any function $u=\varphi(x^k)$ for which the relation $F[u]=0$ is satisfied.

The dependent variable u, as well as all functions $F[u]$ in the jet space, will generally be assumed to be square-matrix-valued of fixed (but otherwise unspecified) matrix dimensions. In particular, we require that, in its most general form, a function $F[u]$ in the jet space is expressible as a finite or an infinite sum of products of alternating x-dependent and u-dependent terms, of the form

$$F[u] = \sum a(x^k)\,\Pi[u]\,b(x^k)\,\Pi'[u]\,c(x^k)\cdots \tag{2.1}$$

where the $a(x^k)$, $b(x^k)$, $c(x^k)$, etc., are (generally) matrix-valued and where the matrices $\Pi[u]$, $\Pi'[u]$, etc., are products of variables u, u_k, u_{kl}, etc., of the "fiber" space (or, more generally, products of powers of these variables). The set of all functions (2.1) is thus a (generally) non-commutative algebra.

If u is a scalar quantity, a total derivative operator can be defined in the usual differential-operator form

$$D_i = \frac{\partial}{\partial x^i} + u_i \frac{\partial}{\partial u} + u_{ij} \frac{\partial}{\partial u_j} + u_{ijk} \frac{\partial}{\partial u_{jk}} + \cdots \tag{2.2}$$

where the summation convention over repeated up-and-down indices (such as j and k in this equation) has been adopted and will be used throughout. If, however, u is matrix-valued, the above expression is obviously not valid. A generalization of the definition of the total derivative is thus necessary for matrix-valued PDEs.

Definition 2.1. The *total derivative operator* with respect to the variable x^i is a linear operator D_i acting on functions $F[u]$ of the form (2.1) in the jet space and having the following properties:

1. On functions $f(x^k)$ in the base space, $D_i\, f(x^k) = \partial f/\partial x^i \equiv \partial_i\, f(x^k)$.

2. For $F[u]=u, u_i, u_{ij}$, etc., we have: $D_i\, u = u_i$, $D_i\, u_j = D_j\, u_i = u_{ij} = u_{ji}$, etc.

3. The operator D_i is a *derivation* on the algebra of all matrix-valued functions of the form (2.1) in the jet space; i.e., the *Leibniz rule* is satisfied:

$$D_i\left(F[u]G[u]\right) = \left(D_i F[u]\right)G[u] + F[u]\,D_i G[u] \tag{2.3}$$

Higher-order total derivatives $D_{ij}=D_iD_j$ may similarly be defined but they no longer possess the derivation property. Given that $\partial_i\partial_j=\partial_j\partial_i$ and that $u_{ij}=u_{ji}$, it follows that $D_iD_j = D_jD_i \Leftrightarrow D_{ij} = D_{ji}$; that is, total derivatives commute. We write: $[D_i\, , D_j]=0$, where, in general, $[A, B] \equiv AB–BA$ will denote the *commutator* of two operators or two matrices, as the case may be.

If u^{-1} is the inverse of u, such that $uu^{-1}= u^{-1}u = \mathbf{1}$, then we can define

$$D_i\left(u^{-1}\right) \equiv -u^{-1}\left(D_i u\right)u^{-1} \tag{2.4}$$

Moreover, for any functions $A[u]$ and $B[u]$ in the jet space it can be shown that

$$D_i\,[A,B] = [D_i A,\, B] + [A,\, D_i B] \tag{2.5}$$

As an example, let $(x^1, x^2) \equiv (x, t)$ and let $F[u]=xtu_x^2$, where u is matrix-valued. Writing $F[u]=xtu_xu_x$, we have: $D_t F[u]=xu_x^2 + xt\,(u_{xt}\, u_x + u_x\, u_{xt})$.

Let now $Q[u] \equiv Q\,(x^k, u, u_k\, , u_{kl}\, , ...)$ be a function in the jet space. We will call this a *characteristic function* (or simply a *characteristic*) of a certain derivative, defined as follows:

Definition 2.2. The *characteristic derivative* with respect to $Q[u]$ is a linear operator Δ_Q acting on functions $F[u]$ in the jet space and having the following properties:

1. On functions $f(x^k)$ in the base space,

$$\Delta_Q\, f(x^k) = 0 \tag{2.6}$$

(that is, Δ_Q acts only in the fiber space).

2. For $F[u]=u$,

$$\Delta_Q\, u = Q[u] \tag{2.7}$$

3. Δ_Q commutes with total derivatives:

$$\Delta_Q D_i = D_i \Delta_Q \iff [\Delta_Q\, , D_i\,] = 0 \quad (\text{all } i) \tag{2.8}$$

4. The operator Δ_Q is a *derivation* on the algebra of all matrix-valued functions of the form (2.1) in the jet space (the Leibniz rule is satisfied):

$$\Delta_Q(F[u]G[u]) = (\Delta_Q F[u])\,G[u] + F[u]\,\Delta_Q G[u] \qquad (2.9)$$

Corollary: By (2.7) and (2.8) we have:

$$\Delta_Q u_i = \Delta_Q D_i u = D_i Q[u] \qquad (2.10)$$

We note that the operator Δ_Q is a well-defined quantity, in the sense that the action of Δ_Q on u uniquely determines the action of Δ_Q on any function $F[u]$ of the form (2.1) in the jet space. Moreover, since, by assumption, u and $Q[u]$ are matrices of equal dimensions, it follows from (2.7) that Δ_Q preserves the matrix character of u, as well as of any function $F[u]$ on which this operator acts.

We also remark that we have imposed conditions (2.6) and (2.8) having a certain property of symmetries of PDEs in mind; namely, *every* symmetry of a PDE can be represented by a transformation of the dependent variable u alone, i.e., can be expressed as a transformation in the fiber space (see [1], Chap. 5).

The following formulas, analogous to (2.4) and (2.5), may be written:

$$\Delta_Q\left(u^{-1}\right) \equiv -u^{-1}\left(\Delta_Q u\right)u^{-1} \qquad (2.11)$$

$$\Delta_Q[A,B] = [\Delta_Q A, B] + [A, \Delta_Q B] \qquad (2.12)$$

As an example, let $(x^1, x^2) \equiv (x, t)$ and let $F[u]=a(x,t)u^2 b(x,t)+[u_x , u_t]$, where a, b and u are matrices of equal dimensions. Writing $u^2=uu$ and using properties (2.7), (2.9), (2.10) and (2.12), we find: $\Delta_Q F[u]=a(x,t)(Qu+uQ)b(x,t)+[D_x Q, u_t]+[u_x , D_t Q]$.

In the case where u is scalar-valued (thus so is $Q[u]$) the characteristic derivative Δ_Q admits a differential-operator representation of the form

$$\Delta_Q = Q[u]\frac{\partial}{\partial u} + \left(D_i Q[u]\right)\frac{\partial}{\partial u_i} + \left(D_i D_j Q[u]\right)\frac{\partial}{\partial u_{ij}} + \cdots \qquad (2.13)$$

[See [1], Chap. 5, for an analytic proof of property (2.8) in this case.]

3. The Lie algebra of characteristic derivatives

The characteristic derivatives Δ_Q acting on functions $F[u]$ of the form (2.1) in the jet space constitute a *Lie algebra of derivations* on the algebra of the $F[u]$. The proof of this statement is contained in the following three Propositions.

Proposition 3.1. Let Δ_Q be a characteristic derivative with respect to the characteristic $Q[u]$; i.e., $\Delta_Q u=Q[u]$ [cf. Eq. (2.7)]. Let λ be a constant (real or complex). We define the operator $\lambda\Delta_Q$ by the relation

$$(\lambda\Delta_Q)F[u] \equiv \lambda\,(\Delta_Q F[u])\,.$$

Then, $\lambda\Delta_Q$ is a characteristic derivative with characteristic $\lambda Q[u]$. That is,

$$\lambda\,\Delta_Q = \Delta_{\lambda Q} \qquad (3.1)$$

Proof. (*a*) The operator $\lambda\Delta_Q$ is linear, since so is Δ_Q .

(*b*) For $F[u]=u$, $(\lambda\Delta_Q)u=\lambda(\Delta_Q u)=\lambda Q[u]$.

(*c*) $\lambda\Delta_Q$ commutes with total derivatives D_i , since so does Δ_Q .

(*d*) Given that Δ_Q satisfies the Leibniz rule (2.9), it is easily shown that so does $\lambda\Delta_Q$.

Comment: Condition (*c*) would not be satisfied if we allowed λ to be a function of the x^k instead of being a constant, since $\lambda(x^k)$ generally does not commute with the D_i. Therefore, relation (3.1) is not valid for a non-constant λ.

Proposition 3.2. Let Δ_1 and Δ_2 be characteristic derivatives with respect to the characteristics $Q_1[u]$ and $Q_2[u]$, respectively; i.e., $\Delta_1 u=Q_1[u]$, $\Delta_2 u=Q_2[u]$. We define the operator $\Delta_1+\Delta_2$ by

$$(\Delta_1+\Delta_2)\,F[u] \equiv \Delta_1 F[u] + \Delta_2 F[u] \;.$$

Then, $\Delta_1+\Delta_2$ is a characteristic derivative with characteristic $Q_1[u]+Q_2[u]$. That is,

$$\Delta_1 + \Delta_2 = \Delta_Q \quad \text{with} \quad Q[u] = Q_1[u] + Q_2[u] \tag{3.2}$$

Proof. (*a*) The operator $\Delta_1+\Delta_2$ is linear, as a sum of linear operators.

(*b*) For $F[u]=u$, $(\Delta_1+\Delta_2)u = \Delta_1 u+\Delta_2 u = Q_1[u]+Q_2[u]$.

(*c*) $\Delta_1+\Delta_2$ commutes with total derivatives D_i , since so do Δ_1 and Δ_2 .

(*d*) Given that each of Δ_1 and Δ_2 satisfies the Leibniz rule (2.9), it is not hard to show that the same is true for $\Delta_1+\Delta_2$.

Proposition 3.3. Let Δ_1 and Δ_2 be characteristic derivatives with respect to the characteristics $Q_1[u]$ and $Q_2[u]$, respectively; i.e., $\Delta_1 u=Q_1[u]$, $\Delta_2 u=Q_2[u]$. We define the operator $[\Delta_1 , \Delta_2]$ (*Lie bracket* of Δ_1 and Δ_2) by

$$[\Delta_1 , \Delta_2]\, F[u] \equiv \Delta_1 (\Delta_2 F[u]) - \Delta_2 (\Delta_1 F[u]) \;.$$

Then, $[\Delta_1 , \Delta_2]$ is a characteristic derivative with characteristic $\Delta_1 Q_2[u]-\Delta_2 Q_1[u]$. That is,

$$[\Delta_1, \Delta_2] = \Delta_Q \quad \text{with} \quad Q[u] = \Delta_1 Q_2[u] - \Delta_2 Q_1[u] \equiv Q_{1,2}[u] \tag{3.3}$$

Proof. (*a*) The linearity of $[\Delta_1 , \Delta_2]$ follows from the linearity of Δ_1 and Δ_2 .

(*b*) For $F[u]=u$, $[\Delta_1 , \Delta_2]u = \Delta_1 (\Delta_2 u) - \Delta_2 (\Delta_1 u) = \Delta_1 Q_2[u]-\Delta_2 Q_1[u] \equiv Q_{1,2}[u]$.

(*c*) $[\Delta_1 , \Delta_2]$ commutes with total derivatives D_i , since so do Δ_1 and Δ_2 .

(*d*) Given that each of Δ_1 and Δ_2 satisfies the Leibniz rule (2.9), one can show (after some algebra and cancellation of terms) that the same is true for $[\Delta_1 , \Delta_2]$.

In the case where u (thus the Q's also) is scalar-valued, the Lie bracket admits a standard differential-operator representation:

$$[\Delta_1, \Delta_2] = Q_{1,2}[u]\frac{\partial}{\partial u} + \left(D_i Q_{1,2}\right)\frac{\partial}{\partial u_i} + \left(D_i D_j Q_{1,2}\right)\frac{\partial}{\partial u_{ij}} + \cdots \qquad (3.4)$$

where $Q_{1,2}[u] = [\Delta_1 , \Delta_2]\, u = \Delta_1 Q_2[u] - \Delta_2 Q_1[u]$.

The Lie bracket $[\Delta_1 , \Delta_2]$ has the following properties:

1. $[\Delta_1 , a\Delta_2+b\Delta_3] = a\,[\Delta_1 , \Delta_2] + b\,[\Delta_1 , \Delta_3]$;
   $[a\Delta_1+b\Delta_2 , \Delta_3] = a\,[\Delta_1 , \Delta_3] + b\,[\Delta_2 , \Delta_3]$ $\qquad$ (a, b = const.)

2. $[\Delta_1 , \Delta_2] = -[\Delta_2 , \Delta_1]$ $\qquad$ (*antisymmetry*)

3. $[\Delta_1 , [\Delta_2 , \Delta_3]] + [\Delta_2 , [\Delta_3 , \Delta_1]] + [\Delta_3 , [\Delta_1 , \Delta_2]] = 0$;
   $[[\Delta_1 , \Delta_2] , \Delta_3] + [[\Delta_2 , \Delta_3] , \Delta_1] + [[\Delta_3 , \Delta_1] , \Delta_2] = 0$ $\qquad$ (*Jacobi identity*)

4. Infinitesimal symmetry transformations of a PDE

Let $F[u]=0$ be a PDE in the independent variables $x^k \equiv x^1, x^2, \ldots$, and the (generally) matrix-valued dependent variable u. A transformation $u(x^k) \rightarrow u'(x^k)$ from the function u to a new function u' represents a *symmetry* of the PDE if the following condition is satisfied: $u'(x^k)$ is a solution of $F[u]=0$ *when* $u(x^k)$ is a solution; that is, $F[u']=0$ when $F[u]=0$.

We will restrict our attention to *continuous symmetries* and, for the moment, to infinitesimal transformations. Although such symmetries may involve transformations of the independent variables (x^k), they may equivalently be expressed as transformations of u alone (see [1], Chap. 5), i.e., as transformations in the fiber space.

An infinitesimal symmetry transformation is written symbolically as

$$u \rightarrow u' = u + \delta u$$

where δu is an infinitesimal quantity, in the sense that all powers $(\delta u)^n$ with $n>1$ may be neglected. The *symmetry condition* is thus written

$$F[u+\delta u] = 0 \quad \text{when} \quad F[u] = 0 \qquad (4.1)$$

An infinitesimal change δu of u induces a change $\delta F[u]$ of $F[u]$, where

$$\delta F[u] = F[u+\delta u] - F[u] \;\Leftrightarrow\; F[u+\delta u] = F[u] + \delta F[u] \qquad (4.2)$$

Now, if δu is an infinitesimal symmetry and if u is a solution of $F[u]=0$, then $u+\delta u$ also is a solution; that is, $F[u+\delta u]=0$. This means that $\delta F[u]=0$ when $F[u]=0$. The symmetry condition (4.1) is thus written as follows:

$$\delta F[u] = 0 \mod F[u] \qquad (4.3)$$

A finite symmetry transformation (we denote it M) of the PDE $F[u]=0$ produces a one-parameter family of solutions of the PDE from any given solution $u(x^k)$. We express this by writing

$$M: u(x^k) \to \bar{u}(x^k;\alpha) \quad \text{with} \quad \bar{u}(x^k;0) = u(x^k) \tag{4.4}$$

For infinitesimal values of the parameter α,

$$\bar{u}(x^k;\alpha) \simeq u(x^k) + \alpha Q[u] \quad \text{where} \quad Q[u] = \left.\frac{d\bar{u}}{d\alpha}\right|_{\alpha=0} \tag{4.5}$$

The function $Q[u] \equiv Q(x^k, u, u_k, u_{kl}, ...)$ in the jet space is called the *characteristic* of the symmetry (or, the *symmetry characteristic*). Putting

$$\delta u = \bar{u}(x^k;\alpha) - u(x^k) \tag{4.6}$$

we write, for infinitesimal α,

$$\delta u = \alpha Q[u] \tag{4.7}$$

We notice that the infinitesimal operator δ has the following properties:

1. According to its definition (4.2), δ is a linear operator:

$$\delta(F[u]+G[u]) = (F[u+\delta u]+ G[u+\delta u]) - (F[u]+G[u]) = \delta F[u]+\delta G[u] \,.$$

2. By the nature of our symmetry transformations, δ produces changes in the fiber space while it doesn't affect functions $f(x^k)$ in the base space [this is implicitly stated in (4.6)].

3. Since δ represents a difference, it commutes with all total derivatives D_i:

$$\delta(D_i A[u]) = D_i(\delta A[u]) \,.$$

In particular, for $A[u]=u$,

$$\delta u_i = \delta(D_i u) = D_i(\delta u) = \alpha D_i Q[u] \,,$$

where we have used (4.7).

4. Since δ expresses an infinitesimal change, it may be approximated to a differential; in particular, it satisfies the Leibniz rule:

$$\delta(A[u]B[u]) = (\delta A[u])B[u] + A[u]\,\delta B[u] \,.$$

For example, $\delta(u^2) = \delta(uu) = (\delta u)u+u\delta u = \alpha(Qu+uQ)$.

Now, consider the characteristic derivative Δ_Q with respect to the symmetry characteristic $Q[u]$. According to (2.7),

$$\Delta_Q u = Q[u] \tag{4.8}$$

We observe that the infinitesimal operator δ and the characteristic derivative Δ_Q share common properties. From (4.7) and (4.8) it follows that the two linear operators are related by

$$\delta u = \alpha \Delta_Q u \tag{4.9}$$

and, by extension,

$$\delta u_i = \alpha D_i Q[u] = \alpha \Delta_Q u_i \,, \text{ etc.}$$

[see (2.10)]. Moreover, for scalar-valued u and by the infinitesimal character of the operator δ, we may write:

$$\delta F[u] = \frac{\partial F}{\partial u}\delta u + \frac{\partial F}{\partial u_i}\delta u_i + \cdots = \alpha\left(\frac{\partial F}{\partial u}Q[u] + \frac{\partial F}{\partial u_i}D_i Q[u] + \frac{\partial F}{\partial u_{ij}}D_i D_j Q[u] + \cdots\right)$$

while, by (2.13),

$$\Delta_Q F[u] = \frac{\partial F}{\partial u}Q[u] + \frac{\partial F}{\partial u_i}D_i Q[u] + \frac{\partial F}{\partial u_{ij}}D_i D_j Q[u] + \cdots \tag{4.10}$$

The above observations lead us to the conclusion that, in general, the following relation is true:

$$\delta F[u] = \alpha \Delta_Q F[u] \tag{4.11}$$

The symmetry condition (4.3) is thus written:

$$\Delta_Q F[u] = 0 \mod F[u] \tag{4.12}$$

In particular, if u is scalar-valued, the above condition is written

$$\frac{\partial F}{\partial u}Q[u] + \frac{\partial F}{\partial u_i}D_i Q[u] + \frac{\partial F}{\partial u_{ij}}D_i D_j Q[u] + \cdots = 0 \mod F[u] \tag{4.13}$$

which is a linear PDE for $Q[u]$. More generally, for matrix-valued u and for a function $F[u]$ of the form (2.1), the symmetry condition for the PDE $F[u]=0$ is a *linear* PDE for the symmetry characteristic $Q[u]$. We write this PDE symbolically as

$$S(Q;u) \equiv \Delta_Q F[u] = 0 \mod F[u] \tag{4.14}$$

where the function $S(Q;u)$ is linear in Q and all total derivatives of Q. (The linearity of S in Q follows from the Leibniz rule and the specific form (2.1) of $F[u]$.)

Below is a list of formulas that may be useful in calculations:

- $\Delta_Q u_i = D_i Q[u]\,, \quad \Delta_Q u_{ij} = D_i D_j Q[u]\,,$ etc.
- $\Delta_Q u^2 = \Delta_Q (uu) = Q[u]u + uQ[u]$ (etc.)
- $\Delta_Q (u^{-1}) = -u^{-1} (\Delta_Q u)\, u^{-1} = -u^{-1} Q[u]\, u^{-1}$
- $\Delta_Q [A[u], B[u]] = [\Delta_Q A, B] + [A, \Delta_Q B]$

Comment: According to (4.12), $\Delta_Q F[u]$ vanishes if $F[u]$ vanishes. Given that Δ_Q is a linear operator, the reader may wonder whether this condition is identically satisfied (a linear operator acting on a zero function always produces a zero function!). Note, however, that the function $F[u]$ is *not identically* zero; it becomes zero only *for solutions* of the given PDE. What we need to do, therefore, is to first evaluate $\Delta_Q F[u]$ for *arbitrary* u and *then* demand that the result vanish *when* u is a solution of the PDE $F[u]=0$.

An alternative – and perhaps more transparent – version of the symmetry condition (4.12) is the requirement that the following relation be satisfied:

$$\Delta_Q F[u] = \hat{L}\, F[u] \tag{4.15}$$

where $\hat{L}$ is a linear operator acting on functions in the jet space (see, e.g., [1], Chap. 2 and 5, for a rigorous justification of this condition in the case of scalar-valued PDEs). For example, one may have

$$\Delta_Q F[u] = \sum_i \beta_i (x^k) D_i F[u] + \sum_{i,j} \gamma_{ij} (x^k) D_i D_j F[u] + A(x^k) F[u] + F[u] B(x^k)$$

where the β_i and γ_{ij} are scalar-valued while A and B are matrix-valued. Let us see some examples, restricting for the moment our attention to scalar PDEs.

Example 4.1. The *sine-Gordon (s-G) equation* is written

$$F[u] \equiv u_{xt} - \sin u = 0 \,.$$

Here, $(x^1, x^2) \equiv (x, t)$. Since $\sin u$ can be expanded into an infinite series in powers of u, we see that $F[u]$ has the required form (2.1). Moreover, since u is a scalar function, we can write the symmetry condition by using (4.13):

$$S(Q;u) \equiv Q_{xt} - (\cos u)\, Q = 0 \mod F[u]$$

where $S(Q;u) = \Delta_Q F[u]$ and where by subscripts we denote total differentiations with respect to the indicated variables. Let us verify the solution $Q[u] = u_x$. As will be shown in Sec. 9, this characteristic produces the finite symmetry transformation

$$M:\ u(x,t) \to \bar{u}(x,t;\alpha) = u(x+\alpha,\, t) \tag{4.16}$$

which implies that, if $u(x,t)$ is a solution of the s-G equation, then $\bar{u}(x,t) = u(x+\alpha,\, t)$ also is a solution. We have:

$$Q_{xt} - (\cos u)\, Q = (u_x)_{xt} - (\cos u)\, u_x = (u_{xt} - \sin u)_x = D_x F[u] = 0 \mod F[u] \,.$$

Notice that $\Delta_Q F[u]$ is of the form (4.15), with $\hat{L} \equiv D_x$. Similarly, the characteristic $Q[u] = u_t$ corresponds to the symmetry

$$M:\ u(x,t) \to \bar{u}(x,t;\alpha) = u(x,\, t+\alpha) \tag{4.17}$$

That is, if $u(x,t)$ is a solution of the s-G equation, then so is $\bar{u}(x,t)=u(x,\, t+\alpha)$. The symmetries (4.16) and (4.17) reflect the fact that the s-G equation does not contain the variables x and t explicitly. (Of course, this equation has many more symmetries that are not displayed here; see, e.g., [1].)

Example 4.2. The *heat equation* is written

$$F[u] \equiv u_t - u_{xx} = 0 \, .$$

The symmetry condition (4.13) reads

$$S(Q;u) \equiv Q_t - Q_{xx} = 0 \quad \text{mod} \quad F[u]$$

where $S(Q;u)=\Delta_Q F[u]$. As is easy to show, the symmetries (4.16) and (4.17) are valid here, too. Let us now try the solution $Q[u]=u$. We have:

$$Q_t - Q_{xx} = u_t - u_{xx} = F[u] = 0 \quad \text{mod} \quad F[u] \, .$$

As will be shown in Sec. 9, this symmetry corresponds to the transformation

$$M: \; u(x,t) \to \bar{u}(x,t;\alpha) = e^{\alpha} u(x,t) \tag{4.18}$$

and is a consequence of the linearity of the heat equation.

Example 4.3. One form of the *Burgers equation* is

$$F[u] \equiv u_t - u_{xx} - u_x^2 = 0 \, .$$

The symmetry condition (4.13) is written

$$S(Q;u) \equiv Q_t - Q_{xx} - 2u_x Q_x = 0 \quad \text{mod} \quad F[u]$$

where, as always, $S(Q;u)=\Delta_Q F[u]$. Putting $Q=u_x$ and $Q=u_t$, we verify the symmetries (4.16) and (4.17):

$$Q_t - Q_{xx} - 2u_x Q_x = u_{xt} - u_{xxx} - 2u_x u_{xx} = D_x F[u] = 0 \quad \text{mod} \quad F[u]$$
$$Q_t - Q_{xx} - 2u_x Q_x = u_{tt} - u_{xxt} - 2u_x u_{xt} = D_t F[u] = 0 \quad \text{mod} \quad F[u]$$

Note again that $\Delta_Q F[u]$ is of the form (4.15), with $\hat{L} \equiv D_x$ and $\hat{L} \equiv D_t$. Another symmetry is $Q[u]=1$, which corresponds to the transformation (see Sec. 9)

$$M: \; u(x,t) \to \bar{u}(x,t;\alpha) = u(x,t) + \alpha \tag{4.19}$$

and is a consequence of the fact that u enters $F[u]$ only through its derivatives.

Example 4.4. The *wave equation* is written

$$F[u] \equiv u_{tt} - c^2 u_{xx} = 0 \quad (c = \text{const.})$$

and its symmetry condition reads

$$S(Q;u) \equiv Q_{tt} - c^2 Q_{xx} = 0 \mod F[u] .$$

The solution $Q[u] = x\,u_x + t\,u_t$ corresponds to the symmetry transformation (Sec. 9)

$$M: \; u(x,t) \to \bar{u}(x,t;\alpha) = u(e^{\alpha} x, e^{\alpha} t) \tag{4.20}$$

expressing the invariance of the wave equation under a scale change of x and t. [The reader may show that the transformations (4.16) – (4.19) also express symmetries of the wave equation.]

It is remarkable that each of the above PDEs admits an infinite set of symmetry transformations [1]. An effective method for finding such infinite sets is the use of a *recursion operator*, which produces a new symmetry characteristic every time it acts on a known characteristic. More on recursion operators will be said in Sec. 6.

5. The Lie algebra of symmetries

As is well known [1] the set of symmetries of a PDE $F[u]=0$ has the structure of a Lie algebra. Let us demonstrate this property in the context of our formalism.

Proposition 5.1. Let $\mathcal{L}$ be the set of characteristic derivatives Δ_Q with respect to the symmetry characteristics $Q[u]$ of the PDE $F[u]=0$. The set $\mathcal{L}$ is a (finite or infinite-dimensional) Lie subalgebra of the Lie algebra of characteristic derivatives acting on functions $F[u]$ in the jet space (cf. Sec. 3).

Proof. (*a*) Let $\Delta_Q \in \mathcal{L} \Leftrightarrow \Delta_Q F[u]=0$ (mod $F[u]$). If λ is a constant (real or complex, depending on the nature of the problem) then $(\lambda\Delta_Q)F[u] \equiv \lambda\Delta_Q F[u]=0$, which means that $\lambda\Delta_Q \in \mathcal{L}$. Given that $\lambda\Delta_Q = \Delta_{\lambda Q}$ [see Eq. (3.1)] we conclude that, if $Q[u]$ is a symmetry characteristic of $F[u]=0$, then so is $\lambda Q[u]$.

(*b*) Let $\Delta_1 \in \mathcal{L}$ and $\Delta_2 \in \mathcal{L}$ be characteristic derivatives with respect to the symmetry characteristics $Q_1[u]$ and $Q_2[u]$, respectively. Then, $\Delta_1 F[u]=0$, $\Delta_2 F[u]=0$, and hence $(\Delta_1+\Delta_2)F[u] \equiv \Delta_1 F[u]+\Delta_2 F[u]=0$; therefore, $(\Delta_1+\Delta_2) \in \mathcal{L}$. It also follows from Eq. (3.2) that, if $Q_1[u]$ and $Q_2[u]$ are symmetry characteristics of $F[u]=0$, then so is their sum $Q_1[u]+Q_2[u]$.

(*c*) Let $\Delta_1 \in \mathcal{L}$ and $\Delta_2 \in \mathcal{L}$, as above. Then, by (4.15),

$$\Delta_1 F[u] = \hat{L}_1 F[u] , \quad \Delta_2 F[u] = \hat{L}_2 F[u] .$$

Now, by the definition of the Lie bracket and the linearity of both Δ_i and $\hat{L}_i$ $(i=1,2)$ we have:

$$[\Delta_1 , \Delta_2] F[u] = \Delta_1(\Delta_2 F[u]) - \Delta_2(\Delta_1 F[u]) = \Delta_1(\hat{L}_2 F[u]) - \Delta_2(\hat{L}_1 F[u])$$
$$\equiv (\Delta_1 \hat{L}_2 - \Delta_2 \hat{L}_1) F[u] = 0 \mod F[u]$$

We thus conclude that $[\Delta_1, \Delta_2] \in \mathcal{L}$. Moreover, it follows from Eq. (3.3) that, if $Q_1[u]$ and $Q_2[u]$ are symmetry characteristics of $F[u]=0$, then so is the function

$$Q_{1,2}[u] = \Delta_1 Q_2[u] - \Delta_2 Q_1[u] .$$

Assume now that the PDE $F[u]=0$ has an n-dimensional symmetry algebra $\mathcal{L}$ (which may be a finite subalgebra of an infinite-dimensional symmetry Lie algebra). Let $\{\Delta_1, \Delta_2, \ldots, \Delta_n\} \equiv \{\Delta_k\}$, with corresponding symmetry characteristics $\{Q_k\}$, be a set of n linearly independent operators that constitute a basis of $\mathcal{L}$, and let Δ_i, Δ_j be any two elements of this basis. Given that $[\Delta_i, \Delta_j] \in \mathcal{L}$, this Lie bracket must be expressible as a linear combination of the $\{\Delta_k\}$, with constant coefficients. We write

$$[\Delta_i, \Delta_j] = \sum_{k=1}^{n} c_{ij}^k \Delta_k \tag{5.1}$$

where the coefficients of the Δ_k in the sum are the antisymmetric *structure constants* of the Lie algebra $\mathcal{L}$ in the basis $\{\Delta_k\}$.

The operator relation (5.1) can be expressed in an equivalent, characteristic form by allowing the operators on both sides to act on u and by using the fact that $\Delta_k u = Q_k[u]$:

$$[\Delta_i, \Delta_j] u = \left(\sum_{k=1}^{n} c_{ij}^k \Delta_k \right) u = \sum_{k=1}^{n} c_{ij}^k (\Delta_k u) \Rightarrow$$

$$\Delta_i Q_j[u] - \Delta_j Q_i[u] = \sum_{k=1}^{n} c_{ij}^k Q_k[u] \tag{5.2}$$

Example 5.1. One of the several forms of the *Korteweg-de Vries (KdV) equation* is

$$F[u] \equiv u_t + u u_x + u_{xxx} = 0 .$$

The symmetry condition (4.14) is written

$$S(Q; u) \equiv Q_t + Q u_x + u Q_x + Q_{xxx} = 0 \mod F[u] \tag{5.3}$$

where $S(Q; u) = \Delta_Q F[u]$. The KdV equation admits a symmetry Lie algebra of infinite dimensions [1]. This algebra has a finite, 4-dimensional subalgebra $\mathcal{L}$ of *point transformations*. A symmetry operator (characteristic derivative) Δ_Q is determined by its corresponding characteristic $Q[u] = \Delta_Q u$. Thus, a basis $\{\Delta_1, \ldots, \Delta_4\}$ of $\mathcal{L}$ corresponds to a set of four independent characteristics $\{Q_1, \ldots, Q_4\}$. Such a basis of characteristics is the following:

$$Q_1[u] = u_x , \quad Q_2[u] = u_t , \quad Q_3[u] = t u_x - 1 , \quad Q_4[u] = x u_x + 3 t u_t + 2u$$

The $Q_1, \ldots, Q_4$ satisfy the PDE (5.3), since, as we can show,

$$S(Q_1;u)=D_xF[u]\,,\quad S(Q_2;u)=D_tF[u]\,,\quad S(Q_3;u)=tD_xF[u]\,,$$
$$S(Q_4;u)=(5+xD_x+3tD_t)F[u]$$

[Note once more that $\Delta_Q F[u]$ is of the form (4.15) in each case.] Let us now see two examples of calculating the structure constants of $\mathcal{L}$ by application of (5.2). We have:

$$\Delta_1 Q_2-\Delta_2 Q_1=\Delta_1 u_t-\Delta_2 u_x=(\Delta_1 u)_t-(\Delta_2 u)_x=(Q_1)_t-(Q_2)_x=(u_x)_t-(u_t)_x=0$$
$$\equiv\sum_{k=1}^{4}c_{12}^{k}\,Q_k$$

Since the Q_k are linearly independent, we must necessarily have $c_{12}^{k}=0$, $k=1,2,3,4$. Also,

$$\Delta_2 Q_3-\Delta_3 Q_2=\Delta_2(tu_x-1)-\Delta_3 u_t=t(\Delta_2 u)_x-(\Delta_3 u)_t=t(Q_2)_x-(Q_3)_t$$
$$=tu_{tx}-(u_x+tu_{xt})=-u_x=-Q_1\;\equiv\;\sum_{k=1}^{4}c_{23}^{k}\,Q_k$$

Therefore, $c_{23}^{1}=-1$, $c_{23}^{2}=c_{23}^{3}=c_{23}^{4}=0$.

6. Recursion operators

Let $\delta u=\alpha Q[u]$ be an infinitesimal symmetry of the PDE $F[u]=0$, where $Q[u]$ is the symmetry characteristic. For any solution $u(x^k)$ of this PDE, the function $Q[u]$ satisfies the linear PDE

$$S(Q;u)\equiv\Delta_Q F[u]=0 \qquad (6.1)$$

Because of the linearity of (6.1) in Q, the sum $Q_1[u]+Q_2[u]$ of two solutions of this PDE, as well as the multiple $\lambda Q[u]$ of any solution by a constant, also are solutions of (6.1) for a given u. Thus, for any solution u of $F[u]=0$, the solutions $\{Q[u]\}$ of the PDE (6.1) form a linear space, which we call S_u.

A *recursion operator* $\hat{R}$ is a linear operator that maps the space S_u into itself. Thus, if $Q[u]$ is a symmetry characteristic of $F[u]=0$ (i.e., a solution of (6.1) for a given u) then so is $\hat{R}Q[u]$:

$$S(\hat{R}Q;u)=0 \quad \text{when} \quad S(Q;u)=0 \qquad (6.2)$$

Obviously, any power of a recursion operator also is a recursion operator. Thus, starting with any symmetry characteristic $Q[u]$, one may in principle obtain an infinite set of such characteristics by repeated application of the recursion operator.

A new approach to recursion operators was suggested in the early 1990s [11,15-17] (see also [8-10]) according to which a recursion operator may be viewed as an *auto-Bäcklund transformation* (BT) [18] for the symmetry condition (6.1) of the PDE $F[u]=0$. By integrating the BT, one obtains new solutions $Q'[u]$ of the linear PDE (6.1) from known ones, $Q[u]$. Typically, this type of recursion operator produces *nonlocal* symmetries in which the symmetry characteristic depends on *integrals* (rather than derivatives) of u. This approach proved to be particularly effective for matrix-valued PDEs such as the nonlinear self-dual Yang-Mills equation, of which new infinite-dimensional sets of "potential symmetries" were discovered [9,11,15].

7. An example: The chiral field equation

Let us consider the *chiral field equation*

$$F[g] \equiv (g^{-1}g_x)_x + (g^{-1}g_t)_t = 0 \tag{7.1}$$

where, in general, subscripts x and t denote total derivatives D_x and D_t , respectively, and where g is a $GL(n,C)$-valued function of x and t, i.e., a complex, non-singular $(n\times n)$ matrix function, differentiable for all x and t. Let $\delta g=\alpha Q[g]$ be an infinitesimal symmetry transformation for the PDE (7.1), with symmetry characteristic $Q[g]=\Delta_Q\, g$. It is convenient to put

$$Q[g] = g\,\Phi[g] \quad \Leftrightarrow \quad \Phi[g] = g^{-1}Q[g] \;.$$

The symmetry condition for (7.1) is

$$\Delta_Q F[g] = 0 \mod F[g] \;.$$

This condition will yield a linear PDE for Q or, equivalently, a linear PDE for Φ. By using the properties of the characteristic derivative we find the latter PDE to be

$$S(\Phi; g) \equiv D_x\left(\Phi_x + [g^{-1}g_x, \Phi]\right) + D_t\left(\Phi_t + [g^{-1}g_t, \Phi]\right) = 0 \mod F[g] \tag{7.2}$$

where the square brackets denote commutators of matrices.

A useful identity that will be needed in the sequel is the following:

$$(g^{-1}g_t)_x - (g^{-1}g_x)_t + [g^{-1}g_x\,,\, g^{-1}g_t] = 0 \tag{7.3}$$

Let us first consider symmetry transformations in the base space, i.e., coordinate transformations of x, t. An obvious symmetry is x-translation, $x'=x+\alpha$, given that the PDE (7.1) does not contain the independent variable x explicitly. For infinitesimal values of the parameter α, we write $\delta x=\alpha$. The symmetry characteristic is $Q[g]=g_x$, so that $\Phi[g]=g^{-1}g_x$. By substituting this expression for Φ into the symmetry condition (7.2) and by using the identity (7.3), we can verify that (7.2) is indeed satisfied:

$$S(\Phi; g) = D_x F[g] = 0 \mod F[g] \;.$$

Similarly, for t-translation, $t'=t+\alpha$ (infinitesimally, $\delta t=\alpha$) with $Q[g]=g_t$, we find

$$S(\Phi; g) = D_t F[g] = 0 \mod F[g] \;.$$

Another obvious symmetry of (7.1) is a scale change of both x and t: $x'=\lambda x$, $t'=\lambda t$. Setting $\lambda=1+\alpha$, where α is infinitesimal, we write $\delta x=\alpha x$, $\delta t=\alpha t$. The symmetry characteristic is $Q[g]=xg_x+tg_t$, so that $\Phi[g]=xg^{-1}g_x+tg^{-1}g_t$. Substituting for Φ into the symmetry condition (7.2) and using the identity (7.3) where necessary, we find that

$$S(\Phi;g) = \left(2+xD_x+tD_t\right)F[g] = 0 \mod F[g] .$$

Let us call $Q_1[g]=g_x$, $Q_2[g]=g_t$, $Q_3[g]=xg_x+tg_t$, and let us consider the corresponding characteristic derivative operators Δ_i defined by $\Delta_i\, g=Q_i$ (i=1,2,3). It is then straightforward to verify the following commutation relations:

$$[\Delta_1,\Delta_2]\,g = \Delta_1 Q_2 - \Delta_2 Q_1 = 0 \quad\Leftrightarrow\quad [\Delta_1,\Delta_2]=0\,;$$

$$[\Delta_1,\Delta_3]\,g = \Delta_1 Q_3 - \Delta_3 Q_1 = -g_x = -Q_1 = -\Delta_1 g \quad\Leftrightarrow\quad [\Delta_1,\Delta_3]=-\Delta_1\,;$$

$$[\Delta_2,\Delta_3]\,g = \Delta_2 Q_3 - \Delta_3 Q_2 = -g_t = -Q_2 = -\Delta_2 g \quad\Leftrightarrow\quad [\Delta_2,\Delta_3]=-\Delta_2\,.$$

Next, we consider the "internal" transformation (i.e., transformation in the fiber space) $g'=g\Lambda$, where Λ is a non-singular constant matrix. Then,

$$F[g'] = \Lambda^{-1}F[g]\,\Lambda = 0 \mod F[g] ,$$

which indicates that this transformation is a symmetry of (7.1). Setting $\Lambda=1+\alpha M$, where α is an infinitesimal parameter while M is a constant matrix, we write, in infinitesimal form, $\delta g=\alpha gM$. The symmetry characteristic is $Q[g]=gM$, so that $\Phi[g]=M$. Substituting for Φ into the symmetry condition (7.2), we find:

$$S(\Phi;g) = \left[F[g], M\right] = 0 \mod F[g] .$$

Given a matrix function $g(x,t)$ satisfying the PDE (7.1), consider the following system of PDEs for two functions $\Phi[g]$ and $\Phi'[g]$:

$$\begin{aligned} \Phi'_x &= \Phi_t + [g^{-1}g_t\,,\,\Phi] \\ -\Phi'_t &= \Phi_x + [g^{-1}g_x\,,\,\Phi] \end{aligned} \qquad (7.4)$$

The *integrability condition* (or consistency condition) $(\Phi'_x)_t=(\Phi'_t)_x$ of this system requires that Φ satisfy the symmetry condition (7.2); i.e., $S(\Phi;g)=0$. Conversely, by applying the integrability condition $(\Phi_t)_x=(\Phi_x)_t$ and by using the fact that g is a solution of $F[g]=0$, one finds that Φ' must also satisfy (7.2); i.e., $S(\Phi';g)=0$.

We conclude that, for any function $g(x,t)$ satisfying the PDE (7.1), the system (7.4) is an *auto-Bäcklund transformation* (BT) [18] relating solutions Φ and Φ' of the symmetry condition (7.2) of this PDE; that is, relating different symmetries of the chiral field equation. Thus, if a symmetry characteristic $Q=g\Phi$ of the PDE (7.1) is known, a new characteristic $Q'=g\Phi'$ may be found by integrating the BT (7.4); the converse is also true. Since the BT (7.4) produces new symmetries from old ones, it may be regarded as a *recursion operator* for the PDE (7.1) [8-11,15-17].

As an example, consider the internal-symmetry characteristic $Q[g]=gM$ (where M is a constant matrix) corresponding to $\Phi[g]=M$. By integrating the BT (7.4) for Φ', we get $\Phi'=[X, M]$ and thus $Q'=g[X, M]$, where X is the "potential" of the PDE (7.1), defined by the system of PDEs

$$X_x = g^{-1} g_t \quad , \quad - X_t = g^{-1} g_x \tag{7.5}$$

Note the *nonlocal* character of the BT-produced symmetry Q', due to the presence of the potential X. Indeed, as seen from (7.5), in order to find X one has to *integrate* the chiral field g with respect to the independent variables x and t. The above process can be continued indefinitely by repeated application of the recursion operator (7.4), leading to an infinite sequence of increasingly nonlocal symmetries.

Unfortunately, as the reader may check, no new information is furnished by the BT (7.4) in the case of coordinate symmetries (for example, by applying the BT for $Q=g_x$ we get the known symmetry $Q'=g_t$). A recursion operator of the form (7.4), however, does produce new nonlocal symmetries from coordinate symmetries in related problems with more than two independent variables, such as the self-dual Yang-Mills equation [8-11,15].

8. Generation of finite symmetry transformations

As we saw in Sec. 4, given a symmetry operator Δ_Q one may immediately define an infinitesimal symmetry of a PDE. Our starting point, however, was the idea of using a *finite* symmetry transformation to generate a one-parameter family of solutions of the PDE. We thus need to generalize the discussion of Sec. 4 by allowing the parameter α to assume finite values.

According to (2.7), the characteristic derivative Δ_Q with respect to the characteristic function $Q[u]$ satisfies the relation

$$\Delta_{Q[u]} u = Q[u] \tag{8.1}$$

By (8.1) and the properties of Δ_Q one may determine the action of Δ_Q on any function $F[u]$ of the form (2.1), thus construct a new function $\Delta_{Q[u]}F[u]$.

Obviously, a change of u will induce a corresponding change on any function $F[u]$. Given a function $u(x^k)$, a continuous change of u may be expressed as a one-parameter family of transformations

$$M : u(x^k) \to \bar{u}(x^k;\alpha) \quad \text{with} \quad \bar{u}(x^k;0) = u(x^k) \tag{8.2}$$

where $\alpha \geq 0$. We suppress the independent variables x^k, which are unaffected by the transformation M, and write, simply,

$$M : u \to \bar{u}(\alpha) \quad \text{with} \quad \bar{u}(0) = u \, .$$

Expanding $\bar{u}(\alpha)$ in powers of α, we have:

$$\bar{u}(\alpha) = u + \alpha Q[u] + \text{higher-order terms in } \alpha \tag{8.3}$$

where $Q[u]$ is given by

$$\frac{d}{d\alpha}\bar{u}(\alpha)\Big|_{\alpha=0} = Q[u] = \Delta_{Q[u]}u \tag{8.4}$$

Now, we assume that, for finite values of the parameter α,

$$\frac{d}{d\alpha}\bar{u}(\alpha) = Q\big[\bar{u}(\alpha)\big] = \Delta_{Q[\bar{u}(\alpha)]}\,\bar{u}(\alpha) \tag{8.5}$$

which is obviously consistent with (8.4). By the properties of the characteristic derivative it then follows that, for any function $F[u]$ of the form (2.1),

$$\frac{d}{d\alpha}F\big[\bar{u}(\alpha)\big] = \Delta_{Q[\bar{u}(\alpha)]}\,F\big[\bar{u}(\alpha)\big] \tag{8.6}$$

As an example, let $F[u]=u^2 \Rightarrow F[\bar{u}(\alpha)]=\bar{u}(\alpha)\bar{u}(\alpha)$. Then, by (8.5) and by using the Leibniz rule,

$$\begin{aligned}\frac{d}{d\alpha}F[\bar{u}(\alpha)] &= \frac{d\bar{u}(\alpha)}{d\alpha}\,\bar{u}(\alpha)+\bar{u}(\alpha)\frac{d\bar{u}(\alpha)}{d\alpha} \\ &= \left\{\Delta_{Q[\bar{u}(\alpha)]}\bar{u}(\alpha)\right\}\bar{u}(\alpha)+\bar{u}(\alpha)\Delta_{Q[\bar{u}(\alpha)]}\bar{u}(\alpha) \\ &= \Delta_{Q[\bar{u}(\alpha)]}\big(\bar{u}(\alpha)\bar{u}(\alpha)\big) = \Delta_{Q[\bar{u}(\alpha)]}\,F\big[\bar{u}(\alpha)\big]\end{aligned}$$

Equation (8.5), together with the initial condition $\bar{u}(0)=u$, is an initial-value problem that, upon integration, yields a one-parameter transformation of the form (8.2). We say that the operator Δ_Q is the *generator* of this transformation. As regards its action on functions, the operator Δ_Q is seen to be equivalent to the *Lie derivative* of differential geometry (see, e.g., Chap. 5 of [19]). And, the latter derivative plays a key role in the differential-geometric methods for studying invariance properties of PDEs [2-4]. We now revisit the symmetry problem for PDEs in the context of our present, more abstract algebraic formalism.

The transformation M of Eq. (8.2) is a *symmetry transformation* for the PDE $F[u]=0$ if it leaves this PDE invariant, in the sense that $F[\bar{u}(\alpha)]=0$ *if* $F[u]=0$. We write:

$$F[\bar{u}(\alpha)]=0 \quad \text{mod} \quad F[u] \tag{8.7}$$

So, if $\bar{u}(x^k;0)=u(x^k)$ is a solution of $F[u]=0$, then so is $\bar{u}(x^k;\alpha)$ *for all values of the parameter* $\alpha>0$. This means that $F[\bar{u}(\alpha)]$, viewed as a function of α, retains a constant (zero) value under continuous changes of α. In mathematical terms,

$$\frac{d}{d\alpha}F\big[\bar{u}(\alpha)\big]=0 \quad \text{mod} \quad F\big[\bar{u}(\alpha)\big]$$

or, in view of (8.6),

$$\Delta_{Q[\bar{u}(\alpha)]}F\big[\bar{u}(\alpha)\big]=0 \quad \text{mod} \quad F\big[\bar{u}(\alpha)\big]\ .$$

Since this must be valid for any value of α, the above relation will still be true if we replace α by a new parameter $\beta=\alpha+c$, where c is any constant such that $\beta\geq 0$. In particular, by choosing $c=-\alpha \Rightarrow \beta=0$, we rewrite the above equation in the simpler form

$$\Delta_{Q[u]} F[u]=0 \mod F[u] \tag{8.8}$$

which is the condition for invariance of the PDE $F[u]=0$. As we have seen, this condition yields a linear PDE for the symmetry characteristic $Q[u]$, of the form

$$S(Q;u) \equiv \Delta_{Q[u]} F[u]=0 \mod F[u] \tag{8.9}$$

where the expression $S(Q;u)$ is linear in Q and all total derivatives of Q. In particular, for scalar-valued u (thus scalar $Q[u]$ also) the operator Δ_Q has the form (2.13) and the symmetry condition (8.9) takes on the form

$$\frac{\partial F}{\partial u} Q[u] + \frac{\partial F}{\partial u_i} D_i Q[u] + \frac{\partial F}{\partial u_{ij}} D_i D_j Q[u] + \cdots = 0 \mod F[u] \tag{8.10}$$

An important class of symmetries is *local* (point) symmetries. As discussed in [1], the symmetry characteristic $Q[u]$ of a local symmetry cannot depend on second or higher-order derivatives of u with respect to the x^k, while the dependence of Q on the first-order derivatives u_k is also subject to certain restrictions. Once a local symmetry characteristic $Q[u]$ is found by solving (8.9) or (8.10), a one-parameter family of symmetry transformations of the PDE $F[u]=0$, of the form

$$M: u(x^k) \to \bar{u}(x^k;\alpha) \;\; ; \quad \bar{u}(x^k;0)=u(x^k) \tag{8.11}$$

is obtained by solving the initial-value problem [cf. Eq. (8.5)]

$$\begin{aligned} &\frac{d}{d\alpha}\bar{u}(x^k;\alpha) = Q[\bar{u}] \\ &\bar{u}(x^k;0)=u(x^k) \end{aligned} \tag{8.12}$$

9. Example: The two-dimensional Laplace equation

As an example for illustrating the process of finding one-parameter symmetry transformations of a PDE, we choose the two-dimensional *Laplace equation* for a scalar function $u(x,t)$:

$$F[u] \equiv u_{xx} + u_{tt} = 0 \tag{9.1}$$

Here, $(x^1, x^2) \equiv (x,t)$. The symmetry condition (8.9) or (8.10) yields the linear PDE

$$S(Q;u) \equiv Q_{xx} + Q_{tt} = 0 \mod F[u] \tag{9.2}$$

where subscripts indicate total differentiations. Each symmetry of the PDE (9.1) corresponds to a solution $Q[u]$ of (9.2) and leads to a one-parameter family of symmetry transformations (8.11) by solving the initial-value problem (8.12). Let us see some examples:

1. $Q[u]=1$ is a solution of (9.2), hence a symmetry characteristic of (9.1). The initial-value problem (8.12) is written

$$\frac{d}{d\alpha}\bar{u}(\alpha)=1 \ ; \quad \bar{u}(0)=u$$

which is easily integrated to give $\bar{u}(x,t;\alpha)=u(x,t)+\alpha$. Thus, if $u(x,t)$ is a solution of (9.1), then so is $u(x,t)+\alpha$. This symmetry reflects the fact that u enters the PDE (9.1) only through its derivatives (i.e., $F[u]$ does not contain u itself).

2. For the symmetry characteristic $Q[u]=u$ we have

$$\frac{d}{d\alpha}\bar{u}(\alpha)=\bar{u} \ ; \quad \bar{u}(0)=u$$

with solution $\bar{u}(x,t;\alpha)=e^{\alpha}u(x,t)$. Thus, if $u(x,t)$ is a solution of (9.1), then so is $\lambda u(x,t)$ for any constant λ. This symmetry reflects the fact that the PDE (9.1) is homogeneous linear in u.

3. $Q[u]=u_x$ is another symmetry characteristic; indeed, note that $S(Q;u)=D_xF[u]=0$ when $F[u]=0$. The initial-value problem is written

$$\frac{d}{d\alpha}\bar{u}(x,t;\alpha)=\bar{u}_x \ ; \quad \bar{u}(x,t;0)=u(x,t) \, .$$

A way to solve this is to consider the parameter α as a variable of equal footing with x and t. The above ordinary differential equation then becomes a homogeneous linear first-order PDE that can be integrated by standard methods (see, e.g., Chap. 4 of [19]):

$$\bar{u}_x-\bar{u}_\alpha=0 \tag{9.3}$$

We form the characteristic system

$$\frac{dx}{1}=\frac{dt}{0}=\frac{d\alpha}{-1}=\frac{d\bar{u}}{0}$$

and seek 3 first integrals of this system. These are $\bar{u}=C_1$, $t=C_2$, $x+\alpha=C_3$. The general solution of (9.3) is then $\Phi(C_1, C_2, C_3)=0$, where the function Φ is arbitrary. That is,

$$\Phi(\bar{u},\, t,\, x+\alpha)=0 \ \Rightarrow \ \bar{u}(x,t;\alpha)=\omega(x+\alpha,\, t) \, .$$

By the initial condition $\bar{u}(x,t;0)=u(x,t)$ we have that $\omega(x,t)=u(x,t)$, and hence $\omega(x+\alpha, t)=u(x+\alpha, t)$. Thus, finally, $\bar{u}(x,t;\alpha)=u(x+\alpha,\, t)$.

In a similar manner, from the symmetry characteristic $Q[u]=u_t$ we get the transformation $\bar{u}(x,t;\alpha)=u(x,\, t+\alpha)$. The two symmetries found above reflect the fact that the PDE (9.1) does not contain the independent variables x and t explicitly.

4. For the symmetry characteristic $Q[u]=xu_x+tu_t$ we have

$$\frac{d}{d\alpha}\bar{u}(x,t;\alpha)=x\bar{u}_x+t\bar{u}_t \;\; ; \;\; \bar{u}(x,t;0)=u(x,t).$$

Working as in the previous example, we form the first-order PDE

$$x\bar{u}_x+t\bar{u}_t-\bar{u}_\alpha=0 \qquad (9.4)$$

with characteristic system

$$\frac{dx}{x}=\frac{dt}{t}=\frac{d\alpha}{-1}=\frac{d\bar{u}}{0} \;.$$

Three first integrals are $\bar{u}=C_1$, $\ln x+\alpha=C_2$, $\ln t+\alpha=C_3$. The general solution of the PDE (9.4) is $\Phi(C_1,C_2,C_3)=0$, with Φ arbitrary. That is,

$$\Phi(\bar{u},\, \ln x+\alpha,\, \ln t+\alpha)=0 \;\;\Rightarrow$$

$$\bar{u}(x,t;\alpha)=\omega(\ln x+\alpha,\, \ln t+\alpha)=\omega\left[\ln\left(e^\alpha x\right),\, \ln\left(e^\alpha t\right)\right]$$

By the initial condition $\bar{u}(x,t;0)=u(x,t)$ we have that $\omega(\ln x, \ln t)=u(x,t)$, and hence $\omega[\ln(e^\alpha x), \ln(e^\alpha t)]=u(e^\alpha x, e^\alpha t)$. Thus, finally, $\bar{u}(x,t;\alpha)=u(e^\alpha x,\, e^\alpha t)$. This transformation expresses the invariance of the PDE (9.1) under a scale change $x\to\lambda x$, $t\to\lambda t$ of x and t.

5. $Q[u]=tu_x-xu_t$ is a symmetry characteristic since $S(Q;u)=(tD_x-xD_t)F[u]=0$ when $F[u]=0$. We write

$$\frac{d}{d\alpha}\bar{u}(x,t;\alpha)=t\bar{u}_x-x\bar{u}_t \;\; ; \;\; \bar{u}(x,t;0)=u(x,t)$$

and form the PDE

$$t\bar{u}_x-x\bar{u}_t-\bar{u}_\alpha=0 \qquad (9.5)$$

with characteristic system

$$\frac{dx}{t}=\frac{dt}{-x}=\frac{d\alpha}{-1}=\frac{d\bar{u}}{0} \qquad (9.6)$$

One first integral is $\bar{u}=C_1$. Another one is found from $dx/t=-dt/x \Rightarrow d(x^2+t^2)=0 \Rightarrow x^2+t^2=C_2$. A third integral is

$$\alpha+\arctan(x/t)=C_3\,.$$

Let us prove this. Setting $x/t=z$, we have:

$$d(\alpha+\arctan z)=d\alpha+\frac{dz}{1+z^2}=d\alpha+\frac{tdx-xdt}{x^2+t^2}\,.$$

But, by the system (9.6), $dx=-td\alpha$ and $dt=xd\alpha$, so that

$$d\left[\alpha+\arctan(x/t)\right]=d\alpha-\frac{(x^2+t^2)\,d\alpha}{x^2+t^2}=0, \quad \text{q.e.d.}$$

The general solution of the PDE (9.5) is $\Phi(C_1,C_2,C_3)=0$, with Φ arbitrary. That is,

$$\Phi\left[\bar{u},\ x^2+t^2,\ \alpha+\arctan(x/t)\right]=0\ \Rightarrow$$

$$\bar{u}(x,t;\alpha)=\omega\left[x^2+t^2,\ \alpha+\arctan(x/t)\right] \tag{9.7}$$

By the initial condition $\bar{u}(x,t;0)=u(x,t)$ we have that $\omega[x^2+t^2,\ \arctan(x/t)]=u(x,t)$. Putting $x\cos\alpha+t\sin\alpha$ and $t\cos\alpha-x\sin\alpha$ in place of x and t, respectively, we find:

$$u(x\cos\alpha+t\sin\alpha,\ t\cos\alpha-x\sin\alpha)=\omega\left(x^2+t^2,\ \arctan\frac{x\cos\alpha+t\sin\alpha}{t\cos\alpha-x\sin\alpha}\right) \tag{9.8}$$

On the other hand, we can show that

$$\tan\left[\alpha+\arctan(x/t)\right]=\frac{x\cos\alpha+t\sin\alpha}{t\cos\alpha-x\sin\alpha}\ \Rightarrow$$

$$\alpha+\arctan(x/t)=\arctan\frac{x\cos\alpha+t\sin\alpha}{t\cos\alpha-x\sin\alpha} \tag{9.9}$$

From (9.8) and (9.9) we have that

$$u(x\cos\alpha+t\sin\alpha,\ t\cos\alpha-x\sin\alpha)=\omega\,[x^2+t^2,\ \alpha+\arctan(x/t)]\ .$$

Thus (9.7) assumes the final form

$$\bar{u}(x,t;\alpha)=u(x\cos\alpha+t\sin\alpha,\ -x\sin\alpha+t\cos\alpha) \tag{9.10}$$

The transformation (9.10) admits a certain geometrical interpretation that becomes evident if we define the new variables $x'=x\cos\alpha+t\sin\alpha$ and $t'=-x\sin\alpha+t\cos\alpha$. In matrix form,

$$\begin{bmatrix} x' \\ t' \end{bmatrix}=\begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}\begin{bmatrix} x \\ t \end{bmatrix}\ .$$

This relation describes a rotation of the vector (x,t) on the xt-plane by an angle α. The PDE $F[u]=0$ is thus invariant under such a rotation on the plane of the independent variables.

10. Concluding remarks

The algebraic approach to the symmetry problem of PDEs, presented in this article, is, in a sense, an extension to matrix-valued problems of the ideas contained in [1], in much the same way as [4] and [5] constitute a generalization of the Harrison-Estabrook geometrical approach [2] to matrix-valued (as well as vector-valued and Lie-algebra-valued) PDEs.

The symmetry transformations we have considered involve only the change of the dependent variable u of the PDE, while leaving the independent variables x^k unchanged. Indeed, as Olver [1] has shown, *every* symmetry of a PDE may be expressed as a transformation of the dependent variable alone, i.e., as a transformation in the fiber space. The symmetry-generating characteristic derivative Δ_Q corresponds to Olver's *evolutionary vector field* with characteristic $Q[u]$.

In *local* (point) symmetries of the PDE $F[u]=0$ the symmetry characteristics $Q[u]$ contain at most first-order derivatives of u with respect to the x^k (a number of such symmetries were considered in the last section in connection with the Laplace equation). The case of *generalized* (non-local) symmetries is more complex; a formal solution to the problem of obtaining one-parameter families of generalized symmetry transformations of PDEs is given in Sec. 5.1 of [1].

Admittedly, the abstract algebraic formalism we have presented does not exhibit significant advantages over the standard geometrical methods (in particular, those described in [4]) with regard to *finding* symmetries of PDEs. However, by employing the concept of the characteristic derivative one is able to bypass the difficulty of having to represent symmetry-generating operators as vector fields in the form of differential operators when matrix-valued variables are involved, which situation can only be handled by making certain *ad hoc* assumptions regarding the action of such "unnatural" operators. The algebraic view offers a more rigorous framework for identifying symmetry operators and finding infinitesimal symmetry transformations of matrix-valued PDEs, as well as for studying the Lie-algebraic structure of the set of symmetry generators (see, e.g., [9]).

References

1. P. J. Olver, *Applications of Lie Groups to Differential Equations* (Springer-Verlag, 1993).
2. B. K. Harrison, F. B. Estabrook, *Geometric approach to invariance groups and solutions of partial differential equations*, J. Math. Phys. 12 (1971) 653.
3. D. G. B. Edelen, *Applied Exterior Calculus* (Wiley, 1985).
4. C. J. Papachristou, B. K. Harrison, *Isogroups of differential ideals of vector-valued differential forms: Application to partial differential equations*, Acta Appl. Math. 11 (1988) 155.
5. C. J. Papachristou, B. K. Harrison, *Some aspects of the isogroup of the self-dual Yang-Mills system*, J. Math. Phys. 28 (1987) 1261.
6. C. J. Papachristou, B. K. Harrison, *Nonlocal symmetries and Bäcklund transformations for the self-dual Yang-Mills system*, J. Math. Phys. 29 (1988) 238.
7. C. J. Papachristou, *Symmetry and integrability of classical field equations*, arXiv:0803.3688 (https://arxiv.org/abs/0803.3688).
8. C. J. Papachristou, *Symmetry, conserved charges, and Lax representations of nonlinear field equations: A unified approach*, Electron. J. Theor. Phys. 7, No. 23 (2010) 1.
9. C. J. Papachristou, B. K. Harrison, *Bäcklund-transformation-related recursion operators: Application to the self-dual Yang-Mills equation*, J. Nonlin. Math. Phys., Vol. 17, No. 1 (2010) 35.
10. C. J. Papachristou, *Symmetry and integrability of a reduced, 3-dimensional self-dual gauge field model*, Electron. J. Theor. Phys. 9, No. 26 (2012) 119.
11. C. J. Papachristou, *Lax pair, hidden symmetries, and infinite sequences of conserved currents for self-dual Yang-Mills fields*, J. Phys. A 24 (1991) L 1051.
12. M. Marvan, A. Sergyeyev, *Recursion operators for dispersionless integrable systems in any dimension*, Inverse Problems 28 (2012) 025011.
13. S. Igonin, M. Marvan, *On construction of symmetries and recursion operators from zero-curvature representations and the Darboux–Egoroff system*, J. Geom. Phys. 85 (2014) 106.

14. A. Sergyeyev, *A simple construction of recursion operators for multidimensional dispersionless integrable systems*, J. Math. Anal. Appl. 454 (2017) 468.

15. C. J. Papachristou, *Potential symmetries for self-dual gauge fields*, Phys. Lett. A 145 (1990) 250.

16. G. A. Guthrie, *Recursion operators and non-local symmetries*, Proc. R. Soc. Lond. A 446 (1994) 107.

17. M. Marvan, *Another look on recursion operators*, in *Differential Geometry and Applications* (Proc. Conf., Brno, 1995) 393.

18. C. J. Papachristou, A. N. Magoulas, *Bäcklund transformations: Some old and new perspectives*, Nausivios Chora Vol. 6 (2016) pp. C3-C17 (https://nausivios.snd.edu.gr/docs/2016C.pdf).

19. C. J. Papachristou, *Aspects of Integrability of Differential Systems and Fields: A Mathematical Primer for Physicists* (Springer, 2019).[1]

[1] Manuscript: https://arxiv.org/abs/1511.01788

Bäcklund-Transformation-Related Recursion Operators: Application to the Self-Dual Yang-Mills Equation

C. J. Papachristou * , B. Kent Harrison **

*Department of Physical Sciences, Naval Academy of Greece, Piraeus 18539, Greece
E-mail: papachristou@snd.edu.gr

**Department of Physics and Astronomy, Brigham Young University, Provo, UT 84602, USA
E-mail: bkentharrison@comcast.net , bkh@byu.edu

Abstract. By using the self-dual Yang-Mills (SDYM) equation as an example, we study a method for relating symmetries and recursion operators of two partial differential equations connected to each other by a non-auto-Bäcklund transformation. We prove the Lie-algebra isomorphism between the symmetries of the SDYM equation and those of the potential SDYM (PSDYM) equation, and we describe the construction of the recursion operators for these two systems. Using certain known aspects of the PSDYM symmetry algebra, we draw conclusions regarding the Lie algebraic structure of the "potential symmetries" of the SDYM equation.

Keywords: Bäcklund transformations, recursion operators, self-dual Yang-Mills equation

Published in 2010 in the *Journal of Nonlinear Mathematical Physics*

I. Introduction

Recursion operators are powerful tools for the study of symmetries of partial differential equations (PDEs). Roughly speaking, a recursion operator is a linear operator which produces a new symmetry characteristic of a PDE whenever it acts on an "old" characteristic (see Appendix). The concept was first introduced by Olver [1, 2] and subsequently used by many authors (see, e.g., [2, 3] and the references therein). An alternative view, based on the concept of a Bäcklund transformation (BT), was developed in a series of papers by the present authors [4-6] who studied the symmetry problem for the self-dual Yang-Mills equation (SDYM). The general idea is that a recursion operator can be viewed as an auto-BT for the "linearization equation" (or symmetry condition) of a (generally nonlinear) PDE. This idea was later further developed and put into formal mathematical perspective by Marvan [7].

It has been known for some time (see, e.g., Section 7.4 of [3] and the references therein) that, when two nonlinear PDEs are connected by a non-auto-BT, symmetries of either PDE may yield symmetries of the other. This can be achieved by using the original BT to construct another non-auto-BT which relates solutions of the linearization equations of the two PDEs. In the particular case of the SDYM equation, the original BT associates this PDE with the "potential SDYM equation" (PSDYM). The symmetries of the latter PDE can then be used to construct the "potential symmetries" of SDYM [5, 8]. We now attempt to go one step further: Can we find a BT which relates *recursion operators* of two PDEs? Given that, as said above, a recursion operator is itself an auto-BT, what we are after is a BT connecting two auto-BTs, each of which produces solutions of a respective linear PDE (symmetry condition). Thus, we are looking for "a transformation of transformations" rather than a transformation of functions.

Our "laboratory" model will again be SDYM, for good reasons. First, it possesses a rich symmetry structure; second, this PDE has been shown to constitute a sort of prototype equation from which several other integrable PDEs are derived by reduction (see, e.g., [9, 10]). By employing a non-auto-BT that connects SDYM with PSDYM, we will show how symmetries and recursion operators of either system can be associated with symmetries and recursion operators, respectively, of the other system. Moreover, we will prove that the symmetry Lie algebras of these two PDEs are isomorphic to each other. This conclusion is more than of academic importance, since it allows us to investigate the symmetry structure of the SDYM problem by studying the relatively easier PSDYM problem. As an example, we will recover the known infinite-dimensional symmetry algebras of SDYM [11-13] from the symmetry structure of PSDYM [8] and show how these algebras are related to potential symmetries.

II. The Symmetry Problem for the SDYM-PSDYM System

We write the SDYM equation in the form

$$F[J] \equiv D_{\bar{y}}(J^{-1}J_y) + D_{\bar{z}}(J^{-1}J_z) = 0 \tag{1}$$

We denote by $x^\mu \equiv y, z, \bar{y}, \bar{z}$ $(\mu = 1, \cdots, 4)$ the independent variables, and by D_y, D_z, etc., the total derivatives with respect to these variables. We will also use the notation $D_y F \equiv F_y$, etc., for any function *F*. We assume that *J* is *SL*(*N*,*C*)-valued (i.e., $\det J = 1$).

We consider the non-auto-BT

$$J^{-1}J_y = X_{\bar{z}}\ , \quad J^{-1}J_z = -X_{\bar{y}} \tag{2}$$

The integrability condition $(X_{\bar{y}})_{\bar{z}} = (X_{\bar{z}})_{\bar{y}}$ yields the SDYM equation (1). The integrability condition $(J_y)_z = (J_z)_y$, which is equivalent to

$$D_y(J^{-1}J_z) - D_z(J^{-1}J_y) + [J^{-1}J_y\, ,\, J^{-1}J_z] = 0\, ,$$

yields a nonlinear PDE for the "potential" *X* of (1), called the "potential SDYM equation" or PSDYM:

$$G[X] \equiv X_{y\bar{y}} + X_{z\bar{z}} - [X_{\bar{y}}\, ,\, X_{\bar{z}}] = 0 \tag{3}$$

Noting that, according to (2), $(trX)_{\bar{z}} = [tr(\ln J)]_y = [\ln(\det J)]_y$, etc., we see that the condition $\det J = 1$ can be satisfied by requiring that $trX = 0$ [this requirement is compatible with the PSDYM equation (3)]. Hence, *SL*(*N*,*C*) SDYM solutions correspond to *sl*(*N*,*C*) PSDYM solutions.

Let $\delta J=\alpha Q$ and $\delta X=\alpha \Phi$ be an infinitesimal symmetry of system (2) (α is an infinitesimal parameter). This means that $(J+\delta J,\ X+\delta X)$ is a solution to the system when $(J,\ X)$ is a solution. This suggests that the integrability conditions $F[J+\delta J]=0$ and $G[X+\delta X]=0$ are satisfied when the integrability conditions $F[J]=0$ and $G[X]=0$ are satisfied; that is, $J+\delta J$ and $X+\delta X$ are solutions of (1) and (3), respectively. The functions Q and Φ are *symmetry characteristics* for the above PDEs. Geometrically, the symmetries of system (2) are realized as transformations in the jet-like space of the variables $\{x^\mu, J, X\}$ and the various derivatives of J and X with respect to the x^μ. These transformations are generated by vector fields which, without loss of generality, may be considered "vertical", i.e., with vanishing projections on the base space of the x^μ [2]. We formally represent these vectors by differential operators of the form

$$V = Q\frac{\partial}{\partial J} + \Phi\frac{\partial}{\partial X} \quad (+\, prolongation\ terms) \tag{4}$$

Consider a function $M(J, X)$. Denote by $\Delta M(J, X)$ the Fréchet derivative [2] of M with respect to V (which in this context is locally the same as the Lie derivative). The infinitesimal variation of M in the "direction" of V is then $\delta M=\alpha\Delta M$. The linear operator Δ is a derivation on the algebra of all $gl(N,C)$-valued functions. The Leibniz rule is written

$$\Delta(M\,N) = (\Delta M)\,N + M\,\Delta N \tag{5}$$

In particular, for the Lie algebra of $sl(N,C)$-valued functions, the Leibniz rule may also be written as

$$\Delta[M,N] = [\Delta M,N] + [M,\Delta N] \tag{6}$$

By definition, the Fréchet derivatives of the fundamental variables J and X are given by

$$\Delta J = Q\ , \quad \Delta X = \Phi \tag{7}$$

We also note that the Fréchet derivative with respect to a *vertical* vector field commutes with all total derivative operators [2]. Finally, for an invertible matrix M,

$$\Delta(M^{-1}) = -M^{-1}(\Delta M)\,M^{-1} \tag{8}$$

(For a discussion of the general symmetry problem for matrix-valued PDEs, see [14].)

We introduce the covariant derivative operators

$$\begin{aligned} \hat{A}_y &\equiv D_y + [J^{-1}J_y,\ \ ] = D_y + [X_{\bar z},\ \ ] \\ \hat{A}_z &\equiv D_z + [J^{-1}J_z,\ \ ] = D_z - [X_{\bar y},\ \ ] \end{aligned} \tag{9}$$

where the BT (2) has been taken into account. By using (3) and the Jacobi identity, the zero-curvature condition $[\hat{A}_y, \hat{A}_z]=0$ is shown to be satisfied, as expected in view of the fact that the "connections" $J^{-1}J_y$ and $J^{-1}J_z$ are pure gauges. Moreover, the linear operators of (9) are derivations on the Lie algebra of *sl*(*N*,*C*)-valued functions, satisfying a Leibniz rule of the form (6):

$$\begin{aligned}\hat{A}_y\,[M,N] &= [\hat{A}_y M,N]+[M,\hat{A}_y N] \\ \hat{A}_z\,[M,N] &= [\hat{A}_z M,N]+[M,\hat{A}_z N]\end{aligned} \qquad (10)$$

If Eqs. (1)-(3) are satisfied, then so must be their Fréchet derivatives with respect to the symmetry vector field *V* of (4). We now derive the symmetry condition for each of the above three systems. For SDYM (1), the symmetry condition is $\Delta F[J]=0$, or

$$D_{\bar{y}}\,\Delta(J^{-1}J_y) + D_{\bar{z}}\,\Delta(J^{-1}J_z) = 0 \qquad (11)$$

(since the Fréchet derivative Δ commutes with total derivatives). By using (5), (7), (8) and (9), it can be shown that

$$\Delta(J^{-1}J_y) = \hat{A}_y\,(J^{-1}Q)\,, \quad \Delta(J^{-1}J_z) = \hat{A}_z\,(J^{-1}Q) \qquad (12)$$

The SDYM symmetry condition (11) then becomes

$$(D_{\bar{y}}\hat{A}_y + D_{\bar{z}}\hat{A}_z)(J^{-1}Q) = 0 \qquad (13)$$

The symmetry condition for PSDYM (3) is $\Delta G[X]=0$, or, by using (6), (7) and (9),

$$\hat{A}_y\Phi_{\bar{y}} + \hat{A}_z\Phi_{\bar{z}} \equiv (\hat{A}_y D_{\bar{y}} + \hat{A}_z D_{\bar{z}})\Phi = 0 \qquad (14)$$

We note the operator identity

$$\hat{A}_y D_{\bar{y}} + \hat{A}_z D_{\bar{z}} = D_{\bar{y}}\hat{A}_y + D_{\bar{z}}\hat{A}_z \qquad (15)$$

which is easily verified by letting the right-hand side act on an arbitrary function *M*. Then, (14) is written in the alternate form,

$$(D_{\bar{y}}\hat{A}_y + D_{\bar{z}}\hat{A}_z)\Phi = 0 \qquad (16)$$

Comparing (13) and (16), we observe that the symmetry characteristic Φ of PSDYM, and the function $J^{-1}Q$, where Q is an SDYM symmetry characteristic, satisfy the same symmetry condition. We thus conclude the following (see also [5]):

- If Q is an SDYM characteristic, then $\Phi = J^{-1}Q$ is a PSDYM characteristic.

Conversely,

- if Φ is a PSDYM characteristic, then $Q=J\Phi$ is an SDYM characteristic.

Finally, the Fréchet derivative with respect to V also leaves the system of PDEs (2) invariant: $\Delta(J^{-1}J_y)=(\Delta X)_{\bar z}$, $\Delta(J^{-1}J_z)=-(\Delta X)_{\bar y}$. With the aid of (12) and (7) we are thus led to a pair of PDEs,

$$\hat{A}_y(J^{-1}Q)=\Phi_{\bar z}\ , \quad \hat{A}_z(J^{-1}Q)=-\Phi_{\bar y} \tag{17}$$

Equation (17) is a BT connecting the symmetry characteristic Φ of PSDYM with the symmetry characteristic Q of SDYM. Indeed, the integrability condition $(\Phi_{\bar z})_{\bar y}=(\Phi_{\bar y})_{\bar z}$ yields the symmetry condition (13) for SDYM. So, when Q is an SDYM symmetry characteristic, the BT (17) is integrable for Φ. Conversely, the integrability condition $[\hat{A}_z\, ,\hat{A}_y](J^{-1}Q)=0$, valid in view of the zero-curvature condition, yields the PSDYM symmetry condition (14) for Φ and guarantees integrability for Q.

We note that, for a given Q, the solution of the BT (17) for Φ is not unique, and vice versa. To achieve uniqueness we thus need to make some additional assumptions: (*a*) If Φ is a solution for a given Q, then so is $\Phi+M(y,z)$, where M is an arbitrary matrix function. We make the agreement that any arbitrary additive term of the form $M(y,z)$ will be ignored when it appears in the solution for Φ. (*b*) If Q is a solution for a given Φ, then so is $Q+\varepsilon(\bar y,\bar z)J$, where $\varepsilon(\bar y,\bar z)$ is an arbitrary matrix function. We agree that any arbitrary additive term of the form $\varepsilon(\bar y,\bar z)J$ will be ignored when it appears in the solution for Q.

With the above conventions, the BT (17) establishes a 1-1 correspondence between the symmetries of SDYM and those of PSDYM. In particular, the SDYM characteristic Q=0 corresponds to the PSDYM characteristic Φ=0. It will be shown below that this correspondence between the two symmetry sets is a Lie algebra isomorphism.

III. Recursion Operators and Lie-Algebra Isomorphism

Since the two PDEs in (17) are consistent with each other and solvable for Φ when Q is an SDYM symmetry characteristic, we may use, say, the first equation to formally express Φ in terms of Q:

$$\Phi=D_{\bar z}^{-1}\hat{A}_y(J^{-1}Q)\equiv\hat{R}(J^{-1}Q) \tag{18}$$

where we have introduced the linear operator

$$\hat{R}=D_{\bar z}^{-1}\hat{A}_y \tag{19}$$

Proposition 1: The operator (19) is a recursion operator for PSDYM.

Proof : Let Φ be a symmetry characteristic for PSDYM. Then, Φ satisfies the symmetry conditions (14) or (16). We will show that $\Phi'\equiv\hat{R}\Phi$ also is a symmetry characteristic. Indeed,

$$\begin{aligned}(\hat{A}_y D_{\bar{y}} + \hat{A}_z D_{\bar{z}})\Phi' &\equiv (\hat{A}_y D_{\bar{y}} + \hat{A}_z D_{\bar{z}})\hat{R}\Phi \\ &= \hat{A}_y D_{\bar{z}}^{-1} D_{\bar{y}} \hat{A}_y \Phi + \hat{A}_z \hat{A}_y \Phi \\ &= \hat{A}_y D_{\bar{z}}^{-1}(D_{\bar{y}} \hat{A}_y + D_{\bar{z}} \hat{A}_z)\Phi + [\hat{A}_z , \hat{A}_y]\Phi = 0 ,\end{aligned}$$

in view of (16) and the zero-curvature condition $[\hat{A}_y , \hat{A}_z]=0$. ■

For $sl(N,C)$ PSDYM solutions, the symmetry characteristic Φ must be traceless. Then, so is the characteristic $\Phi' = \hat{R}\Phi$. That is, the recursion operator (19) preserves the $sl(N,C)$ character of PSDYM.

Is there a systematic process by which one could *derive* the recursion operator (19)? To this end, we seek an auto-BT relating solutions of the PSDYM symmetry condition (14). As shown in [5], such a BT is

$$\hat{A}_y \Phi = \Phi'_{\bar{z}} \;, \quad \hat{A}_z \Phi = -\Phi'_{\bar{y}} \tag{19a}$$

The first of these equations can then be re-expressed as $\Phi' = \hat{R}\Phi$, with $\hat{R}$ given by (19).

Consider now a symmetry characteristic Q of SDYM, i.e., a solution of the symmetry condition (13). Also, consider the transformation

$$Q' = J\hat{R}\,(J^{-1}Q) \equiv \hat{T}Q \tag{20}$$

where we have introduced the linear operator

$$\hat{T} = J\hat{R}J^{-1} \tag{21}$$

Proposition 2: The operator (21) is a recursion operator for SDYM.

Proof : By assumption, Q is an SDYM symmetry characteristic. Then, as shown above, $\Phi = J^{-1}Q$ is a PSDYM characteristic. Since $\hat{R}$ is a PSDYM recursion operator, $\Phi' \equiv \hat{R}\Phi = \hat{R}(J^{-1}Q)$ also is a PSDYM characteristic. Then, finally, $Q' = J\Phi'$, given by (20), is an SDYM characteristic. ■

For $SL(N,C)$ SDYM solutions, the symmetry characteristic Q must satisfy the condition $tr(J^{-1}Q)=0$. As can be seen, this condition is preserved by the recursion operator (21). [Note, in this connection, that the BT (17) or (18) properly associates $SL(N,C)$ SDYM characteristics Q with $sl(N,C)$ PSDYM characteristics Φ.]

The recursion operator (21) also can be derived from an auto-BT for the SDYM symmetry condition (13). This BT was constructed in [6] by using a properly chosen Lax pair for SDYM (we refer the reader to this paper for details). We may thus conclude that recursion operators such as (19) or (21) in effect represent auto-BTs for symmetry conditions of respective nonlinear PDEs (see also [7]).

Lemma: The Fréchet derivative $\varDelta$ with respect to the vector V of (4), and the recursion operator $\hat{R}$ of (19), satisfy the commutation relation

$$[\varDelta, \hat{R}] = D_{\bar{z}}^{-1}[\varPhi_{\bar{z}} \,,\ ] \tag{22}$$

where $\varPhi = \varDelta X$, according to (7).

Proof: Introducing an auxiliary function F, and using the derivation property (6) of $\varDelta$ and the commutativity of $\varDelta$ with all total derivatives (as well as all powers of such derivatives), we have:

$$\begin{aligned}\varDelta \hat{R} F &= \varDelta D_{\bar{z}}^{-1} \hat{A}_y F = D_{\bar{z}}^{-1} \varDelta (D_y F + [X_{\bar{z}}, F]) \\ &= D_{\bar{z}}^{-1} (D_y \varDelta F + [(\varDelta X)_{\bar{z}}, F] + [X_{\bar{z}}, \varDelta F]) \\ &= D_{\bar{z}}^{-1} (\hat{A}_y \varDelta F + [\varPhi_{\bar{z}}, F]) = \hat{R} \varDelta F + D_{\bar{z}}^{-1} [\varPhi_{\bar{z}}, F] \,,\end{aligned}$$

from which there follows (22). ■

Proposition 3: The BT (17), or equivalently, its solution (18), establishes an isomorphism between the symmetry Lie algebras of SDYM and PSDYM.

Proof: Let V be a vector field of the form (4), generating a symmetry of the BT (2). As explained previously, since this BT is invariant under V, the same will be true with regard to its integrability conditions. Hence, V also represents a symmetry of the SDYM-PSDYM system of equations (1) and (3). The SDYM and PSDYM characteristics are $Q = \varDelta J$ and $\varPhi = \varDelta X$, respectively, where $\varDelta$ denotes the Fréchet derivative with respect to V. Consider the linear map I defined by (18):

$$I: \ \varPhi = I\{Q\} = \hat{R} J^{-1} Q \tag{23}$$

or

$$I: \ \varDelta X = I\{\varDelta J\} = \hat{R} J^{-1} \varDelta J \tag{24}$$

Consider also a pair of symmetries of system (2), indexed by i and j. These are generated by vector fields $V^{(r)}$, where $r = i, j$. The Fréchet derivatives with respect to the $V^{(r)}$ will be denoted $\varDelta^{(r)}$. The SDYM and PSDYM symmetry characteristics are $Q^{(r)} = \varDelta^{(r)} J$ and $\varPhi^{(r)} = \varDelta^{(r)} X$, respectively. According to (24),

$$\varDelta^{(r)} X = I\{\varDelta^{(r)} J\} = \hat{R} J^{-1} \varDelta^{(r)} J = \hat{R} J^{-1} Q^{(r)} \ ; \quad r = i, j \tag{25}$$

By the Lie-algebraic property of symmetries of PDEs, the functions $[\varDelta^{(i)}, \varDelta^{(j)}] J$ and $[\varDelta^{(i)}, \varDelta^{(j)}] X$ also represent symmetry characteristics for SDYM and PSDYM, respectively, where we have put

$$\begin{aligned}&[\varDelta^{(i)}, \varDelta^{(j)}] J \equiv \varDelta^{(i)} \varDelta^{(j)} J - \varDelta^{(j)} \varDelta^{(i)} J = \varDelta^{(i)} Q^{(j)} - \varDelta^{(j)} Q^{(i)} \,, \\ &[\varDelta^{(i)}, \varDelta^{(j)}] X \equiv \varDelta^{(i)} \varDelta^{(j)} X - \varDelta^{(j)} \varDelta^{(i)} X = \varDelta^{(i)} \varPhi^{(j)} - \varDelta^{(j)} \varPhi^{(i)} \,.\end{aligned}$$

We must now verify that

$$[\varDelta^{(i)}, \varDelta^{(j)}]\,X = I\,\{[\varDelta^{(i)}, \varDelta^{(j)}]\,J\} = \hat{R}\,J^{-1}[\varDelta^{(i)}, \varDelta^{(j)}]\,J \qquad (26)$$

Putting $r=j$ into (25), and applying the Fréchet derivative $\varDelta^{(i)}$, we have:

$$\begin{aligned}\varDelta^{(i)}\varDelta^{(j)}X &= \varDelta^{(i)}\hat{R}\,J^{-1}Q^{(j)} = [\varDelta^{(i)}, \hat{R}]J^{-1}Q^{(j)} + \hat{R}\,\varDelta^{(i)}J^{-1}Q^{(j)} \\ &= D_{\bar{z}}^{-1}[\varPhi_{\bar{z}}^{(i)}, J^{-1}Q^{(j)}] + \hat{R}\,\varDelta^{(i)}J^{-1}Q^{(j)},\end{aligned}$$

where we have used the commutation relation (22). By (23) and (19),

$$\varPhi_{\bar{z}}^{(i)} = D_{\bar{z}}\,\hat{R}\,J^{-1}Q^{(i)} = \hat{A}_y\,J^{-1}Q^{(i)}.$$

Moreover, by properties (5) and (8) of the Fréchet derivative,

$$\begin{aligned}\varDelta^{(i)}J^{-1}Q^{(j)} &= -J^{-1}(\varDelta^{(i)}J)\,J^{-1}Q^{(j)} + J^{-1}\varDelta^{(i)}Q^{(j)} \\ &= -\,J^{-1}Q^{(i)}J^{-1}Q^{(j)} + J^{-1}\varDelta^{(i)}Q^{(j)}.\end{aligned}$$

So,

$$\varDelta^{(i)}\varDelta^{(j)}X = D_{\bar{z}}^{-1}[\hat{A}_y J^{-1}Q^{(i)}, J^{-1}Q^{(j)}] - \hat{R}\,J^{-1}Q^{(i)}J^{-1}Q^{(j)} + \hat{R}\,J^{-1}\varDelta^{(i)}Q^{(j)}.$$

Subtracting from this the analogous expression for $\varDelta^{(j)}\varDelta^{(i)}X$, we have:

$$\begin{aligned}[\varDelta^{(i)}, \varDelta^{(j)}]X &\equiv \varDelta^{(i)}\varDelta^{(j)}X - \varDelta^{(j)}\varDelta^{(i)}X \\ &= D_{\bar{z}}^{-1}\Big([\hat{A}_y J^{-1}Q^{(i)}, J^{-1}Q^{(j)}] + [J^{-1}Q^{(i)}, \hat{A}_y J^{-1}Q^{(j)}]\Big) \\ &\qquad - \hat{R}\,[J^{-1}Q^{(i)}, J^{-1}Q^{(j)}] + \hat{R}\,J^{-1}(\varDelta^{(i)}Q^{(j)} - \varDelta^{(j)}Q^{(i)}) \\ &= D_{\bar{z}}^{-1}\hat{A}_y\,[J^{-1}Q^{(i)}, J^{-1}Q^{(j)}] - \hat{R}\,[J^{-1}Q^{(i)}, J^{-1}Q^{(j)}] \\ &\qquad + \hat{R}\,J^{-1}(\varDelta^{(i)}\varDelta^{(j)}J - \varDelta^{(j)}\varDelta^{(i)}J) \\ &= \hat{R}\,J^{-1}[\varDelta^{(i)}, \varDelta^{(j)}]\,J\end{aligned}$$

where we have used the derivation property (10) of $\hat{A}_y$ and we have taken (19) into account. Thus, (26) has been proven. ■

Now, suppose $\hat{P}$ is a recursion operator for SDYM, while $\hat{S}$ is a recursion operator for PSDYM. Thus, if Q and $\varPhi$ are symmetry characteristics for SDYM and PSDYM, respectively, then $Q' = \hat{P}Q$ and $\varPhi' = \hat{S}\varPhi$ also are symmetry characteristics.

Definition: The linear operators $\hat{P}$ and $\hat{S}$ will be called *equivalent with respect to the isomorphism I* (or *I-equivalent*) if the following condition is satisfied:

$$\hat{S}\varPhi = I\,\{\hat{P}Q\} \quad \textit{when} \quad \varPhi = I\,\{Q\} \qquad (27)$$

By using (23), the above condition is written

$$\hat{S}\Phi = \hat{R}J^{-1}\hat{P}Q \quad when \quad \Phi = \hat{R}J^{-1}Q \;\Rightarrow\; \hat{S}\hat{R}J^{-1}Q = \hat{R}J^{-1}\hat{P}Q \ .$$

Thus, in order that $\hat{P}$ and $\hat{S}$ be *I*-equivalent recursion operators, the following operator equation must be satisfied on the infinite-dimensional linear space of all SDYM symmetry characteristics:

$$\hat{S}\hat{R}J^{-1} = \hat{R}J^{-1}\hat{P} \qquad (28)$$

Having already found a PSDYM recursion operator $\hat{S}=\hat{R}$, we now want to evaluate the *I*-equivalent SDYM recursion operator $\hat{P}$. To this end, we put $\hat{S}=\hat{R}$ in (28) and write

$$\hat{R}\,(\hat{R}J^{-1} - J^{-1}\hat{P}) = 0 \ .$$

As is easy to see, this is satisfied for $\hat{P}=\hat{T}$, in view of (21). We thus conclude that

- the recursion operators $\hat{R}$ and $\hat{T}$, defined by (19) and (21), are *I*-equivalent.

We note that (28) is a sort of BT relating recursion operators of different PDEs, rather than solutions or symmetries of these PDEs. Thus, if a recursion operator is known for either PDE, this BT will yield a corresponding operator for the other PDE. Note that we have encountered BTs at various levels: (*a*) The non-auto-BT (2), relating solutions of two different nonlinear PDEs (1) and (3); (*b*) the BT (17), or equivalently (18), relating symmetry characteristics of these PDEs; (*c*) the recursion operators (19) and (21), which can be re-expressed as auto-BTs for the symmetry conditions (14) and (13), respectively; and (*d*) the BT (28), relating recursion operators for the original, nonlinear PDEs. (We make the technical observation that the first three BTs are "strong", while the last one is "weak"; see Appendix.)

Example: Consider the PSDYM symmetry characteristic $\Phi = X_z$ (*z*-translation). To find the *I*-related SDYM characteristic Q, we use (23):

$$\hat{R}J^{-1}Q = \Phi \;\Rightarrow\; D_{\bar{z}}^{-1}\hat{A}_y(J^{-1}Q) = X_z \;\Rightarrow\; \hat{A}_y(J^{-1}Q) = X_{z\bar{z}} \overset{(2)}{\Rightarrow}$$
$$(J^{-1}Q)_y + [J^{-1}J_y\,,\, J^{-1}Q] = (J^{-1}J_y)_z \ ,$$

which is satisfied for $Q = J_z$. By applying the recursion operator $\hat{T}$ on Q,

$$Q' = \hat{T}Q = J\hat{R}J^{-1}Q = JD_{\bar{z}}^{-1}\hat{A}_y(J^{-1}J_z) = JD_{\bar{z}}^{-1}\left\{(J^{-1}J_z)_y + [J^{-1}J_y, J^{-1}J_z]\right\}$$
$$= JD_{\bar{z}}^{-1}(J^{-1}J_y)_z \overset{(2)}{=} JD_{\bar{z}}^{-1}X_{z\bar{z}} = JX_z \ .$$

To find the *I*-related PSDYM characteristic Φ', we use (23) once more:

$$\Phi' = \hat{R}J^{-1}Q' = \hat{R}X_z = \hat{R}\Phi \ .$$

We notice that $\hat{R}\Phi = I\{\hat{T}Q\}$ when $\Phi = I\{Q\}$, as expected by the fact that $\hat{R}$ and $\hat{T}$ are *I*-equivalent recursion operators. ■

Now, let $Q^{(0)}$ be some SDYM symmetry characteristic. By repeated application of the recursion operator $\hat{T}$, we obtain an infinite sequence of such characteristics:

$$Q^{(1)} = \hat{T}Q^{(0)},\ \ Q^{(2)} = \hat{T}Q^{(1)} = \hat{T}^2 Q^{(0)},\ \cdots\ ,\ \ Q^{(n)} = \hat{T}Q^{(n-1)} = \hat{T}^n Q^{(0)},\ \cdots$$

(we note that any power of a recursion operator also is a recursion operator). Also, let

$$\Phi^{(0)} = I\{Q^{(0)}\} = \hat{R}J^{-1}Q^{(0)} \tag{29}$$

be the PSDYM characteristic which is *I*-related to $Q^{(0)}$. Repeated application of the PSDYM recursion operator $\hat{R}$ will now yield an infinite sequence of PSDYM characteristics. Taking into account that $\hat{R}$ and $\hat{T}$ are *I*-equivalent recursion operators, we can write this sequence as follows:

$$\Phi^{(1)} = \hat{R}\Phi^{(0)} = I\{\hat{T}Q^{(0)}\}\,,\quad \Phi^{(2)} = \hat{R}^2\Phi^{(0)} = I\{\hat{T}^2 Q^{(0)}\}\,,\ \cdots\ ,$$
$$\Phi^{(n)} = \hat{R}^n\Phi^{(0)} = I\{\hat{T}^n Q^{(0)}\}\,,\ \cdots$$

Assume now that the infinite set of SDYM symmetries represented by the characteristics $\{Q^{(n)}\}$ $(n=0,1,2,\cdots)$ has the structure of a Lie algebra. This set then constitutes a symmetry subalgebra of SDYM. Given that the set $\{\Phi^{(n)}\}$ is *I*-related to $\{Q^{(n)}\}$ and that *I* is a Lie-algebra isomorphism, we conclude that the infinite set of characteristics $\{\Phi^{(n)}\}$ corresponds to a symmetry subalgebra of PSDYM which is isomorphic to the associated subalgebra $\{Q^{(n)}\}$ of SDYM.

More generally, let $\{Q_k^{(0)} / k=1,2,\cdots,p\}$ be a finite set of SDYM symmetry characteristics, and let $\{\Phi_k^{(0)} / k=1,2,\cdots,p\}$ be the *I*-related set of PSDYM characteristics, where

$$\Phi_k^{(0)} = I\{Q_k^{(0)}\} = \hat{R}J^{-1}Q_k^{(0)}\ ;\ \ k=1,2,\cdots,p \tag{30}$$

Assume that the infinite set of characteristics

$$\{Q_k^{(n)} = \hat{T}^n Q_k^{(0)} / n=0,1,2,\cdots;\ k=1,2,\cdots,p\} \tag{31}$$

corresponds to a Lie subalgebra of SDYM symmetries. Then, the *I*-related set of characteristics

$$\{\Phi_k^{(n)} = \hat{R}^n \Phi_k^{(0)} / n=0,1,2,\cdots;\ k=1,2,\cdots,p\} \tag{32}$$

corresponds to a PSDYM symmetry subalgebra which is isomorphic to that of (31).

Let us summarize our main conclusions:

- The infinite-dimensional symmetry Lie algebras of SDYM and PSDYM are isomorphic, the isomorphism I being defined by (23) or (24).
- The recursion operators $\hat{T}$ and $\hat{R}$, defined in (21) and (19), when applied to I-related symmetry characteristics [such as those in (29) or (30)], may generate isomorphic, infinite-dimensional symmetry subalgebras of SDYM and PSDYM, respectively.
- Since the structures of the symmetry Lie algebras of SDYM and PSDYM are similar, all results regarding the latter structure are also applicable to the SDYM case.

Comment: At this point the reader may wonder whether it is really necessary to go through the PSDYM symmetry problem in order to solve the corresponding SDYM problem. In principle, of course, the SDYM case can be treated on its own. In practice, however, it is easier to study the symmetry structure of PSDYM first and then take advantage of the isomorphism between this structure and that of SDYM. This statement is justified by the fact that the PSDYM recursion operator is considerably easier to handle compared to the corresponding SDYM operator. This property of the former operator is of great value in the interest of computational simplicity (in particular, for the purpose of deriving various commutation relations; cf. [8]).

IV. Potential Symmetries and Current Algebras

We recall that every SDYM symmetry characteristic can be expressed as $Q=J\Phi$, where Φ is a PSDYM characteristic (we note that Φ is *not* I-related to Q). Let Φ be a characteristic which depends locally or nonlocally on X and/or various derivatives of X. By the BT (2), X must be an integral of J and its derivatives, and so it and its derivatives X_y and X_z are nonlocal in J. On the other hand, according to (2), the quantities $X_{\bar{y}}$ and $X_{\bar{z}}$ depend locally on J. Thus, in general, Φ can be local or nonlocal in J. In the case where Φ is *nonlocal* in J, we say that the characteristic $Q=J\Phi$ expresses a *potential symmetry* of SDYM [3, 5]. (See Appendix for a general definition of locality and nonlocality of symmetries.)

A. Internal Symmetries. The PSDYM equation is generally invariant under a transformation of the form

$$\Delta^{(0)}X = \Phi^{(0)} = [X, M] \tag{33}$$

where M is any constant $sl(N,C)$ matrix. Since the characteristic $\Phi^{(0)}$ is nonlocal in J, the transformation

$$Q = J\Phi^{(0)} = J\,[X, M]$$

is a genuine potential symmetry of SDYM. Note that the SDYM characteristic which is I-related to $\Phi^{(0)}$ is not Q, but rather $Q^{(0)}= JM$, since we then have

$$\hat{R} J^{-1} Q^{(0)} = \hat{R} M = D_{\bar{z}}^{-1} [X_{\bar{z}}, M] = [X, M] = \Phi^{(0)} .$$

Let $\{\tau_k\}$ be a basis for $sl(N,C)$:

$$[\tau_i, \tau_j] = C_{ij}^k \tau_k \ .$$

Then M is expanded as $M = \alpha^k \tau_k$, and (33) is resolved into a set of independent basis transformations

$$\Delta_k^{(0)} X = \Phi_k^{(0)} = [X, \tau_k]$$

corresponding to the SDYM potential symmetries

$$Q_k = J \Phi_k^{(0)} = J [X, \tau_k] \ .$$

These are not the same as the I-related characteristics

$$\Delta_k^{(0)} J = Q_k^{(0)} = J \tau_k \ .$$

Consider now the infinite set of transformations

$$\Delta_k^{(n)} X = \Phi_k^{(n)} = \hat{R}^n \Phi_k^{(0)} = \hat{R}^n [X, \tau_k] \quad (n = 0, 1, 2, \cdots) \qquad (34)$$

As can be shown, they satisfy the commutation relations of a Kac-Moody algebra:

$$[\Delta_i^{(m)}, \Delta_j^{(n)}] X = C_{ij}^k \Delta_k^{(m+n)} X \ .$$

In view of the isomorphism I, this structure is also present in SDYM. Indeed, this is precisely the familiar hidden symmetry of SDYM [11, 12]. The SDYM transformations which are I-related to those in (34) are given by

$$\Delta_k^{(n)} J = Q_k^{(n)} = \hat{T}^n Q_k^{(0)} = \hat{T}^n J \tau_k \quad (n = 0, 1, 2, \cdots) \ .$$

They constitute an infinite set of potential symmetries (note, for example, that $\Delta_k^{(1)} J = J [X, \tau_k] = J \Phi_k^{(0)}$) and they satisfy the commutation relations

$$[\Delta_i^{(m)}, \Delta_j^{(n)}] J = C_{ij}^k \Delta_k^{(m+n)} J \ .$$

B. Symmetries in the Base Space. A number of local PSDYM symmetries corresponding to coordinate transformations are nonlocal in J, hence lead to potential symmetries of SDYM. By using isovector methods [4, 15], nine such PSDYM symmetries can be found. They can be expressed as follows:

$$\Delta_k^{(0)} X = \Phi_k^{(0)} = \hat{L}_k X \quad (k = 1, 2, \cdots, 9) \tag{35}$$

where the $\hat{L}_k$ are nine linear operators which are explicitly given by

$$\hat{L}_1 = D_y \,, \quad \hat{L}_2 = D_z \,, \quad \hat{L}_3 = z D_y - \bar{y} D_{\bar{z}} \,, \quad \hat{L}_4 = y D_z - \bar{z} D_{\bar{y}} \,,$$
$$\hat{L}_5 = y D_y - z D_z - \bar{y} D_{\bar{y}} + \bar{z} D_{\bar{z}} \,, \quad \hat{L}_6 = 1 + y D_y + z D_z \,,$$
$$\hat{L}_7 = 1 - \bar{y} D_{\bar{y}} - \bar{z} D_{\bar{z}} \,, \quad \hat{L}_8 = y \hat{L}_6 + \bar{z} (y D_{\bar{z}} - z D_{\bar{y}}) \,,$$
$$\hat{L}_9 = z \hat{L}_6 + \bar{y} (z D_{\bar{y}} - y D_{\bar{z}}) \,.$$

The $\hat{L}_1$, $\hat{L}_2$ represent translations of y and z, respectively, while the $\hat{L}_3$, $\hat{L}_4$ represent rotational symmetries. The $\hat{L}_5$, $\hat{L}_6$, $\hat{L}_7$ express scale transformations, while $\hat{L}_8$ and $\hat{L}_9$ represent nonlinear coordinate transformations which presumably reflect the special conformal invariance of the SDYM equations in their original, covariant form.

The first five operators $\hat{L}_1, \cdots, \hat{L}_5$ form the basis of a Lie algebra, the commutation relations of which we write in the form

$$[\hat{L}_i \,, \hat{L}_j] = - f_{ij}^{\,k} \hat{L}_k \quad (k = 1, \cdots, 5) \,.$$

Consider now the infinite set of transformations

$$\Delta_k^{(n)} X = \Phi_k^{(n)} = \hat{R}^n \Phi_k^{(0)} = \hat{R}^n \hat{L}_k X \quad (k = 1, \cdots, 5) \tag{36}$$

As can be shown [8], these form a Kac-Moody algebra:

$$[\Delta_i^{(m)}, \Delta_j^{(n)}] X = f_{ij}^{\,k} \Delta_k^{(m+n)} X \,.$$

Consider also the infinite sets of transformations

$$\Delta^{(n)} X = \hat{R}^n \hat{L}_6 X \quad \text{and} \quad \Delta^{(n)} X = \hat{R}^n \hat{L}_7 X \tag{37}$$

As can be proven [8], each set forms a Virasoro algebra (apart from a sign):

$$[\Delta^{(m)}, \Delta^{(n)}] X = -(m-n) \Delta^{(m+n)} X \,.$$

Taking the isomorphism I into account, we conclude that the SDYM symmetry algebra possesses both Kac-Moody and Virasoro subalgebras ("current algebras" [16]), both of which are associated with infinite sets of potential symmetries. The former subalgebras are associated with both internal and coordinate transformations, while the latter ones are related to coordinate transformations only. These conclusions are in agreement with those of [13], although the mathematical approach there is different from ours.

V. Summary

By using the SDYM-PSDYM system as a model, we have studied a process for associating symmetries and recursion operators of two nonlinear PDEs related to each other by a non-auto-BT. The concept of a BT itself enters our analysis at various levels: (*a*) The non-auto BT (2) relates solutions of the nonlinear PDEs (1) and (3); (*b*) the non-auto-BT (17) or (18) relates symmetry characteristics of these PDEs; (*c*) the auto-BTs for the symmetry conditions (14) and (13) lead to the recursion operators (19) and (21), respectively; and (*d*) the transformation (28) may be perceived as a BT associating recursion operators for the original, nonlinear PDEs. We have proven the isomorphism between the infinite-dimensional symmetry Lie algebras of SDYM and PSDYM, and we have used this property to draw several conclusions regarding the Lie-algebraic structure of the potential symmetries of SDYM.

For further reading on recursion operators, the reader is referred to [17-22]. A nice discussion of the SDYM symmetry structure and its connection to the existence of infinitely many conservation laws can be found in the paper by Adam *et al.* [23].

VI. Appendix: Some Basic Definitions

To make the paper as self-contained as possible, basic definitions of some key concepts that are being used are given below:

A. Recursion Operators. Consider a PDE $F[u]=0$, in the dependent variable u and the independent variables x^μ ($\mu=1,2,...$). Let $\delta u=\alpha Q[u]$ be an infinitesimal symmetry transformation of the PDE, where $Q[u]$ is the *symmetry characteristic*. The symmetry is generated by the (formal) vector field

$$V = Q[u]\frac{\partial}{\partial u} + \textit{prolongation} = Q\frac{\partial}{\partial u} + Q_\mu \frac{\partial}{\partial u_\mu} + Q_{\mu\nu}\frac{\partial}{\partial u_{\mu\nu}} + \cdots \qquad (A.1)$$

(where the $Q_\mu \equiv D_\mu Q$, etc., denote total derivatives of Q). The symmetry condition is expressed by a PDE, linear in Q:

$$S(Q;u) \equiv \varDelta F[u] = 0 \mod F[u] \qquad (A.2)$$

where $\varDelta$ denotes the Fréchet derivative with respect to V. If u is a scalar quantity, then (A.2) takes the form

$$S(Q;u) = V F[u] = Q\frac{\partial F}{\partial u} + Q_\mu \frac{\partial F}{\partial u_\mu} + Q_{\mu\nu}\frac{\partial F}{\partial u_{\mu\nu}} + \cdots = 0 \mod F[u] \qquad (A.3)$$

Since the PDE (A.2) is linear in Q, the sum of two solutions (for the same u) also is a solution. Thus, for any given u, the solutions $\{Q[u]\}$ of (A.2) form a linear space S_u. A *recursion operator* $\hat{R}$ is a linear operator which maps the space S_u into itself. Thus, if Q is a symmetry characteristic of $F[u]=0$ [i.e., a solution of (A.2)], then so is $\hat{R}Q$:

$$S(\hat{R}Q;u) = 0 \quad \textit{when} \quad S(Q;u) = 0 \qquad (A.4)$$

We note that $\hat{R}^2 Q, \hat{R}^3 Q, \cdots, \hat{R}^n Q$ also are symmetry characteristics. This means that

- any power $\hat{R}^n$ of a recursion operator also is a recursion operator.

Thus, starting with any symmetry characteristic Q, we can obtain an infinite set of such characteristics by repeated application of the recursion operator.

A *symmetry operator* $\hat{L}$ is a linear operator, independent of u, which produces a symmetry characteristic $Q[u]$ when it acts on u. Thus, $\hat{L}u = Q[u]$. We note that $\hat{R}\hat{L}u$ is a symmetry characteristic, which means that

- the product $\hat{R}\hat{L}$ of a recursion operator and a symmetry operator is a symmetry operator.

Thus, given that $\hat{R}^n$ is a recursion operator, we conclude that $\hat{R}^n\hat{L}u$ is a member of S_u. Examples of symmetry operators are the nine operators $\hat{L}_k$ that appear in (35), as well as the operator $\hat{L} = [\ \ , M\,]$ which is implicitly defined in (33).

B. "Strong" and "Weak" Bäcklund Transformations. In the most general sense, a BT is a set of relations (typically differential, although in certain cases algebraic ones are also considered) which connect solutions of two different PDEs (non-auto-BT) or of the same PDE (auto-BT). The technical distinction between "strong" and "weak" BTs [24, 25] can be roughly described as follows: In a strong BT connecting, say, the variables u and v, integrability of the differential system for either variable *demands* that the other variable satisfy a certain PDE. A weak BT, on the other hand, is much like a symmetry transformation: u and v are not, *a priori*, required to satisfy any particular PDEs for integrability. *If*, however, u satisfies some specific PDE, *then* v satisfies some related PDE. (An example is the Cole-Hopf transformation, connecting solutions of Burgers' equation to solutions of the heat equation.)

The BT (2) is strong, since its integrability conditions *force* the functions J and X to satisfy the PDEs (1) and (3), respectively. Similar remarks apply to the BTs (17) and (19*a*). On the other hand, transformation (28) does not *a priori* impose any specific properties on the operators $\hat{P}$ and $\hat{S}$. *If*, however, $\hat{P}$ is an SDYM recursion operator, *then* $\hat{S}$ is the *I*-equivalent PSDYM recursion operator. Thus, equation (28) is a Bäcklund-like transformation of the weak type, although this particular transformation relates operators rather than functions.

C. Local and Nonlocal Symmetries. Let $F[u]=0$ be a PDE in the dependent variable u and the independent variables x^μ (μ=1,2,...). A symmetry characteristic $Q[u]$ represents a *local* symmetry of the PDE if Q depends, at most, on x^μ, u, and derivatives of u with respect to the x^μ. A symmetry is *nonlocal* if the corresponding characteristic Q contains additional variables expressed as *integrals* of u with respect to the x^μ (or, more generally, integrals of local functions of u). As an example, the PSDYM characteristic $\Phi = [X, M\,]$ (where M is a constant matrix) represents a local symmetry of this PDE (since it depends locally on the PSDYM variable X), whereas the SDYM characteristic $Q = J\,[X, M\,]$ represents a nonlocal symmetry of that PDE since it contains an additional variable X which is expressed as an integral of a local function of the principal SDYM variable J [this follows from the BT (2)]. The infinite symmetries (34), (36) and (37) are increasingly nonlocal in X for n>0, since they are produced by repeated application of the integro-differential recursion operator $\hat{R}$.

Acknowledgment

We appreciate the kind assistance of Kathleen O'Shea-Arapoglou.

References

1. P.J. Olver, J. Math. Phys. 18 (1977) 1212.
2. P.J. Olver, *Applications of Lie Groups to Differential Equations*, 2nd ed. (Springer-Verlag, 1993).
3. G.W. Bluman and S. Kumei, *Symmetries and Differential Equations* (Springer-Verlag, 1989).
4. C.J. Papachristou and B.K. Harrison, J. Math. Phys. 28 (1987) 1261; J. Math. Phys. 29 (1988) 238.
5. C.J. Papachristou, Phys. Lett. A 145 (1990) 250.
6. C.J. Papachristou, J. Phys. A: Math. Gen. 24 (1991) L 1051.
7. M. Marvan, *Another Look on Recursion Operators*, in *Differential Geometry and Applications* (Proc. Conf., Brno, 1995) 393.
8. C.J. Papachristou, Phys. Lett. A 154 (1991) 29.
9. R.S. Ward, Phil. Trans. R. Soc. A 315 (1985) 451.
10. M.J. Ablowitz and P.A. Clarkson, *Solitons, Nonlinear Evolution Equations and Inverse Scattering* (Cambridge University Press, 1991).
11. L.-L. Chau, M.L. Ge and Y.S. Wu, Phys. Rev. D 25 (1982) 1086.
12. L. Dolan, Phys. Rep. 109 (1984) 1.
13. A.D. Popov and C.R. Preitschopf, Phys. Lett. B 374 (1996) 71.
14. C.J. Papachristou, arXiv: 0803.3688 (2008).
15. C.J. Papachristou and B.K. Harrison, Acta Appl. Math. 11 (1988) 155.
16. J. Mickelsson, *Current Algebras and Groups* (Plenum Press, 1989).
17. G.A. Guthrie, Proc. R. Soc. Lond. A 446 (1994) 107.
18. A.H. Bilge, J. Phys. A: Math. Gen. 26 (1993) 7511.
19. M. Gürses, A. Karasu and V.V. Sokolov, J. Math. Phys. 40 (1999) 6473.
20. M. Marvan, Acta Appl. Math. (2009) DOI 10.1007/s10440-009-9450-4.
21. A. Sergyeyev, J. Phys. A: Math. Gen. 38 (2005) 3397.
22. J.P. Wang, J. Math. Phys. 50 (2009) 023506.
23. C. Adam, J. Sánchez-Guillén and A. Wereszczyński, JHEP09 (2008) 014.
24. P.J. McCarthy, Lett. Math. Phys. 2 (1978) 493.
25. D. Levi and Z. Popowicz, J. Phys. A: Math. Gen. 18 (1985) 215.