P. D. Gerber

Linearization of Cauchy's Problem for Quadratic Semilinear Partial Differential Equations

Abstract: The technique of linearization is applied to quadratic semilinear systems of n first-order partial differential equations. By introducing a related linear algebra $\mathscr{A}$, we combine analytic and algebraic arguments to obtain a class of linearizable systems. We relate the coefficients of such systems to the center of $\mathscr{A}$ and show that, for hyperbolic systems, the ideals of $\mathscr{A}$ decouple the system into disjoint subsystems each having its own single wave number.

Introduction to quadratic systems and their related algebras

A partial differential equation is *quadratic* if at most it is of degree two in the dependent variables and their derivatives. It is *semilinear* if it is linear in the highest-order derivatives. This paper treats systems of n first-order equations that are both semilinear and quadratic. If we confine our attention to equations in two independent variables, x and t, the most general such system for the n unknown functions $v^i = v^i(t,x)$ is

$$c^i_\alpha v^\alpha_t + a^i_\alpha v^\alpha_x = a^i_{\alpha\beta} v^\alpha v^\beta + b^i_\alpha v^\alpha. \tag{Q}$$

In (Q) the coefficients are independent of the v^i and we use the notation of tensor algebra: All free indices range from 1 to n; the summation convention applies to an index repeated as a subscript and as a superscript. Also, the subscript notation for partial derivatives is used. Thus, (Q) is a concise notation for

$$\sum_{\alpha=1}^{n} \left[c^i_\alpha \frac{\partial v^\alpha}{\partial t} + a^i_\alpha \frac{\partial v^\alpha}{\partial x} \right] = \sum_{\alpha,\beta=1}^{n} a^i_{\alpha\beta} v^\alpha v^\beta + \sum_{\alpha=1}^{n} b^i_\alpha v^\alpha, \qquad i = 1, \cdots, n.$$

We use the term *quadratic system* for any system of the form (Q). The purpose of this paper is to establish initially the great practical and mathematical importance of quadratic systems and to explore in detail the technique for "solving" them by transformation to linear systems. Future papers will explore other techniques, such as nonlinear superposition and approximation by families of particular solutions.

Some indication of the practical importance and diversity of quadratic systems is given by the following four examples:

$$u_t - \varepsilon v_x = -uv,$$
$$u_x = v;$$
$$u_t - v_x = \alpha u^2 + \beta u,$$
$$u_x = v;$$
$$u_t - w_x = -uv,$$
$$u_x = v,$$
$$v_x = w;$$

and

$$u^1_t + k^1 u^1_x = a^2 u^2 u^3 + b^1 u^1,$$
$$u^2_t + k^2 u^2_x = a^2 u^1 u^3 + b^2 u^2,$$
$$u^3_t + k^3 u^3_x = a^3 u^1 u^2 + b^3 u^3. \tag{1}$$

If we solve the first three of these systems for the variable u, we obtain, respectively,

$$u_t + uu_x = \varepsilon u_{xx}, \tag{2}$$

$$u_t = u_{xx} + \alpha u^2 + \beta u, \tag{3}$$

$$u_t + uu_x = u_{xxx}. \tag{4}$$

Equation (1) is a basic equation in nonlinear optics, which is discussed in detail by Lamb [1] and by Bloembergen [2]. Equation (2) is *Burgers' equation*, a fundamental model of dissipative ($\varepsilon \neq 0$) and conservative ($\varepsilon = 0$) systems. Equations (4) is the *Korteweg-deVries equation*, which is a model of dispersion. An excellent exposition on (2) and (4) is given by Lax [3]. Equation (3) is *Fisher's equation*, a basic model in nonlinear diffusion that is discussed by Montroll [4]. These examples, along with others from continuum mechanics and kinetic theory, indicate the practical importance of quadratic systems.

As differential equations, quadratic systems already occupy a diverse and interesting branch of analysis. Quadratic systems also occupy an important branch of algebra, because each system is closely related to a linear algebra, the *related algebra* of the system. For quadratic systems, the related algebra plays the same role as does the vector space for linear systems: The theory divides naturally into an *analytic* part—the study of the function-theoretic part, and an *algebraic* part—the study of the related algebra. (This idea on the theory of quadratic systems is due to Marcus [5].)

To introduce the related algebra, we need some preliminary algebraic material, given by Definitions 1 and 2.

Definition 1 An *algebra* $\mathscr{G}$ over a field F is a vector space over F with a binary composition (multiplication) which is bilinear with respect to the vector operations.

Thus, if $\mathscr{G}$ is an algebra and $\mathbf{u},\mathbf{v} \in \mathscr{G}$, the *product* $\mathbf{uv} \in \mathscr{G}$ is defined and satisfies

$$\mathbf{u}(\mathbf{v}+\mathbf{w}) = \mathbf{uv}+\mathbf{uw}, \qquad (\mathbf{v}+\mathbf{w})\mathbf{u} = \mathbf{vu}+\mathbf{wu},$$

$$\alpha(\mathbf{uv}) = (\alpha\mathbf{u})\mathbf{v} = \mathbf{u}(\alpha\mathbf{v}), \qquad \mathbf{u},\mathbf{v},\mathbf{w} \in \mathscr{G}, \qquad \alpha \in F.$$

The notation $\mathscr{G}$, $\mathscr{G}(*)$, $\mathscr{G}(\#)$, etc. indicates that the product of $\mathbf{u}$ and $\mathbf{v}$ is to be written $\mathbf{uv}$, $\mathbf{u} \quad \mathbf{v}$, $\mathbf{u} \# \mathbf{v}$, etc.

If $\mathscr{G}$ is an n-dimensional algebra with a basis $\mathbf{e}_i$, then $\mathbf{e}_j\mathbf{e}_k$ is defined and must be a linear combination of the $\mathbf{e}_i$. Therefore, there exist n^3 scalars $g^i_{jk} \in F$ such that

$$\mathbf{e}_j\mathbf{e}_k = g^i_{jk}\mathbf{e}_i. \tag{5}$$

The g^i_{jk} are the *structure constants of* $\mathscr{G}$ (in the basis $\mathbf{e}_i$).

Conversely, suppose $\mathscr{G}$ is an n-dimensional vector space over F with a basis $\mathbf{e}_i$. If $g^i_{jk} \in F$ are *any* n^3 scalars, then (5) *defines* a bilinear multiplication on $\mathscr{G}$ with structure constants g^i_{jk}. Equation (5) defines the product of any two basis vectors and bilinearity defines the product of arbitrary vectors. If $\mathbf{x} = x^j\mathbf{e}_j$, $\quad \mathbf{y} = y^k\mathbf{e}_k \; x^j \; \mathscr{G}$, then

$$\mathbf{xy} = (g^i_{jk}x^jy^k)\mathbf{e}_i. \tag{6}$$

The notation $\mathscr{G} = (g^i_{jk})$ indicates that $\mathscr{G}$ is the algebra with structure constants g^i_{jk} (in some fixed basis). Generally, algebras are *not* associative, nor are they commutative. See Schafer [6] for an excellent coverage of algebras.

Definition 2 The *related algebra* of (Q) is the commutative algebra $\mathscr{A} = \mathscr{A}(*) = (a^i_{jk})$.

$\mathscr{A}$ is commutative because there is no loss in generality in assuming that $a^i_{jk} = a^i_{kj}$ in (Q), and this, in (6), leads to $\mathbf{x} * \mathbf{y} = \mathbf{y} * \mathbf{x}$.

The goal of research in quadratic systems is to relate the analytic and qualitative properties of the differential equations and the algebraic properties of the related algebra.

For the remainder of this paper we treat the initial value problem for (Q), when the matrix $C = (c^i_j)$ is the identity matrix $I = (\delta^i_j)$. When the vectors $\mathbf{v} = [v^i(t,x)]$, $\boldsymbol{\psi} = \boldsymbol{\psi}(x) = [\psi^i(x)]$ are introduced, the matrices $A = (a^i_j)$, $B = (b^i_j)$, and the related algebra $\mathscr{A} = \mathscr{A}(*) = (a^i_{jk})$, (Q) becomes

$$\mathbf{v}_t + A\mathbf{v}_x = \mathbf{v} * \mathbf{v} + B\mathbf{v}, \tag{7}$$

$$\mathbf{v}(0,x) = \boldsymbol{\psi}(x). \tag{8}$$

In the second section, we treat the problem of solving (7) and (8) with the technique of linearization. For example, find a function $\mathbf{f} = \mathbf{f}(t,\mathbf{u})$ and matrices $\hat{A},\hat{B}$ $[\mathbf{f} = (f^i), \quad \mathbf{u} = (u^i), \quad \hat{A} = (\hat{a}^i_j), \quad \hat{B} = (\hat{b}^i_j)]$ such that the transformation $\mathbf{v} = f(t,\mathbf{u})$ reduces (7) to the linear system $\mathbf{u}_t + \hat{A}\mathbf{u}_x = \hat{B}\mathbf{u}$. Analytic and algebraic criteria for the existence of a linearization are then obtained. In particular, if the system (7) is *A-associative*, then $\mathscr{A}$, A,B satisfy

$$e^{-Bt}Ae^{Bt}(\mathbf{x} * \mathbf{y}) = (e^{-Bt}Ae^{Bt}\mathbf{x}), \qquad \mathbf{x},\mathbf{y} \in \mathscr{A}, \tag{9}$$

and we show that a linearization can be constructed by solving the ordinary differential equation

$$\dot{\mathbf{h}} = \mathbf{h} * \mathbf{h} + B\mathbf{h}, \qquad \mathbf{h}(0) = \boldsymbol{\xi}.$$

The third section relates matrices A,B that satisfy (9) to $\mathscr{Z}(\mathscr{A})$, the *center of* $\mathscr{A}$, i.e., the set of elements which associate with all elements of $\mathscr{A}$. An algebra is termed *unital* if it contains a unity. If $\mathscr{A}$ is unital, all the matrices $e^{-Bt}Ae^{Bt}$ satisfying (9) lie in the *multiplication algebra* of $\mathscr{Z}$, i.e., the set of all linear operators $M_z : \mathbf{x} \to \mathbf{z} \quad \mathbf{x}$, $\quad \mathbf{z} \in \mathscr{Z}$, $\quad \mathbf{x} \in \mathscr{A}$. A curious outcome of this is that these matrices form a field when $\mathscr{A}$ is simple (i.e., when $\mathscr{A}$ has no proper ideals).

We consider in the fourth section systems which are A-associative wave equations. For such systems the eigenspaces of A are ideas of $\mathscr{A}$ and these spaces decouple the system into subsystems, each with a single wave velocity. We conclude the paper by showing that the *only* optics equations of the type (1) that can be linearized are those with $k^1 = k^2 = k^3$.

Complete linearizations of the Cauchy problem

Cauchy's problem for the system of n first-order partial differential equations

$$v^i_t + a^i_\alpha v^\alpha_x = a^i_{\alpha\beta}v^\alpha v^\beta + b^i_\alpha v^\alpha$$

is directed toward an analytic solution $v^i = v^i(t,x)$ that satisfies the initial condition $v^i(0,x) = \psi^i(x)$. We assume that the tensors $\mathscr{A} = (a^i_{jk})$, $A = (a^i_j)$, $B = (b^i_j)$ are constant and that the initial function $\boldsymbol{\psi} = (\psi^i(x))$ is an analytic function of x.

In the notation of the first section these equations become

$$\mathbf{v}_t + A\mathbf{v}_x = \mathbf{v} * \mathbf{v} + B\mathbf{v}, \tag{10}$$

$$\mathbf{v}(0,x) = \boldsymbol{\psi}(x), \tag{11}$$

with $\mathbf{v} = (v^i)$, $\boldsymbol{\psi} = (\psi^i)$, and $*$ the multiplication in the related algebra $\mathscr{A} = \mathscr{A}(*) = (a^i_{jk})$.

It follows from the Cauchy-Kowalewsky theorem [7] that this Cauchy problem has a unique solution $\mathbf{v} = \mathbf{v}(t,x;\boldsymbol{\psi})$. We define the general solution of (10) to be this quantity $\mathbf{v}$, where $\boldsymbol{\psi}$ ranges over all functions which are analytic on $(-\infty,\infty)$.

This method of Cauchy-Kowalewsky establishes the existence of $\mathbf{v}(t,x;\boldsymbol{\psi})$ by showing that a unique formal power series for $\mathbf{v}$ exists and converges. As a practical method this approach has serious drawbacks. For example, in practice the power series cannot be constructed nor summed; little or no qualitative information about the solution is revealed; and the method does not lead to a representation of the general solution.

Our goal is to find an approach to the Cauchy problem, (10) and (11), that overcomes the drawbacks of the method of power series. To obtain some idea as to how to start, we consider two special cases: $n = 1$, when (10) is a scalar equation; and $A = 0$, when (10) is a system of ordinary differential equations.

Taking first the scalar case, we observe that (10), (11) can be solved by the *method of characteristics* [8]. If A, B, and v are scalars and we retain the notation $v * v$ for av^2, the characteristic equations of (10), (11) are

$$\dot{t} = 1, \qquad t(0) = 0,$$

$$\dot{x} = A, \qquad x(0) = \sigma,$$

$$\dot{v} = v * v + Bv, \qquad v(0) = \psi(\sigma).$$

Here x and t are functions of s and σ and the derivatives are with respect to s.

The first equation gives $s = t$, while the second gives $x = As + \sigma$, which together give $\sigma = x - At$. The solution of the third equation is $v = h[s,\psi(\sigma)]$, in which $h = h(t,\xi)$ is the solution of $\dot{h} = h * h + Bh$, $h(0,\xi) = \xi$.

Combining these results gives

$$v(t,x) = h[t,\psi(x - At)], \tag{12}$$

this being the general solution of the Cauchy problem (10), (11) when $n = 1$.

In this scalar case, the equation for h is really the Riccati equation $\dot{h} = ah^2 + Bh$, $h(0) = \xi$, which is easily integrated to give

$$h(t,\xi) = Be^{Bt}\xi/[-a(e^{Bt} - 1)\xi + B].$$

This representation for v admits an entirely different interpretation. If we let $\phi = \phi(u)$ be an invertible function with inverse ϕ^{-1}, then (12) is equivalent to

$$v(x,t) = h[t,\phi(u)] = f(t,u), \tag{13}$$

where $u = u(t,x)$ is the solution of the *linear* Cauchy problem

$$u_t + Au_x = 0, \tag{14}$$

$$u(0,x) = \phi^{-1}[\psi(x)]. \tag{15}$$

Regarding (13) as a transformation on the solutions of (14), we see that the pair (13), (14) is a linearization of (10): If u is a solution of the linear equation (14), $f(t,u)$ is a solution of the nonlinear equation (10). The linearization is also *complete*, in the sense that $f(t,u)$ generates *all* solutions of (10) as u ranges over all solutions of (14). To solve (10), (11), take u to be the solution of (14) which satisfies (15).

For $n = 1$, the method of characteristics, which led to the representation (12), and the method of linearization (13), (14) are equivalent. For $n \neq 0$, neither the method of characteristics nor the representation (12) makes sense, but the transformation scheme does. This suggests that we attempt to solve (10) by finding a complete linearization of the form (13), (14).

In an attempt to linearize a partial differential equation, the choice of the linear equation is critical. To see if (14) is up to the task, we consider the second special case of (10) in which $A = 0$. The Cauchy problem then becomes the initial value problem $\dot{\mathbf{v}} = \mathbf{v} * \mathbf{v} + B\mathbf{v}$, $\mathbf{v}(0) = \boldsymbol{\psi}$.

Here it seems natural to regard this equation as a nonlinear (quadratic) deformation of a linear equation. The solution then ought to be a deformation of the solution of the linear equation. This suggests the transformation scheme $\mathbf{v} = \mathbf{f}(t,\mathbf{u})$, $\dot{\mathbf{u}} = B\mathbf{u}$.

This representation of $\mathbf{v}$ is also natural if we expect the behavior of $\mathbf{v}$ to decompose into a linear part, described by $\mathbf{u}$, and a nonlinear part, described by $\mathbf{f}$. Whatever the motivation, it is a sound one, for complete linearizations of this form abound:

Lemma 1 Let $\mathbf{f} = \mathbf{f}(t,\mathbf{u})$ be analytic in t,$\mathbf{u}$ for all $\mathbf{u}$ and all t near 0 and let $\boldsymbol{\phi} = \boldsymbol{\phi}(\mathbf{u}) = \mathbf{f}(0,\mathbf{u})$. Then the transformation scheme

$$\mathbf{v} = \mathbf{f}(t,\mathbf{u}), \tag{16}$$

$$\dot{\mathbf{u}} = B\mathbf{u}, \qquad \mathbf{u}(0) = \boldsymbol{\eta} \tag{17}$$

is a linearization of

$$\dot{\mathbf{v}} = \mathbf{v} * \mathbf{v} + B\mathbf{v} \tag{18}$$

if and only if

$$\mathbf{f}(t,\mathbf{u}) = \mathbf{h}[t,\boldsymbol{\phi}(e^{-Bt}\mathbf{u})], \tag{19}$$

where $\mathbf{h} = \mathbf{h}(t,\boldsymbol{\xi})$ is the solution of

$$\dot{\mathbf{h}} = \mathbf{h} * \mathbf{h} + B\mathbf{h}, \qquad \mathbf{h}(0, \xi) = \xi. \tag{20}$$

Proof The steps which lead to the formula (19) appear in the proof of Theorem 1. A simpler, direct proof follows.

Assume $\mathbf{f}$ is a linearization. The general solution of (17) is

$$\mathbf{u} = e^{Bt}\boldsymbol{\eta}, \tag{21}$$

so $\mathbf{v} = \mathbf{f}(t, e^{Bt}\boldsymbol{\eta})$ is the unique solution of (18) that satisfies $\mathbf{v}(0) = \boldsymbol{\phi}(\boldsymbol{\eta})$. But (20) then implies $\mathbf{v} = \mathbf{h}[t, \boldsymbol{\phi}(\boldsymbol{\eta})] = \mathbf{f}(t, e^{Bt}\boldsymbol{\eta})$. This and (21) imply (19).

Conversely, if $\mathbf{f}$ is given by (19) and $\mathbf{u}$ is the solution of (17), then $\mathbf{v}$ is given by $\mathbf{v}(t) = \mathbf{h}[t, \boldsymbol{\phi}(e^{-Bt}e^{Bt}\boldsymbol{\eta})] = \mathbf{h}[t, \boldsymbol{\phi}(\boldsymbol{\eta})]$, which is a solution of (18) because of (20). Therefore $\mathbf{f}$ is a linearization, which completes the proof.

Unfortunately, Lemma 1 is of little practical use in solving (18). Construction of the linearization by formula (19) requires the function $\mathbf{h}$, and $\mathbf{h}$ is already the general solution (18)! However, the lemma does provide a representation of linearizations and this can be useful.

The lemma also shows that the number of linearizations is immense. There is an n-function-parameter family of them, since the n functions ϕ^i in (19) are arbitrary. Within this multitude we would expect particular ones to be relatively simple. This suggests approaching (18) by seeking side conditions on the linearization that select the simple ones. An example is Euler's linearization of the Riccati equation; see Watson [9]. The author has successfully applied this approach [10] to a large class of quadratic systems, obtaining a generalization of Euler's linearization.

Both special cases of (10) considered so far indicate that the method of linearization can be a successful approach to (10). The obvious transformation scheme to try on the general Cauchy problem is $\mathbf{v} = \mathbf{f}(t, \mathbf{u})$, $\mathbf{u}_t + \hat{A}\mathbf{u}_x = \hat{B}\mathbf{u}$, where $\hat{A}, \hat{B}$ are constant matrices. The first of these coincides with (13) and (16) and the second is the simplest linear equation that can be made to specialize to both (14) and (17).

Formalizing our earlier discussion, we now give a definition of the linearization of (10).

Definition 3 The transformation scheme (22) and (23) is a linearization of (10) near $t = 0$ if and only if $\mathbf{f}(t, \mathbf{u})$ is a solution of (10) near 0 whenever $\mathbf{u}$ is a solution of (23). The linearization is complete if and only if every solution of (10) and (11) is given by $\mathbf{v} = \mathbf{f}(t, \mathbf{u})$ for some $\mathbf{u}$ satisfying (23).

In sharp contrast to Lemma 1, the very existence of a linearization of (10) is in doubt, and so we proceed by trying to identify those equations that can be linearized. Our success in this search is guaranteed, since the case $A = 0$ and the case for which (10) is n decoupled scalar equations are examples.

The first step is to reduce the existence question to one that satifies some analytic criteria. The following theorem does this and should be regarded as a direct generalization of Lemma 1.

First we give some convenient notation. If $\mathbf{u} = (u^i)$ and $\mathbf{g} = (g^i(\mathbf{u}))$, we denote the matrix of partial derivatives of $\mathbf{g}$ by $\mathbf{g}_{\mathbf{u}}$:

$$\mathbf{g}_{\mathbf{u}}(\mathbf{u}) = \left(\frac{\partial g^i(\mathbf{u})}{\partial u^j}\right).$$

Theorem 1 Let $\mathbf{f} = \mathbf{f}(t, \mathbf{u})$ be analytic in $t, \mathbf{u}$ for all t near 0 and all $\mathbf{u}$, and let $\boldsymbol{\phi} = \boldsymbol{\phi}(\mathbf{u}) = \mathbf{f}(0, \mathbf{u})$. Then the transformation scheme

$$\mathbf{v} = \mathbf{f}(t, \mathbf{u}), \tag{22}$$

$$\mathbf{u}_t + \hat{A}\mathbf{u}_x = \hat{B}\mathbf{u}, \tag{23}$$

is a linearization of

$$\mathbf{v}_t + A\mathbf{v}_x = \mathbf{v} * \mathbf{v} + B\mathbf{v} \tag{24}$$

if and only if

$$\mathbf{f}(\tau, \mathbf{w}) = \mathbf{h}[\tau, \boldsymbol{\phi}(e^{-\hat{B}\tau}\mathbf{w})], \tag{25}$$

where $\mathbf{h} = \mathbf{h}(t, \xi)$ is the solution of

$$\dot{\mathbf{h}} = \mathbf{h} * \mathbf{h} + B\mathbf{h}, \qquad \mathbf{h}(0, \xi) = \xi, \tag{26}$$

and for all τ near 0 and all $\mathbf{w}$, $\boldsymbol{\phi}$ satisfies

$$A\mathbf{h}_\xi[\tau, \boldsymbol{\phi}(\mathbf{w})]\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{w})e^{-\hat{B}\tau} = \mathbf{h}_\xi[\tau, \boldsymbol{\phi}(\mathbf{w})]\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{w})e^{-\hat{B}\tau}\hat{A}. \tag{27}$$

Proof There are two steps in the proof. The first step is to show that f is a linearization if and only if f satisfies the two equations

$$A\mathbf{f}_{\mathbf{u}}(\tau, \mathbf{w}) = \mathbf{f}_{\mathbf{u}}(\tau, \mathbf{w})\hat{\mathbf{A}}, \tag{28}$$

$$\mathbf{f}_t(\tau, \mathbf{w}) + \mathbf{f}_{\mathbf{u}}(\tau, \mathbf{w})\hat{B}\mathbf{w} = \mathbf{f}(\tau, \mathbf{w}) * \mathbf{f}(\tau, \mathbf{w}) + B\mathbf{f}(\tau, \mathbf{w}). \tag{29}$$

The second step is to solve (29) by the method of characteristics to obtain (25). Then (27) is obtained by substituting (25) into (28).

If we substitute (22) into (24) and use (23) to eliminate $\mathbf{u}_t$, we see that $\mathbf{f}$ is a linearization of (24) if and only if

$$[A\mathbf{f}_{\mathbf{u}}(t, \mathbf{u}) - \mathbf{f}_{\mathbf{u}}(t, \mathbf{u})\hat{A}]\mathbf{u}_x + \mathbf{f}_t(t, \mathbf{u}) + \mathbf{f}(t, \mathbf{u})\hat{B}\mathbf{u} = \mathbf{f}(t, \mathbf{u}) * \mathbf{f}(t, \mathbf{u}) + B\mathbf{f}(t, \mathbf{u}) \tag{30}$$

holds whenever $\mathbf{u}$ is a solution of (23).

We claim that (30) holds whenever $\mathbf{u}$ is a solution of (23) if and only if (28) and (29) hold for all $\mathbf{w}$ and all small τ. Clearly, (28) and (29) imply (30) and so we need only to establish their necessity.

Intuitively we expect that as $\mathbf{u}$ ranges over all solutions of (23) $\mathbf{u}$ and $\mathbf{u}_x$ will vary independently. Equation (30) can then hold only if the coefficient of $\mathbf{u}_x$ vanishes [(Eq. 28)] and the remaining terms form an equality [(Eq. 29)].

More precisely, we have that $\mathbf{u}$ and $\mathbf{u}_x$ are independent for fixed t,x: For each x_0, small τ, and arbitrary constant vectors $\mathbf{w},\mathbf{z}$, there exists a function $\hat{\mathbf{u}} = \hat{\mathbf{u}}(t,x)$ satisfying (23) and

$$\hat{\mathbf{u}}(\tau,x_0) = \mathbf{w}, \qquad \hat{\mathbf{u}}_x(\tau,x_0) = \mathbf{z}. \tag{31}$$

Suppose for the moment that such a $\hat{\mathbf{u}}$ exists. Then (30) must hold for $\mathbf{u} = \hat{\mathbf{u}}$; and if we make this substitution and set $t = \tau$ and $x = x_0$, (30) becomes

$$[A\mathbf{f}_{\mathbf{u}}(\tau,\mathbf{w}) - \mathbf{f}_{\mathbf{u}}(\tau,\mathbf{w})\hat{A}]\mathbf{z}$$
$$+ \mathbf{f}_t(\tau,\mathbf{w}) + \mathbf{f}_{\mathbf{u}}(\tau,\mathbf{w})\hat{B}\mathbf{w}$$
$$= \mathbf{f}(\tau,\mathbf{w}) * \mathbf{f}(\tau,\mathbf{w}) + B\mathbf{f}(\tau,\mathbf{w}).$$

This equation must hold for all $\mathbf{z}$ and $\mathbf{w}$, which implies (28) and (29).

To establish the existence of $\hat{\mathbf{u}}$, let $\bar{\boldsymbol{\psi}} = [\bar{\psi}^i(x)]$ be an analytic function such that

$$\bar{\boldsymbol{\psi}}(x_0) = \mathbf{w}, \qquad \bar{\boldsymbol{\psi}}_x(x_0) = \mathbf{z}, \tag{32}$$

and let $\bar{\mathbf{u}}$ be the solution of

$$-\bar{\mathbf{u}}_t + \hat{A}\bar{\mathbf{u}}_x = \hat{B}\bar{\mathbf{u}}, \qquad \bar{\mathbf{u}}(0,x) = \bar{\boldsymbol{\psi}}(x). \tag{33}$$

The Cauchy-Kowalewsky theorem guarantees the existence of such a $\bar{\mathbf{u}}$, and for x near x_0, $\bar{\mathbf{u}}$ exists for t near 0, e.g., for $0 \le t < 2\tau$.

Now let $\hat{\mathbf{u}}(t,x) = \bar{\mathbf{u}}(\tau - t,x)$. Equation (33) implies that $\hat{\mathbf{u}}$ is a solution of (23) and that $\hat{\mathbf{u}}(\tau,x) = \bar{\boldsymbol{\psi}}(x)$. Equation (31) follows directly from this equation and (32).

We have now established that $\mathbf{f}$ is a linearization if and only if it satisfies (28) and (29). The particular form of (29) allows us to solve it by the method of characteristics. The characteristic equations are

$$\dot{\tau} = 1, \qquad \tau(0) = 0; \qquad \Rightarrow \tau = s;$$
$$\dot{\mathbf{w}} = \hat{B}\mathbf{w}, \qquad \mathbf{w}(0) = \boldsymbol{\sigma}; \qquad \Rightarrow \mathbf{w} = e^{\hat{B}s}\boldsymbol{\sigma};$$
$$\dot{\mathbf{f}} = \mathbf{f} * \mathbf{f} + B\mathbf{f}, \qquad \mathbf{f})0) = \boldsymbol{\phi}(\boldsymbol{\sigma});$$
$$\Rightarrow \mathbf{f} = \mathbf{h}[t,\boldsymbol{\phi}(\boldsymbol{\sigma})].$$

Solving the first two of these for s and σ and substituting them into the third gives (25). Equation (27) follows from substituting (25) into (28) and replacing $\mathbf{w}\exp(-\hat{B}\tau)$ by $\mathbf{w}$. This procedure is allowed because (28) holds for all $\mathbf{w},t$. This completes the proof.

Theorem 1 provides the promised analytic criterion for the existence of a linearization: It exists if and only if there exist $\hat{A},\hat{B}$ and $\boldsymbol{\phi}$ such that (27) holds. This is a one-parameter family of *overdetermined* systems of partial differential equations for $\boldsymbol{\phi}$. In general, such a system has *no* solution, which justifies the earlier remark questioning the very existence of a linearization of the partial differential equation. When $A = 0$, however, (27) can be satisfied by taking $\hat{A} = 0$ and in this case Theorem 1 becomes Lemma 1.

The problem of identifying those systems (10) that can be linearized has now been reduced to finding a set of *algebraic* conditions on $A,B,\mathscr{A}$ which are the necessary and sufficient conditions for (27) to have a solution. A method of constructing $\boldsymbol{\phi}$ should also be provided.

One approach to this problem is to expand both sides of (27) in a Taylor series in $\mathbf{w}$. Equating coefficients of like powers of $\mathbf{w}$ then gives an infinite sequence of necessary algebraic conditions. The problem reduces to selecting a finite number of these conditions that are sufficient. Rather than pursue this highly involved approach to completion, we derive the first two such necessary conditions and a separate sufficient condition.

From now on we consider only complete linearizations, since only these solve all Cauchy problems and provide a single representation of the general solution. In this case we can assume without loss of generality that $\boldsymbol{\phi}(\mathbf{0}) = \mathbf{0}$: If $\mathbf{f}$ is complete, $\boldsymbol{\phi}$ is onto and has a root $\mathbf{r}$; (27) is invariant under $\mathbf{u}$-translations, so we can replace $\boldsymbol{\phi}(\mathbf{u})$ by $\boldsymbol{\phi}(\mathbf{u} + \mathbf{r})$.

Recall that the notation $\mathscr{G} = (g^i_{jk})$ means that $\mathscr{G}$ is the algebra with structure constants g^i_{jk} in the usual basis and that the notation $\mathscr{G} = \mathscr{G}(\#)$, $\mathscr{H} = \mathscr{H}(\wedge)$ indicates that the products in $\mathscr{G}$ and $\mathscr{H}$ are respectively denoted by $\mathbf{x} \# \mathbf{y}$, $\mathbf{x} \wedge \mathbf{y}$.

Introduce the following matrices and algebras:

$$G = (g^i_j) = \left(\frac{\partial\phi^i(\mathbf{0})}{\partial u^j}\right) = \phi_{\bar{u}}(\mathbf{0}), \tag{34}$$

$$\mathscr{G}(\#) = (g^i_{jk}) = \left(\frac{\partial^2\phi^i(\mathbf{0})}{\partial u^j \partial u^k}\right) = \phi_{\mathbf{u}\mathbf{u}}(\mathbf{0}), \tag{35}$$

$$H = [h^i_j(t)] = e^{Bt} = \mathbf{h}_\xi(t,\mathbf{0}), \tag{36}$$

$$\mathscr{H}(\wedge) = [h^i_{jk}(t)]$$
$$= \int_0^t h^i_a(-\tau)a^a_{bc}h^b_j(\tau)h^c_k(\tau)d\tau = \tfrac{1}{2}e^{-Bt}\mathbf{h}_{\xi\xi}(t,\mathbf{0}). \tag{37}$$

We must first establish the rightmost equalities claimed in (36) and (37). From (26) we have

$$\dot{\mathbf{h}}(t,\boldsymbol{\xi}) = \mathbf{h}(t,\boldsymbol{\xi}) * \mathbf{h}(t,\boldsymbol{\xi}) + B\mathbf{h}(t,\boldsymbol{\xi}), \qquad \mathbf{h}(0,\boldsymbol{\xi}) = \boldsymbol{\xi}.$$

Taking the $\boldsymbol{\xi}$-derivative of this and setting $\boldsymbol{\xi} = \mathbf{0}$ gives

$$\dot{\mathbf{h}}_\xi(t,\mathbf{0}) = B\mathbf{h}_\xi(t,\mathbf{0}), \qquad \mathbf{h}_\xi(0,\mathbf{0}) = I;$$

and the solution of this linear ordinary differential equation is (36).

Similarly, taking the second $\boldsymbol{\xi}$-derivative and setting $\boldsymbol{\xi} = \mathbf{0}$ gives

$$\dot{\mathbf{h}}_{\xi\xi}(t,\mathbf{0}) = 2\mathbf{h}_{\xi}(t,\mathbf{0}) * \mathbf{h}_{\xi}(t,\mathbf{0}) + B\mathbf{h}_{\xi\xi}(t,\mathbf{0}), \qquad \mathbf{h}_{\xi\xi}(0,\mathbf{0}) = \mathbf{0};$$

and the solution of this linear ordinary differential equation is

$$\mathbf{h}_{\xi\xi}(t,\mathbf{0}) = 2e^{Bt}\int_0^t e^{-B\tau}\mathbf{h}_{\xi}(\tau,\mathbf{0}) * \mathbf{h}_{\xi}(\tau,\mathbf{0})\,d\tau = 2e^{Bt}\int_0^t e^{-B\tau}(e^{B\tau} * e^{B\tau})\,d\tau.$$

This establishes (37).

The algebra $\mathscr{H}$ is time dependent and we can express the $\mathscr{H}$ multiplication in terms of the $\mathscr{A} = \mathscr{A}(*)$ multiplication directly:

$$\mathbf{x} \wedge \mathbf{y} = \int_0^t e^{-B\tau}[(e^{B\tau}\mathbf{x}) * (e^{B\tau}\mathbf{y})]\,d\tau. \tag{38}$$

From this we have $\mathbf{x} \wedge \mathbf{y} = \mathbf{0}$ when $t = 0$;

$$d/dt(\mathbf{x} \wedge \mathbf{y}) = e^{-Bt}[(e^{Bt}\mathbf{x}) * (e^{Bt}\mathbf{y})]; \tag{39}$$

and

$$d/dt(\mathbf{x} \wedge \mathbf{y})|_{t=0} = \mathbf{x} * \mathbf{y}.$$

If we utilize these notations and evaluate (27) and its $\mathbf{w}$-derivative at $\mathbf{w} = \mathbf{0}$, we obtain two necessary conditions:

Lemma 2 If $\boldsymbol{\phi} = \boldsymbol{\phi}(\mathbf{u})$ is an analytic solution of (27) near zero and $\boldsymbol{\phi}(\mathbf{0}) = \mathbf{0}$, then

$$e^{-Bt}Ae^{Bt}G = Ge^{-\hat{B}t}\hat{A}e^{\hat{B}t}, \tag{40}$$

$$e^{-Bt}Ae^{Bt}[2(G\mathbf{x}) \wedge (G\mathbf{y}) + \mathbf{x}\#\mathbf{y}] = 2(Ge^{-\hat{B}t}\hat{A}e^{\hat{B}t}\mathbf{x}) \wedge (G\mathbf{y}) + (e^{-\hat{B}t}\hat{A}e^{\hat{B}t}\mathbf{x}) \# \mathbf{y}. \tag{41}$$

Proof Setting $\mathbf{w} = \mathbf{0}$ and $\tau = t$ in (27) gives

$$A\mathbf{h}_{\xi}(t,\mathbf{0})\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{0})e^{-\hat{B}t} = \mathbf{h}_{\xi}(t,\mathbf{0})\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{0})e^{-\hat{B}t}\hat{A};$$

and substituting (34) and (36) into this gives

$$Ae^{Bt}Ge^{-\hat{B}t} = e^{Bt}Ge^{-\hat{B}t}\hat{A}.$$

This implies (40).

If we differentiate (27) with respect to $\mathbf{w}$ and set $\mathbf{w} = \mathbf{0}$ and $\tau = t$, we obtain

$$A[\mathbf{h}_{\xi\xi}(t,\mathbf{0})\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{0})(\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{0})e^{-\hat{B}t}) + \mathbf{h}_{\xi}(t,\mathbf{0})\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{0})e^{-\hat{B}t}] = [\mathbf{h}_{\xi\xi}(t,\mathbf{0})\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{0})(\boldsymbol{\phi}_{\mathbf{u}}(\mathbf{0})e^{-\hat{B}t}) + \mathbf{h}_{\xi}(t,\mathbf{0})\boldsymbol{\phi}_{\mathbf{uu}}(\mathbf{0})e^{-\hat{B}t}]\hat{A}.$$

Substituting (34) through (37) into this and expressing the resulting expression in component notation gives

$$a^i_p[2h^p_a(t)h^a_{bc}(t)\,g^b_k g^c_d \hat{h}^d_j(-t) + h^p_a(t)g^a_{kb}\hat{h}^b_j(-t)] = [2h^i_a(t)h^a_{bc}(t)g^b_k g^c_d\hat{h}^d_p(-t) + h^i_a(t)g^a_{kb}\hat{h}^b_p(-t)]\hat{a}^p_j, \tag{42}$$

where we have introduced $e^{\hat{B}t} = (\hat{h}^i_j(t))$.

After multiplying both sides of (42) by $x^j y^k$ (and summing over j,k) and converting back to vector notation, we see that (42) holds if and only if for all $\mathbf{x} = (x^i)$ and $\mathbf{y} = (y^i)$ we have

$$Ae^{Bt}[2(G\mathbf{y}) \wedge (Ge^{-\hat{B}t}\mathbf{x}) + \mathbf{y} \# (e^{-\hat{B}t}\mathbf{x})] = e^{Bt}[2(G\mathbf{y}) \wedge (Ge^{-\hat{B}t}\hat{A}\mathbf{x}) + \mathbf{y} \# (e^{-\hat{B}t}\hat{A}x)].$$

Multiplying both sides by e^{-Bt} and replacing $\mathbf{x}$ by $e^{\hat{B}t}\mathbf{x}$ gives (41), which completes the proof.

Neither (40) nor (41) is an algebraic condition because of the dependence on t. Each, however, implies the infinite sequence of algebraic conditions obtained by expanding in powers of t and equating coefficients.

For example, if we equate coefficients of the first two terms in (40) and (41) and introduce the usual commutator notation $[A,B] = AB - BA$, we obtain the four algebraic necessary conditions

$$AG = G\hat{A}, \tag{43}$$

$$[A,B]G = G[\hat{A},\hat{B}], \tag{44}$$

$$A(\mathbf{x} \# \mathbf{y}) = (\hat{A}\mathbf{x}) \# \mathbf{y}, \tag{45}$$

$$2A[(G\mathbf{x}) * (G\mathbf{y})] + [A,B](\mathbf{x} \# \mathbf{y}) = 2(G\hat{A}\mathbf{x}) * (G\mathbf{y}) + [\hat{A},\hat{B}]\mathbf{x} \# \mathbf{y}. \tag{46}$$

Of the two necessary conditions (40) and (41), the second is by far the more indirect and complex: $\mathscr{A},A,B$ have the property that there exist $\mathscr{G},\hat{A},\hat{B}$ such that (41) holds. We have not found conditions on $\mathscr{A},A,B$ which are both necessary and sufficient for (41) to hold. An obviously sufficient condition is

$$e^{-Bt}Ae^{Bt}[(G\mathbf{x}) \wedge (G\mathbf{y})] = (e^{-Bt}Ae^{Bt}G\mathbf{x}) \wedge (G\mathbf{y}). \tag{47}$$

This equation, (40), and

$$e^{-Bt}Ae^{Bt}(\mathbf{x} \# \mathbf{y}) = (e^{-\hat{B}t}\hat{A}e^{\hat{B}t}\mathbf{x}) \# \mathbf{y} \tag{48}$$

imply (41); and (48) is satisfied by the zero algebra: $\mathbf{x} \# \mathbf{y} = \mathbf{0}$. We can show that when G is nonsingular, (40) and (47) are also sufficient conditions for (27) to have a solution, i.e., that (10) has a linearization.

We first replace (47) by a simpler condition:

Lemma 4 Assume that G is nonsingular. Then if $\mathscr{A},A,B$ satisfy

$$e^{-Bt}Ae^{Bt}[(G\mathbf{x}) \wedge (G\mathbf{y})] = (e^{-Bt}Ae^{Bt}G\mathbf{x}) \wedge (G\mathbf{y}), \tag{49}$$

they also satisfy

$$e^{-Bt}Ae^{Bt}[(G\mathbf{x}) * (G\mathbf{y})] = (e^{-Bt}Ae^{Bt}G\mathbf{x}) * (G\mathbf{y}). \quad (50)$$

Proof If we differentiate (49) with respect to t and use (39), we obtain

$$e^{-Bt}[A,B]e^{Bt}[(G\mathbf{x}) \wedge (G\mathbf{y})] + e^{-Bt}A[(e^{Bt}G\mathbf{x}) * (e^{Bt}G\mathbf{y})]$$
$$= \{e^{-Bt}[A,B]e^{Bt}G\mathbf{x}\} \wedge (G\mathbf{y}) + e^{-Bt}[(Ae^{Bt}G\mathbf{x}) * (e^{Bt}G\mathbf{y})]. \quad (51)$$

Setting $t = 0$ in this gives $A[(G\mathbf{x}) * (G\mathbf{y})] = (AG\mathbf{x}) * (G\mathbf{y})$, and since G is nonsingular, this implies that

$$A(\mathbf{x} * \mathbf{y}) = (A\mathbf{x}) * \mathbf{y}. \quad (52)$$

Substituting (52) into (51) gives

$$e^{-Bt}[A,B]e^{Bt}[(G\mathbf{x}) \wedge (G\mathbf{y})] = \{e^{-Bt}[A,B]e^{Bt}G\mathbf{x}\} \wedge (G\mathbf{y}).$$

We express this as

$$e^{-Bt}ad(B)Ae^{Bt}[(G\mathbf{x}) \wedge (G\mathbf{y})]$$
$$= [e^{-Bt}ad(B)Ae^{Bt}G\mathbf{x}] \wedge (G\mathbf{y}), \quad (53)$$

where we have introduced the linear operator $ad(B)$ by $ad(B)A = [B,A]$.

Equation (53) shows that if A satisfies (49), so does $ad(B)A$. By induction it follows that

$$e^{-Bt}ad^n(B)Ae^{Bt}[(G\mathbf{x}) \wedge (G\mathbf{y})]$$
$$= [e^{-Bt}ad^n(B)Ae^{Bt}G\mathbf{x}] \wedge (G\mathbf{y}) \quad (54)$$

for all n.

Differentiating (54) with respect to t, setting $t = 0$, and using (39) implies that

$$ad^n(B)A[(G\mathbf{x}) * (G\mathbf{y})] = [ad^n(B)AG\mathbf{x}] * (G\mathbf{y}). \quad (55)$$

The Campbell-Baker-Hausdorf formula (see Hausner and Schwartz [11]) states that

$$e^{Bt}Ae^{-Bt} = e^{ad(Bt)}A. \quad (56)$$

If we multiply both sides of (55) by $t^n/n!$ and then sum the resulting equations from $n = 1$ to $n = \infty$, we obtain

$$e^{ad(Bt)}A[(G\mathbf{x}) * (G\mathbf{y})] = [e^{ad(Bt)}AG\mathbf{x}] * (G\mathbf{y}).$$

This and (56) imply (50), which completes the proof.

We can now provide sufficient conditions for the existence of a linearization:

Theorem 2 If $\mathscr{A},B$ satisfy

$$e^{-Bt}Ae^{Bt}(\mathbf{x} * \mathbf{y}) = (e^{-Bt}Ae^{Bt}\mathbf{x}) * \mathbf{y}, \quad (57)$$

then

$$\mathbf{v} = \mathbf{f}(t,\mathbf{u}) = \mathbf{h}(t,e^{-Bt}\mathbf{u}), \quad (58)$$

$$\mathbf{u}_t + A\mathbf{u}_x = B\mathbf{u} \quad (59)$$

is a complete linearization of $\mathbf{v}_t + A\mathbf{v}_x = \mathbf{v} * \mathbf{v} + B\mathbf{v}$.

Proof The transformation (58) is just (25) with $\boldsymbol{\phi}(\mathbf{u}) = \mathbf{u}$. According to Theorem 1, this transformation is a linearization if and only if $\boldsymbol{\phi}$ satisfies (27); and for this particular $\boldsymbol{\phi},\hat{A}$, and $\hat{B}$ (27) becomes

$$A\mathbf{h}_\xi(t,\mathbf{w})e^{-Bt} = \mathbf{h}_\xi(t,\mathbf{w})e^{-Bt}A. \quad (60)$$

We first show that (57) implies that

$$Ae^{-Bt}\mathbf{h}_\xi(t,\mathbf{w}) = e^{-Bt}\mathbf{h}_\xi(t,\mathbf{w})A, \quad (61)$$

and that this in turn implies (60).

To establish (61), we introduce

$$L(t) = Ae^{-Bt}\mathbf{h}_\xi(t,\mathbf{w}),$$
$$R(t) = e^{-Bt}\mathbf{h}_\xi(t,\mathbf{w})A.$$

Then

$$\dot{L} = Ae^{-Bt}(\dot{\mathbf{h}}_\xi - B\mathbf{h}_\xi) = Ae^{-Bt}(2\mathbf{h} * \mathbf{h}_\xi)$$
$$= e^{-Bt}[2\mathbf{h} * (e^{Bt}Ae^{-Bt}\mathbf{h}_\xi)],$$
$$\Rightarrow \dot{L} = 2e^{-Bt}[\mathbf{h} * (e^{Bt}L)], \qquad L(0) = A. \quad (62)$$

Similarly,

$$\dot{R} = e^{-Bt}(\dot{\mathbf{h}}_\xi A - B\mathbf{h}_\xi A),$$
$$= e^{-Bt}[2\mathbf{h} * (\mathbf{h}_\xi A)] = 2e^{-Bt}[\mathbf{h} * (e^{Bt}e^{-Bt}\mathbf{h}_\xi A)],$$
$$\Rightarrow \dot{R} = 2e^{-Bt}[\mathbf{h} * (e^{Bt}R)], \qquad R(0) = A. \quad (63)$$

Here we have used (57) and (26). Equations (62) and (63) show that L and R satisfy the same linear ordinary differential equation and the same initial condition. Therefore $R = L$ because of the uniqueness theorem for the initial value problem for ordinary differential equations. This shows that (57) implies (61).

To show that (61) implies (60), we first observe that since $\mathbf{h}$ is the solution of $\mathbf{h} = \mathbf{h} * \mathbf{h} + B\mathbf{h}$, $\mathbf{h}(0,\boldsymbol{\xi}) = \boldsymbol{\xi}$, we have $\mathbf{h}(t + \tau,\boldsymbol{\xi}) = \mathbf{h}[t,\mathbf{h}(\tau,\boldsymbol{\xi})]$.

Taking $\tau = -t$ gives

$$\mathbf{h}[t,\mathbf{h}(-t,\boldsymbol{\xi})] = \boldsymbol{\xi},$$
$$\Rightarrow \mathbf{h}_\xi[t,\mathbf{h}(-t,\boldsymbol{\xi})]\mathbf{h}_\xi(-t,\boldsymbol{\xi}) = I,$$
$$\Rightarrow \mathbf{h}_\xi^{-1}(-t,\boldsymbol{\xi}) = \mathbf{h}_\xi[t,\mathbf{h}(-t,\boldsymbol{\xi})]. \quad (64)$$

Now, since (61) holds for all $t,\mathbf{w}$, it holds when we replace $\mathbf{w}$ by $\mathbf{h}(-t,\mathbf{w})$, giving

$$Ae^{-Bt}\mathbf{h}_\xi[t,\mathbf{h}(-t,\mathbf{w})] = e^{-Bt}\mathbf{h}_\xi[t,\mathbf{h}(-t,\mathbf{w})]A.$$

Substituting (64) into this equation gives

$$Ae^{-Bt}\mathbf{h}_\xi^{-1}(-t,\mathbf{w}) = e^{-Bt}\mathbf{h}_\xi^{-1}(-t,\mathbf{w})A,$$
$$\Rightarrow \mathbf{h}_\xi(-t,\mathbf{w})e^{Bt}A = A\mathbf{h}_\xi(-t,\mathbf{w})e^{Bt}.$$

The last equation is (60), which was a sufficient condition for (58), (59) to be a linearization. The linearization is complete because $\phi(\mathbf{u})$ is invertible. This proves the theorem.

The condition (57) would be an associative law if the operation $*$ and matrix multiplication were the same. For this reason, we say that (10) is A-associative if (57) holds. If A and B commute, (57) simplifies to a t-independent relation. In this case we can show that (57) is necessary as well.

Theorem 3 Assume that A and B commute: $[A,B] = AB - BA = 0$. Then a linearization with $\phi(\mathbf{0}) = \mathbf{0}$ and $\phi_{\mathbf{u}}(\mathbf{0})$ nonsingular exists if and only if

$$A(\mathbf{x} * \mathbf{y}) = (A\mathbf{x}) * \mathbf{y}. \tag{65}$$

Proof When $[A,B] = 0$, (65) is the same as (57) and is therefore a sufficient condition for the existence of a linearization.

If $[A,B] = 0$, and $G = \phi_{\mathbf{u}}(\mathbf{0})$ is nonsingular, then (44) implies $[\hat{A},\hat{B}] = 0$; and (46) then becomes $A[(G\mathbf{x}) * (G\mathbf{y})] = (AG\mathbf{x}) * (G\mathbf{y})$. This equation is equivalent to (65) when G is nonsingular.

With Theorem 2 we have identified a class of systems that can be completely linearized—the A-associative systems. These systems have an attractive property: Solving the Cauchy problem for them reduces to solving a system of nonlinear *ordinary* differential equations (26) and to solving a system of *linear* partial differential equations. This procedure is a great reduction in complexity, and we expect to be able to carry the qualitative theory of such equations quite far.

A-associative systems and the center of $\mathscr{A}$.
We recall that the system

$$\mathbf{v}_t + A\mathbf{v}_x = \mathbf{v} * \mathbf{v} + B\mathbf{v} \tag{10}$$

is A-associative if and only if

$$e^{-Bt}Ae^{Bt}(\mathbf{x} * \mathbf{y}) = (e^{-Bt}Ae^{Bt}\mathbf{x}) * \mathbf{y}. \tag{66}$$

Such systems are completely linearized by

$$\mathbf{v} = \mathbf{h}(t, e^{-Bt}\mathbf{u}), \tag{67}$$

$$\mathbf{u}_t + A\mathbf{u}_x = B\mathbf{u}, \tag{68}$$

where $h = h(t,\xi)$ is the solution of

$$\mathbf{h} = \mathbf{h} * \mathbf{h} + B\mathbf{h}, \qquad \mathbf{h}(0,\xi) = \xi. \tag{69}$$

The solution of (69) depends only on $\mathscr{A}, B$ and once we know this solution we can solve the set of systems (10) defined by (66). This leads naturally to the following problem: Given $\mathscr{A}$, find all A,B that satisfy (66). We shall express the solution of this problem in terms of an important structure of $\mathscr{A}$, i.e., the center of $\mathscr{A}$.

We present the necessary algebraic material before stating the theorem. For details see Schafer [6]. An algebra $\mathscr{G}$ is *unital* if it contains a *unity*: an element $\mathbf{u}$ such that $\mathbf{ux} = \mathbf{xu} = \mathbf{x}$ for all $\mathbf{x} \in \mathscr{G}$.

Let $\mathbf{x}$ be any element in the commutative algebra $\mathscr{G}$. The multiplication operator $M_\mathbf{x} = M(\mathbf{x})$ determined by $\mathbf{x}$ is the linear operator $M_\mathbf{x}: \mathbf{y} \to \mathbf{x} * \mathbf{y}, \quad \mathbf{y} \in \mathscr{G}$. If $\mathscr{S}$ is any subset of $\mathscr{G}$, $\mathscr{S}*$ is the set of all multiplication operators of elements in $\mathscr{S}$, i.e., $\mathscr{S} = \{M_\mathbf{x}: \mathbf{x} \in \mathscr{S}\}$.

If $\mathscr{G}$ is a commutative algebra, the center of $\mathscr{G}$, $\mathscr{Z} = \mathscr{Z}(\mathscr{G})$, is the set of those elements that associate with all elements, i.e., $\mathscr{Z}(\mathscr{G}) = [\mathbf{z} \in \mathscr{G}: \mathbf{z}(\mathbf{xy}) = (\mathbf{zx})\mathbf{y}, \forall \mathbf{x},\mathbf{y} \in \mathscr{G}]$.

We can now express the condition for A-associativity in terms of the center.

Theorem 4 Assume that $\mathscr{A}$ is unital. Then the system $\mathbf{v}_t + A\mathbf{v}_x = \mathbf{v} * \mathbf{v} + B\mathbf{v}$ is A-associative if and only if $e^{-Bt}Ae^{Bt}$ is the multiplication operator of a central element of $\mathscr{A}$:

$$e^{-Bt}Ae^{Bt} \in \mathscr{Z}(\mathscr{A})^*. \tag{70}$$

Proof Assume that (10) is A-associative and that $\mathscr{A}$ is unital. Then (66) most hold and, with $\mathbf{x} = \mathbf{u}$, the unity of $\mathscr{A}$, gives $(e^{-Bt}Ae^{Bt})\mathbf{y} = (e^{-Bt}Ae^{Bt}\mathbf{u}) * \mathbf{y}$. In terms of multiplication operators this is

$$e^{-Bt}Ae^{Bt} = M_\mathbf{z}, \tag{71}$$

where $\mathbf{z} = e^{-Bt}Ae^{Bt}\mathbf{u}$.

Substituting this into (66) gives

$$\mathbf{z} * (\mathbf{x} * \mathbf{y}) = (\mathbf{z} * \mathbf{x}) * \mathbf{y} \tag{72}$$

and this is the condition that $\mathbf{z} \in \mathscr{Z}(\mathscr{A})$. Therefore (70) holds. Conversely, (71) and (72) imply (66), which proves the theorem.

This theorem is an example of the interplay between analytic properties of a quadratic differential equation and algebraic properties of its related algebra. The center is an important structure in the study of algebras and the theorem transfers properties of the center to properties of the differential system.

We briefly mention an example of such a transfer. We recall that $\mathscr{J}$ is an *ideal* of an algebra $\mathscr{G}$ if $\mathscr{J}$ is a linear subspace of $\mathscr{G}$ and $\mathbf{j} \in \mathscr{J}, \quad \mathbf{g} \in \mathscr{G}$ imply that $\mathbf{jg}$ and $\mathbf{gj}$ are contained in $\mathscr{J}$. An ideal is *proper* if it is not $\mathscr{G}$ or $\{\mathbf{0}\}$. An algebra is *simple* if it has no proper ideals.

It is easy to show that $\mathscr{A}$ contains an (n-p)-dimensional ideal if and only if the quadratic part of (10) can be weakly decoupled, i.e., if in some coordinate system the quadratic part of the first p equations involves only the first p variables. Therefore there is no weak decoupling of the quadratic part if and only if $\mathscr{A}$ is simple. But, as shown by Schafer [12], if $\mathscr{A}$ is unital and simple $\mathscr{Z}(\mathscr{A})$

is a field and so then is $\mathscr{X}^*(\mathscr{A})$. This, with Theorem 4, gives the following curious lemma.

Lemma 3 If the system (10) has no weak decoupling of the quadratic part, the set of matrices $e^{-Bt}Ae^{Bt}$ satisfying (66) then forms a matrix field.

A-associative wave equations

By common consent the partial differential equation

$$\mathbf{v}_t + A\mathbf{v}_x = \mathbf{v} * \mathbf{v} + B\mathbf{v} \qquad (10)$$

is termed a wave equation when it is hyperbolic, i.e., when the matrix has only real eigenvalues and a complete set of eigenvectors.

Assume that (10) is hyperbolic and let λ_p, $p = 1, \cdots, m$, be the distinct eigenvalues of A and $\mathscr{E}_p$ the corresponding invariant subspaces, i.e., $\mathscr{E}_p = \{\mathbf{x}{:}A\mathbf{x} = \lambda_p\mathbf{x}\}$. Denote the image of $\mathscr{E}_p$ under the transformation e^{-Bt} by $e^{-Bt}\mathscr{E}_p$, i.e., $e^{-Bt}\mathscr{E}_p = \{e^{-Bt}\mathbf{x}{:}\mathbf{x} \in \mathscr{E}_p\}$.

The algebraic relationship between hyperbolicity and A-associativity is characterized in the following lemma.

Lemma 4 Assume that (10) is hyperbolic. Then it is A-associative if and only if the spaces $e^{-Bt}\mathscr{E}_p$ are ideals of $\mathscr{A}$ for all $p = 1, \cdots, m$ and all t.

Proof Assume that (10) is A-associative, so that

$$e^{-Bt}Ae^{Bt}(\mathbf{x} * \mathbf{y}) = (e^{-Bt}Ae^{Bt}\mathbf{x}) * \mathbf{y}. \qquad (73)$$

Since A and $e^{-Bt}Ae^{Bt}$ are similar, they have the same eigenvalues λ_p. Also, if $\mathbf{x}_p \in \mathscr{E}_p$, then

$$e^{-Bt}Ae^{Bt}(e^{-Bt}\mathbf{x}_p) = e^{-Bt}A\mathbf{x}_p = \lambda_p e^{-Bt}\mathbf{x}_p,$$

which implies that the corresponding invariant subspaces of $e^{-Bt}Ae^{Bt}$ are $e^{-Bt}\mathscr{E}_p$:

$$e^{-Bt}\mathscr{E}_p = \{\mathbf{x}{:}e^{-Bt}Ae^{Bt}\mathbf{x} = \lambda_p\mathbf{x}\}. \qquad (74)$$

Let $\mathbf{x}_p \in e^{-Bt}\mathscr{E}_p$, $\mathbf{y} \in \mathscr{A}$. Then, by (73), (74),

$$e^{-Bt}Ae^{Bt}(\mathbf{x}_p * \mathbf{y}) = (e^{-Bt}Ae^{Bt}\mathbf{x}_p) * \mathbf{y} = (\lambda_p\mathbf{x}_p) * \mathbf{y} = \lambda_p(\mathbf{x}_p * \mathbf{y}),$$

which implies that $\mathbf{x}_p * \mathbf{y} \in e^{-Bt}\mathscr{E}_p$. Therefore $e^{-Bt}\mathscr{E}_p$ is an ideal of $\mathscr{A}$.

Conversely, assume that the $e^{-Bt}\mathscr{E}_p$ are ideals. If $\mathbf{x}_p \in e^{-Bt}\mathscr{E}_p$ and $\mathbf{y} \in \mathscr{A}$, then $\mathbf{x}_p * \mathbf{y} \in e^{-Bt}\mathscr{E}_p$ and, by (74),

$$e^{-Bt}Ae^{Bt}(\mathbf{x}_p * \mathbf{y}) = \lambda_p(\mathbf{x}_p * \mathbf{y}).$$

Also, by (74),

$$(e^{-Bt}Ae^{Bt}\mathbf{x}_p) * \mathbf{y} = (\lambda_p\mathbf{x}_p) * \mathbf{y} = \lambda_p(\mathbf{x}_p * \mathbf{y}),$$

and these together imply

$$e^{-Bt}Ae^{Bt}(\mathbf{x}_p * \mathbf{y}) = (e^{-Bt}Ae^{Bt}\mathbf{x}_p) * \mathbf{y}.$$

when $\mathbf{y} = \mathbf{y}_p \in e^{-Bt}\mathscr{E}_p$, this implies in particular that

$$e^{-Bt}Ae^{Bt}(\mathbf{x}_p * \mathbf{y}_p) = (e^{-Bt}Ae^{Bt}\mathbf{x}_p) * \mathbf{y}_p, \qquad \mathbf{x}_p, \mathbf{y}_p \in e^{-Bt}\mathscr{E}_p. \qquad (75)$$

We claim that (75) implies (73). Since A has a complete set of eigenvectors, the invariant subspaces (74) decompose the space $\mathfrak{R}^n$ into a direct sum:

$$\mathfrak{R}^n = e^{-Bt}\mathscr{E}_1 \oplus \cdots \oplus e^{-Bt}\mathscr{E}_m.$$

This means that the subspaces are disjoint (except for 0) and that every $\mathbf{x} \in \mathfrak{R}^n$, i.e., every $\mathbf{x} \in \mathscr{A}$, has a unique representation

$$\mathbf{x} = \mathbf{x}_1 + \cdots + \mathbf{x}_m, \qquad \mathbf{x}_p \in e^{-Bt}\mathscr{E}_p. \qquad (76)$$

Since the $e^{-Bt}\mathscr{E}_p$ are ideals of $\mathscr{A}$, we also have

$$\mathbf{x} * \mathbf{y} = \mathbf{x}_1 * \mathbf{y}_1 + \cdots + \mathbf{x}_m * \mathbf{y}_m. \qquad (77)$$

This follows from the fact that $\mathbf{j} * \mathbf{k} = 0$ if $\mathbf{j}$ and $\mathbf{k}$ lie in disjoint ideals. For if $\mathscr{J}, \mathscr{K}$ are ideals, $\mathbf{j} \in \mathscr{J}$ and $\mathbf{k} \in \mathscr{K}$ imply that $\mathbf{j} * \mathbf{k}$ lies in both $\mathscr{J}$ and $\mathscr{K}$. If $\mathscr{J}$ and $\mathscr{K}$ are disjoint, $\mathscr{J} \cap \mathscr{K} = \{\mathbf{0}\}$, so $\mathbf{j} * \mathbf{k} = \mathbf{0}$.

If we use (76), (77), and (74) in (73), we obtain an expression that is equivalent to (73):

$$e^{-Bt}Ae^{Bt}(\mathbf{x}_1 * \mathbf{y}_1 + \cdots + \mathbf{x}_m * \mathbf{y}_m). = (e^{-Bt}Ae^{Bt}\mathbf{x}_1) * \mathbf{y}_1 + \cdots + (e^{-Bt}Ae^{Bt}\mathbf{x}_m) * \mathbf{y}_m.$$

This equation follows directly from (75), which proves that (10) is A-associative and proves the lemma.

We now give two applications of this lemma.

Theorem 5 Assume that the system

$$\mathbf{v}_t + A\mathbf{v}_x = \mathbf{v} * \mathbf{v} + B\mathbf{v} \qquad (78)$$

is hyperbolic and that A and B commute. Then (78) has a linearization of the form $\mathbf{v} = \mathbf{f}(t,\mathbf{u})$, $\mathbf{u}_t + \hat{A}\mathbf{u}_x = \hat{B}\mathbf{u}$, with $\mathbf{f}(0,\mathbf{0}) = \mathbf{0}$ and $\mathbf{f}_\mathbf{u}(0,\mathbf{0})$ nonsingular if and only if the $\mathscr{E}_p$ decouple $(*)$. That is, in the unique representation $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_m$, $\mathbf{v}_p \in \mathscr{E}_p$, $(*)$ becomes

$$(\partial\mathbf{v}_p/\partial t) + (\lambda_p\partial\mathbf{v}_p/\partial x) = \mathbf{v}_p * \mathbf{v}_p + B_p\mathbf{y}_p, \qquad p = 1, \cdots, m, \qquad (79)$$

where $\mathbf{v}_p * \mathbf{v}_p \in \mathscr{E}_p$ and B_p is the restriction of B to $\mathscr{E}_p$, $B_p\mathbf{v}_p \in \mathscr{E}_p$.

Proof A has a complete set of eigenvectors, and so $\mathscr{A}$ is the direct sum of the $\mathscr{E}_p$, i.e., $\mathscr{A} = \mathscr{E}_1 \oplus \cdots \oplus \mathscr{E}_m$.

This means that every $\mathbf{v} \in \mathscr{A}$ has a unique direct sum representation $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_m$, $\mathbf{v}_p \in \mathscr{E}_p$. For any $\mathbf{x} \in \mathscr{A}$, denote the projection of $\mathbf{x}$ onto $\mathscr{E}_p$ by $(\mathbf{x})_p$, i.e., $(\mathbf{x})_p$ is the $\mathscr{E}_p$ component of $\mathbf{x}$ in the direct sum representation. In this notation (78) is equivalent to

$$(\mathbf{v}_t)_p + (A\mathbf{v}_x)_p = (\mathbf{v} * \mathbf{v})_p + (B\mathbf{v})_p, \qquad p = 1, \cdots, m.$$

Now $\mathbf{v} = \mathbf{v}_1 + \cdots + \mathbf{v}_m$; from (74) $A\mathbf{v}_p = \lambda_p \mathbf{v}_p$; and also from (74) $B\mathbf{v}_p \in \mathscr{E}_p$ since $AB = BA$. Using this in the above equation, we see that (78) is equivalent to

$$(\partial \mathbf{v}_p/\partial t) + (\lambda_p\, \partial \mathbf{v}_p/\partial x) = (\mathbf{v} * \mathbf{v})_p + B\mathbf{v}_p, \qquad p = 1, \cdots, m. \qquad (80)$$

By Theorem 3 the required linearization exists if and only if $(*)$ is A-associative; by Lemma 4 $(*)$ is A-associative if and only if the $\mathscr{E}_p$ are ideals of $\mathscr{A}$. Therefore we must show that the $\mathscr{E}_p$ decouple (81) into (79) if and only if the $\mathscr{E}_p$ are ideals.

First assume that the $\mathscr{E}_p$ are ideals. Then, by (77),

$$\begin{aligned}(\mathbf{v} * \mathbf{v})_p &= [(\mathbf{v}_1 + \cdots + \mathbf{v}_m) * (\mathbf{v}_1 + \cdots + \mathbf{v}_m)]_p \\ &= (\mathbf{v}_1 * \mathbf{v}_1 + \cdots + \mathbf{v}_m * \mathbf{v}_m)_p = \mathbf{v}_p * \mathbf{v}_p.\end{aligned}$$

Substitution of this into (80) gives (79), where B_p is the restriction of B to $\mathscr{E}_p$.

Conversely, assume that the $\mathscr{E}_p$ decouple (80) into (79). Subtracting these two equations implies that for all $\mathbf{v} \in \mathscr{A}$, $(\mathbf{v} * \mathbf{v})_p = \mathbf{v}_p * \mathbf{v}_p$. Replacing $\mathbf{v}$ by $\mathbf{v} + \mathbf{w}$ gives

$$\begin{aligned}&[(\mathbf{v}+\mathbf{w}) * (\mathbf{v}+\mathbf{w})]_p = (\mathbf{v}_p + \mathbf{w}_p) * (\mathbf{v}_p + \mathbf{w}_p), \\ \Rightarrow\ &(\mathbf{v} * \mathbf{v})_p + 2(\mathbf{v} * \mathbf{w})_p + (\mathbf{w} * \mathbf{w})_p = \mathbf{v}_p * \mathbf{v}_p \\ &\quad + 2\mathbf{v}_p * \mathbf{w}_p + \mathbf{w}_p * \mathbf{w}_p, \\ \Rightarrow\ &(\mathbf{v} * \mathbf{w})_p = \mathbf{v}_p * \mathbf{w}_p \in \mathscr{E}_p, \\ \Rightarrow\ &\mathbf{v} * \mathbf{w} = \mathbf{v}_1 * \mathbf{w}_1 + \cdots + \mathbf{v}_m * \mathbf{w}_m.\end{aligned}$$

If we take $\mathbf{v} = \mathbf{v}_p \in \mathscr{E}_p$, this implies $\mathbf{v}_p * \mathbf{w} = \mathbf{v}_p * \mathbf{w}_p \in \mathscr{E}_p$. Therefore $\mathscr{E}_p$ is an ideal. This completes the proof.

This result is somewhat disappointing, because the systems (80) are really ordinary differential equations. Thus the linearization process does not enlarge the class of solvable hyperbolic systems when $AB = BA$.

Corollary 1 The equations of nonlinear optics,

$$\left.\begin{aligned} v_t^1 + k^1 v_x^1 &= a^1 v^2 v^3 + b^1 v^1 \\ v_t^2 + k^2 v_x^2 &= a^2 v^1 v^3 + b^2 v^2 \\ v_t^3 + k^3 v_x^3 &= a^3 v^1 v^2 + b^3 v^3 \end{aligned}\right\}, \qquad (81)$$

can be linearized as required in Theorem 5 if and only if $k^1 = k^2 = k^3$.

Proof Since A and B in (81) are diagonal matrices they commute. Also, the system is hyperbolic. The system is already decoupled with respect to A and B and, by Theorem 5, the quadratic part must also be decoupled. Each single equation involves variables from the remaining two, and each pair of equations involves the variable from the remaining one. Therefore there is no decoupling into three single equations or into one single equation and one pair. Thus $m = 1$ and $\mathscr{E}_1$ is three-dimensional, so $k^1 = k^2 = k^3$.

Both of Theorems 3 and 5 indicate the need of relaxing the requirements on linearizations to obtain larger classes of solvable systems.

Acknowledgments

The author gratefully acknowledges the very helpful discussions of this material with Peter Lax and Farouk Odeh.

References

1. G. L. Lamb, "Analytical Descriptions of Ultrashort Optical Pulse Propagation in a Resonant Medium," *Rev. Mod. Phys.* **43**, 99 (1971).
2. N. Bloembergen, *Nonlinear Optics*, W. A. Benjamin, Inc., New York, 1965.
3. P. Lax, "Nonlinear Partial Differential Equations of Evolution," *Proceedings of the International Congress of Mathematics*, Nice, 1970, pp. 831–840.
4. E. Montroll, "On Nonlinear Processes Involving Population Growth and Diffusion," *J. Appl. Prob.* **4**, 281 (1967).
5. L. Marcus, "Quadratic Differential Equations and Non-associative Algebras," *Ann. Math. Studies*, **45**, 185 (1960), Princeton University Press, Princeton, N.J.
6. R. D. Schafer, *An Introduction to Nonassociative Algebras*, Academic Press, Inc., New York, 1966.
7. R. Courant and D. Hilbert, *Methods of Mathematical Physics, Volume II*, Interscience Publishers, New York, 1962, pp. 39–56.
8. *Ibid*, pp. 62–131.
9. G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, Cambridge, England, 1966, pp. 1–4, 85–94.
10. P. D. Gerber, "Left Alternative Algebras and Quadratic Differential Equations I," Research Report RC-4011, IBM Thomas J. Watson Research Center, Yorktown Heights, New York, 1972.
11. M. Hausner and J. Schwartz, *Lie Groups; Lie Algebras*, Gordon and Breach, Science Publishers, New York, 1968.
12. Schafer, *op. cit.*, pp. 16–17.

Received January 17, 1973

The author is located at the IBM Thomas J. Watson Research Center, Yorktown Heights, New York, 10598.