SIAM J. Sci. Comput.
Vol. 24, No. 3, pp. 905–923

INDEX AND CHARACTERISTIC ANALYSIS OF LINEAR PDAE SYSTEMS*

WADE S. MARTINSON† AND PAUL I. BARTON‡

Abstract. Automatable analyses of the index and well-posedness of two general classes of linear PDAEs are presented for use in equation-based dynamic simulation packages. These analyses determine whether a problem possesses a unique solution trajectory, the potential effect of noise in the initial and boundary conditions on the calculated solution, and the index of the model equations. Measures of the degeneracy of the system, rather than a differentiation index, bound the order of derivatives of forcing functions that appear in the solution.

Key words. index, characteristics, well-posedness, partial differential equations, partial differential-algebraic equations, differentiation index, perturbation index

AMS subject classifications. 35G99, 35R99

PII. S1064827599363411

1. Introduction.

Efficient and productive dynamic simulation of large-scale systems, such as chemical process flowsheets, often relies on equation-based simulation packages. These packages take as input the differential and algebraic model equations that describe the physical behaviors of the process, processing tasks, and discrete events, and provide tools such as model inheritance and decomposition to construct the large-scale model. They perform several analyses automatically and perform simulation-based calculations without intervention from the engineer. By reducing the amount and level of numerical expertise required to set up, debug, and perform these calculations, a modern simulation package is often the critical element that makes such simulation-based activities economically feasible in industrial settings [8, 14].

The ultimate purpose of equation-based simulation packages is to isolate the engineer from the numerics and to simply return the solution, together with guarantees of its accuracy. In order to accomplish this goal when some unit models include partial differential equations (PDEs), a simulation package should be able to identify models that it cannot solve numerically. Ideally, it would be able to analyze the equations and auxiliary conditions provided by the engineer, determine whether a unique solution trajectory even exists, and determine whether an accurate solution may be obtained by a time evolution method.

An evolution problem, which is considered to be the dynamic model equations together with initial and boundary data, is said to be well-posed if it possesses a unique solution that depends continuously on its data. If the problem is not well-posed because a unique solution does not exist, the meaning of results produced by any discretization technique or solution method is not clear. If the solution is not continuously dependent on its data, the solution computed using typical finite difference or finite element methods will be very sensitive to the particular mesh

*Received by the editors October 29, 1999; accepted for publication (in revised form) April 19, 2002; published electronically November 22, 2002.

http://www.siam.org/journals/sisc/24-3/36341.html

†Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave. 66-464, Cambridge, MA 02139. Current affiliation: Cargill Central Research, Cargill Inc., P.O. Box 5699, Minneapolis, MN 55440-5699 (wade_martinson@cargill.com).

‡Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave. 66-464, Cambridge, MA 02139 (pib@mit.edu).

spacing employed as well as to the roundoff error [22].

If the governing PDE has a differentiation index [17] of 2 or higher with respect to the evolution variable, one expects a high-index differential-algebraic equation (DAE) to result from any reasonable method of lines semidiscretization of the original problem. The problems associated with numerical solution of high-index DAEs are well known [1, 19, 21]. Existence of a strong solution will be shown to depend on sufficient differentiability of the forcing functions, just as for DAEs. The connection between the index and smoothness requirements on the forcing functions and data will be examined. It will be shown that the smoothness required is determined by the degeneracy of the system, rather than by a differentiation index.

In the systems context, analyses that may be performed by a computer have significant value. The analyses presented in this paper may be performed automatically using only routines that parse equations, differentiate algebraic equations, and calculate the generalized eigenvalues of a matrix pair. The paper concludes with two examples.

2. Background.

The analyses presented in this paper draw their inspiration from two sources. The first is index analysis of DAE systems and extensions to partial differential-algebraic equation (PDAE) systems. The second is the classical characteristic analysis of hyperbolic PDEs.

2.1. Basic terms and definitions.

A pair of matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ for which there exists some scalar $\lambda \in \mathbb{C}$ such that $|\mathbf{B} - \lambda\mathbf{A}| \neq 0$ is called *regular*. A one-parameter family of matrices of the form $\mathbf{B} - \lambda\mathbf{A}$ is referred to as a *pencil* of matrices; the pencil is said to be *regular* if it has at least one invertible member. The parameter λ is sometimes given as a ratio of two scalars ρ/τ so that the pencil is given by $\tau\mathbf{B} - \rho\mathbf{A}$. A particular value of λ that satisfies $|\mathbf{B} - \lambda\mathbf{A}| = 0$, or of ρ, τ that satisfies $|\tau\mathbf{B} - \rho\mathbf{A}| = 0$, is a *generalized eigenvalue* of the matrix pair [5].

For every pair of matrices $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$ that forms a regular pencil, there exist invertible matrices $\mathbf{P}$, $\mathbf{Q} \in \mathbb{C}^{n\times n}$ that transform the pair to its Weierstrass canonical form

$$\mathbf{PAQ} = \begin{bmatrix} \mathbf{I} & \\ & \mathbf{N} \end{bmatrix}, \qquad \mathbf{PBQ} = \begin{bmatrix} \mathbf{J} & \\ & \mathbf{I} \end{bmatrix},$$

where $\mathbf{J}$ is a lower Jordan matrix and $\mathbf{N}$ is a lower Jordan matrix of nilpotency k [5]. A matrix $\mathbf{N}$ such that $\mathbf{N}^{k-1} \neq \mathbf{0}$ and $\mathbf{N}^k = \mathbf{0}$ is said to be *nilpotent*, or to have *nilpotency* k.

Let the *degeneracy* of a generalized eigenvalue be defined as one less than the dimension of the associated Jordan block in the Weierstrass canonical form; the degeneracy of a block will refer to the same quantity. The term degeneracy will be used to refer to both the degeneracy of a block and of the associated eigenvalue.

A function of one independent variable $f(x)$ is said to have C^i continuity, or simply to be C^i, if its ith derivative is everywhere continuous. A function of multiple independent variables $g(x, y)$ has C^i continuity in x if its ith partial derivative with respect to x is everywhere continuous.

The terminology used to discuss solution existence, uniqueness, and differentiability for DAEs and PDEs varies. Let a *problem* be defined as a system of governing equations together with auxiliary conditions (initial conditions for DAEs, initial and boundary conditions for PDEs) and the domain over which the equations hold. A DAE is said to be *solvable* if a unique solution exists and is C^1. A PDE problem is said to be *well-posed* if there exists a unique solution that depends continuously on its data. A problem that is not well-posed is called *ill-posed*. A *strong* solution to a PDE system satisfies the governing equations pointwise everywhere.

Auxiliary conditions for a PDE may be used to determine a unique solution in two basic ways. The solution may be constructed from the functional form of the auxiliary conditions. Alternatively, a solution may be built from a family of basis functions, and some auxiliary conditions may be used to restrict eligible members of that family of basis functions. Let an *unrestricted solution* be a solution built strictly from the functional forms of the auxiliary conditions so that all conditions are considered to be data. Let a *restricted solution* be a solution constructed by using only some of the auxiliary conditions as data. In this paper, auxiliary conditions will be assumed to be linear and of Dirichlet type unless otherwise noted.

2.2. DAE systems and index. A *linear time-invariant* DAE is a system of the form

$$\mathbf{A}\dot{\mathbf{u}}(t) + \mathbf{B}\mathbf{u}(t) = \mathbf{f}(t). \tag{2.1}$$

Here $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n\times n}$, $T = [0, t_f]$, $\mathbf{u}, \mathbf{f} : T \to \mathbb{R}^n$, and $\dot{\mathbf{u}} \equiv \frac{d\mathbf{u}}{dt}$.

The *differentiation index* of the system is denoted by ν and is defined [6] as the minimum number of times that some or all of the equations must be differentiated in order to determine $\dot{\mathbf{u}}$ as a continuous function of $\mathbf{u}$ and t [1]. Other indices have been defined for DAE systems, including the perturbation [10] and strangeness [13] indices. Throughout the rest of this paper, the term "index" will refer to the differentiation index unless otherwise stated.

The *derivative array equations* and the notion of a 1-*full system* appear frequently in index analysis [6]. The derivative array is built by successive differentiations of the entire original system and has the form

$$\begin{bmatrix} \mathbf{A} & & & \\ \mathbf{B} & \mathbf{A} & & \\ & \mathbf{B} & \mathbf{A} & \\ & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \dot{\mathbf{u}} \\ \ddot{\mathbf{u}} \\ \dddot{\mathbf{u}} \\ \vdots \end{bmatrix} = - \begin{bmatrix} \mathbf{B}\mathbf{u} \\ \mathbf{0} \\ \vdots \\ \end{bmatrix} + \begin{bmatrix} \mathbf{f} \\ \dot{\mathbf{f}} \\ \ddot{\mathbf{f}} \\ \vdots \end{bmatrix}. \tag{2.2}$$

The ith derivative array equations are the first i block rows of this system and are often written as

$$\mathcal{A}_i \mathbf{u}_i = \mathbf{f}_i. \tag{2.3}$$

The ith derivative array is said to be 1-*full* if there exists a set of row operations represented by the matrix $\mathbf{R} \in \mathbb{R}^{in\times in}$ such that

$$\mathbf{R}\mathcal{A}_i = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix} \tag{2.4}$$

with $\mathbf{I} \in \mathbb{R}^{n\times n}$. If j is the smallest integer such that the jth derivative array is 1-full, then a minimum of $j-1$ differentiations are required to determine $\dot{\mathbf{u}}$ as a continuous function of $\mathbf{u}$ and t, and thus the index of the original system is $j-1$.

Suppose $\mathbf{A}$ and $\mathbf{B}$ form a regular pencil. Then, multiplication on the left by $\mathbf{P}$ and introduction of new variables $\mathbf{v}(t) = \mathbf{Q}^{-1}\mathbf{u}(t)$ produces the following pair of decoupled subsystems:

$$\begin{bmatrix} \mathbf{I} & \\ & \mathbf{N} \end{bmatrix} \begin{bmatrix} \dot{\mathbf{v}}_1(t) \\ \dot{\mathbf{v}}_2(t) \end{bmatrix} + \begin{bmatrix} \mathbf{J} & \\ & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1(t) \\ \mathbf{v}_2(t) \end{bmatrix} = \begin{bmatrix} \mathbf{g}_1(t) \\ \mathbf{g}_2(t) \end{bmatrix}. \tag{2.5}$$

This is called the *canonical form* of the DAE [1]. The first block row, referred to as the *differential part*, is a fully determined linear ordinary differential equation (ODE). A unique C^1 solution $\mathbf{v}_1$ exists on T for any initial values $\mathbf{v}_1(0)$ and continuous forcing function $\mathbf{g}_1$.

The solution of the second block row, called the *algebraic part*, is completely determined by the equations and is given by

$$\mathbf{v}_2(t) = \sum_{i=0}^{k-1} (-1)^i \mathbf{N}^i \left(\frac{d}{dt}\right)^i \mathbf{g}_2(t). \tag{2.6}$$

Provided that $\mathbf{g}_2$ is sufficiently differentiable, a unique solution $\mathbf{v}_2$ exists on T. Clearly, no arbitrary constants of integration appear in the solution of the algebraic part.

What properties of the system or a solution are made obvious through transformation to canonical form? First, the existence of the canonical form is a necessary condition for solvability. If $\mathbf{A}, \mathbf{B}$ do not form a regular pencil, then the system is not solvable [1].

Second, the *differentiation index*, ν, is equal to k, the nilpotency of $\mathbf{N}$. Derivatives of up to order $\nu - 1$ appear in the solution, so a sufficient condition for $\mathbf{u}$ to be C^1, which is in turn necessary for solvability, is that all forcing functions f_i have at least C^ν continuity. In the transformed system, a sufficient condition is that the elements of $\mathbf{g}_2$ have at least C^ν continuity.

Third, determination of a unique solution requires specification of m initial conditions, where m is the dimension of the first block row. This may differ from the number of differential equations or differential variables in the original DAE when $\nu > 0$.

2.3. Hyperbolic PDEs and characteristics. A linear, first order PDE over two independent variables t and x has the form

$$\mathbf{A}\mathbf{u}(t,x)_t + \mathbf{B}\mathbf{u}(t,x)_x = \mathbf{f}(t,x). \tag{2.7}$$

Here $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, $\Omega \subset \mathbb{R}^2$, $\mathbf{u}, \mathbf{f} : \Omega \to \mathbb{R}^n$, $\mathbf{u}_t \equiv \frac{\partial \mathbf{u}}{\partial t}$, and $\mathbf{u}_x \equiv \frac{\partial \mathbf{u}}{\partial x}$. Such a system is said to be *hyperbolic* iff $\mathbf{A}$ is invertible and $\mathbf{A}^{-1}\mathbf{B}$ is diagonalizable with strictly real eigenvalues [3, 11].

Multiplication of a hyperbolic system on the left by $\mathbf{A}^{-1}$, followed by introduction of new variables $\mathbf{u}(t,x) = \mathbf{S}\mathbf{v}(t,x)$ and multiplication on the left by $\mathbf{S}$, where $\mathbf{S}$ diagonalizes $\left(\mathbf{A}^{-1}\mathbf{B}\right)$, produces

$$\mathbf{v}(t,x)_t + \mathbf{\Lambda}\mathbf{v}(t,x)_x = \mathbf{g}(t,x). \tag{2.8}$$

This is the *characteristic form* [3, 11] of the original hyperbolic PDE. Each equation in the transformed system has the form

$$v_i(t,x)_t + \lambda_i v_i(t,x)_x = g_i(t,x). \tag{2.9}$$

Suppose that there is only one truly independent variable; in other words, let $x = x(t)$. More specifically, let $(x - x_0) = \lambda_i(t - t_0)$, so that $\frac{dx}{dt} = \lambda_i$. Then, by the chain rule,

$$\begin{aligned} \frac{dv_i(t, x(t))}{dt} &= v_i(t, x(t))_t + v_i(t, x(t))_x \frac{dx}{dt} \\ &= g_i(t, x). \end{aligned} \tag{2.10}$$

Each equation in the characteristic form is thus equivalent to an ODE along a particular direction in the (t, x)-plane. That direction, $\frac{dx}{dt} = \lambda_i$, is called a *characteristic direction* of the original system, or simply a *characteristic*.

For clarity, let $g_i(t, x) = 0$. The solution to the ODE is then

$$v_i(t, x) = v_i(t_0, x_0) \tag{2.11}$$

along $(x - x_0) = \lambda_i(t - t_0)$. Consider the domain given by $\Omega = [0, t_f] \times [a, b]$. If $\lambda_i > 0$, then specification of v_i on $t = 0$, $x \in [a, b]$ or on $t \in [0, t_f]$, $x = a$ provides an initial condition for the characteristic ODE.

What does the characteristic form tell us about the system? First, the number of boundary conditions that must be enforced along $x = a$ equals the number of positive characteristics ($\lambda_i > 0$). Similarly, the number of boundary conditions that must be enforced on $x = b$ equals the number of negative characteristics. The total number of boundary conditions required to determine a unique solution is equal to the number of nonzero characteristics. These boundary conditions provide “initial” conditions for the characteristic ODEs as they travel into the domain interior.

Second, the eigenvalues of $\mathbf{A}^{-1}\mathbf{B}$ govern the dependence of the solution on its data. The solution *depends continuously on its initial data* if the L_2 norm of the solution at t can be bounded by the L_2 norm of the initial data and some finite scaling factor $C : [0, t_f] \to \mathbb{R}$ that depends only on t:

$$||\mathbf{u}(t, \cdot)|| \leq C(t)||\mathbf{u}(0, \cdot)||. \tag{2.12}$$

Using the Fourier transform and Parseval's equation, one can show that the solution depends continuously on its initial data iff $\mathbf{A}^{-1}\mathbf{B}$ is diagonalizable and all eigenvalues are strictly real. With a suitable transformation of independent variables, the same analysis shows that those same conditions on the eigenvalues of $\mathbf{A}^{-1}\mathbf{B}$ are necessary and sufficient for the solution to depend continuously on its boundary data.

If the solution depends on derivatives of the initial data, the scaling factor C will have a polynomial dependence on ω in the Fourier domain. Such a solution is weakly ill-posed.[1] It might be possible to obtain a useful numerical solution of a weakly ill-posed system with a discretization of sufficiently high order [22]. If the scaling factor has an exponential dependence on ω, the solution is strongly ill-posed, and roundoff error will eventually govern a computed numerical solution [12].

2.4. PDE systems and index. The differentiation index of a PDE is defined with respect to a direction in the independent variable space [17]. Consider the PDE system

$$\mathbf{f}(\mathbf{u}(t, x)_t, \mathbf{u}(t, x)_x, \mathbf{u}(t, x), t, x) = \mathbf{0}. \tag{2.13}$$

[1] Weakly ill-posed systems are sometimes called *weakly well-posed*, although they are in fact ill-posed.

Here $\Omega \subset \mathbb{R}^2$, $\mathbf{u} : \Omega \to \mathbb{R}^n$, $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \Omega \to \mathbb{R}^n$, $\mathbf{u}_t \equiv \frac{\partial \mathbf{u}}{\partial t}$, and $\mathbf{u}_x \equiv \frac{\partial \mathbf{u}}{\partial x}$. The differentiation index with respect to t is defined as the minimum number of times that some or all of the equations must be differentiated with respect to t in order to determine $\mathbf{u}_t$ as a continuous function of $\mathbf{u}$, t, and x. This function may include partial differential operators in x. The differentiation index with respect to x is defined in an analogous fashion.

The derivative array equations may be defined in a manner that is notationally similar to the derivative array for a DAE by using operator-valued matrices. For our linear, first order system (2.7), the derivative array equations in t have the form

$$\begin{bmatrix} \mathbf{A} & & & \\ \mathbf{B}\frac{\partial}{\partial x} & \mathbf{A} & & \\ & \mathbf{B}\frac{\partial}{\partial x} & \mathbf{A} & \\ & & \ddots & \ddots \end{bmatrix} \begin{bmatrix} \mathbf{u}_t \\ \mathbf{u}_{tt} \\ \mathbf{u}_{ttt} \\ \vdots \end{bmatrix} = - \begin{bmatrix} \mathbf{B}\frac{\partial}{\partial x}\mathbf{u} \\ \mathbf{0} \\ \vdots \\ \end{bmatrix} + \begin{bmatrix} \mathbf{f} \\ \mathbf{f}_t \\ \mathbf{f}_{tt} \\ \vdots \end{bmatrix}, \tag{2.14}$$

which is again written more compactly as

$$\mathcal{A}_i \mathbf{u}_i = \mathbf{f}_i. \tag{2.15}$$

Here $\mathcal{A}_i \in P^{in \times in}$, where $P = \{\mathcal{L} : \mathcal{L}u = \sum_{j=1}^{\infty} c_j \left(\frac{\partial}{\partial x}\right)^j u, c_j \in \mathbb{R}\}$, the set of all linear partial differential operators in x.

The ith derivative array in t is said to be 1-full if there exists an $\mathbf{R} \in P^{in \times in}$ such that

$$\mathbf{R}\mathcal{A}_i = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \mathbf{H} \end{bmatrix}, \tag{2.16}$$

where $\mathbf{D} \in P^{n \times n}$ is a diagonal matrix. $\mathbf{D}$ is not necessarily $\mathbf{I}$ because $\langle P, +, \times \rangle$ is only a ring [17]. Note that 1-fullness does not require that $\mathbf{u}_t$ be given as an *explicit* function of $\mathbf{u}$, t, and x.

Index analysis with respect to a particular direction in the independent variable space reveals the number of degrees of freedom that exist on hyperplanes normal to that direction [17]. If the index with respect to t of a system of dimension n is zero, then n degrees of freedom exist on surfaces of the form $t = c$. Similarly, if the index of that same system with respect to x is zero, then n degrees of freedom exist on surfaces of the form $x = c$. These degrees of freedom must be specified by auxiliary conditions in order to determine a unique solution.

Other indices have been defined for PDE systems. They include the perturbation index, the algebraic index, and various other t- and x-indices [2, 9, 15, 16]. In this paper, the term "index with respect to z," where z is an independent variable over which a PDE is defined, will refer to the differentiation index unless otherwise noted.

3. Linear systems with simple forcing. Now consider a linear, not necessarily hyperbolic system over two independent variables, for which the forcing terms are functions of the independent variables only. Such a system will be referred to as a *linear system with simple forcing*:

$$\mathbf{A}\mathbf{u}(t,x)_t + \mathbf{B}\mathbf{u}(t,x)_x = \mathbf{f}(t,x). \tag{3.1}$$

Let $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{n \times n}$, $\Omega \subset \mathbb{R}^2$, and $\mathbf{u}, \mathbf{f} : \Omega \to \mathbb{R}^n$. We will consider domains of the form $\Omega = [0, t_f] \times [a, b]$, with the additional restriction that information at a time t_2 may not be used to determine the solution at t_1 unless $t_1 \geq t_2$.[2]

[2]This does not necessarily indicate that numerical solution will be by an evolution method in t.

Let such a system be called *regular* iff the coefficient matrices form a regular pencil. Regularity is a necessary condition for uniqueness of the solution and thus for well-posedness [2]. Henceforth, all systems with simple forcing will be assumed regular.

Every regular linear system with simple forcing is equivalent to one of the following forms, which will be referred to as its *canonical form*:

$$\begin{bmatrix} \mathbf{I} & & \\ & \mathbf{I} & \\ & & \mathbf{N}_2 \end{bmatrix} \begin{bmatrix} \mathbf{v}_1(t,x) \\ \mathbf{v}_2(t,x) \\ \mathbf{v}_3(t,x) \end{bmatrix}_t + \begin{bmatrix} \mathbf{J} & & \\ & \mathbf{N}_1 & \\ & & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{v}_1(t,x) \\ \mathbf{v}_2(t,x) \\ \mathbf{v}_3(t,x) \end{bmatrix}_x = \begin{bmatrix} \mathbf{f}_1(t,x) \\ \mathbf{f}_2(t,x) \\ \mathbf{f}_3(t,x) \end{bmatrix}. \tag{3.2}$$

Here $\mathbf{J}$ is an invertible lower Jordan matrix, and $\mathbf{N}_1$ and $\mathbf{N}_2$ are lower Jordan matrices of nilpotency ν_1 and ν_2, respectively. This canonical form may be constructed in the same manner as the canonical form of a linear time-invariant DAE. Let the *max degeneracy*, denoted ν_{max}, be defined as one less than the dimension of the largest Jordan block of $\mathbf{J}$, $\mathbf{N}_1$, or $\mathbf{N}_2$.

A hyperbolic PDE in characteristic form is a special case of the first block row, so the first block row will be referred to as the *hyperbolic part* of the canonical form. An ODE system is a special case of the second block row, so it will be referred to as the *differential part* of the canonical form. When the forcing function is allowed a linear dependence on the dependent variables, the heat equation (written as a first order system) will be a special case of the third block row, so let the third block row be called the *parabolic part* of the canonical form.

Can this canonical form provide information about the original system in a manner similar to the canonical form of a DAE or the characteristic form of a hyperbolic PDE? Consider first the differentiation index of the system.

THEOREM 3.1. *The differentiation index with respect to t, ν_t, of a linear system with simple forcing is equal to ν_2.*

Proof. The first and second block rows of the canonical form give $\mathbf{v}_{1_t}$ and $\mathbf{v}_{2_t}$ as continuous functions of $\mathbf{v}_x$, t, and x. The smallest derivative array with respect to t for the third block row that must be 1-full has $\nu_2 + 1$ block rows, so the index of the system with respect to t is ν_2. ∎

COROLLARY 3.2. *The differentiation index with respect to x, ν_x, of a linear system with simple forcing is equal to ν_1.*

Now, can the canonical form provide insight into the dependence of the solution on its data?

LEMMA 3.3. *The unrestricted solution to a regular, first order system with simple forcing depends continuously on its initial data iff the generalized eigenvalues in the differential and hyperbolic parts of the canonical form are strictly real and of degeneracy zero.*

Proof. Because the coefficient matrices form a regular pencil, the homogeneous system is equivalent to one which, in Fourier space, has the form

$$\begin{bmatrix} \tilde{\mathbf{I}} & \\ & \tilde{\mathbf{N}} \end{bmatrix} \hat{\mathbf{v}}(t,\omega)_t + i\omega \begin{bmatrix} \tilde{\mathbf{J}} & \\ & \tilde{\mathbf{I}} \end{bmatrix} \hat{\mathbf{v}}(t,\omega) = \mathbf{0},$$

which may be rewritten as

$$\tilde{\mathbf{A}}\hat{\mathbf{v}}(t,\omega)_t + \tilde{\mathbf{B}}\hat{\mathbf{v}}(t,\omega) = \mathbf{0}.$$

The solution is

$$\hat{\mathbf{v}}(t,\omega) = e^{-\tilde{\mathbf{A}}^D\tilde{\mathbf{B}}t}\tilde{\mathbf{A}}\tilde{\mathbf{A}}^D\hat{\mathbf{v}}(0,\omega),$$

where

$$\tilde{\mathbf{A}}^D\tilde{\mathbf{B}} = \begin{bmatrix} i\omega\tilde{\mathbf{J}} & \\ & \mathbf{0} \end{bmatrix}, \qquad \tilde{\mathbf{A}}\tilde{\mathbf{A}}^D = \begin{bmatrix} \tilde{\mathbf{I}} & \\ & \mathbf{0} \end{bmatrix},$$

and thus

$$||\hat{\mathbf{v}}(t,\omega)|| = ||e^{-\tilde{\mathbf{A}}^D\tilde{\mathbf{B}}t}\tilde{\mathbf{A}}\tilde{\mathbf{A}}^D\hat{\mathbf{v}}(0,\omega)|| \leq ||e^{-\tilde{\mathbf{A}}^D\tilde{\mathbf{B}}t}|| \; ||\hat{\mathbf{v}}(0,\omega)||.$$

First assume that all eigenvalues are strictly real and nondegenerate. Then $\tilde{\mathbf{J}} = \mathbf{\Lambda}$ with $\Lambda_{jj} \in \mathbb{R}$ so $||e^{i\omega\mathbf{\Lambda}t}|| = 1$, and there exists a finite $C(t)$ independent of ω such that

$$||e^{-\tilde{\mathbf{A}}^D\tilde{\mathbf{B}}t}|| \; ||\hat{\mathbf{v}}(0,\cdot)|| \leq C(t)||\hat{\mathbf{v}}(0,\cdot)||.$$

By Parseval's equation, the result holds in (t,x) space as well, and by Duhamel's principle, the result holds for simple forcing.

Second, assume that the system depends continuously on its initial data. Then there exists a finite $C(t)$ independent of ω such that

$$||\mathbf{v}(t,\cdot)|| \leq C(t)||\hat{\mathbf{v}}(0,\cdot)||,$$

so one can choose $C(t)$ such that

$$C(t) \geq ||e^{-\tilde{\mathbf{A}}^D\tilde{\mathbf{B}}t}||$$

for all $\omega \in \mathbb{R}$.

If there is a block of $\tilde{\mathbf{J}}$ that corresponds to an eigenvalue of degeneracy $k > 0$, then recall that the matrix norm is bounded from below by the magnitude of its largest element. Thus we have

$$||e^{-\tilde{\mathbf{A}}^D\tilde{\mathbf{B}}t}|| \geq \frac{c}{k!}|\omega|^k t^k$$

in contradiction to the bound independent of ω.

If $\lambda = a + bi$, $b \neq 0$, is an eigenvalue, then we have

$$||e^{-\tilde{\mathbf{A}}^D\tilde{\mathbf{B}}t}|| \geq |e^{-i\omega\lambda t}| = |e^{-i\omega at}e^{\omega bt}|$$

in contradiction to the bound independent of ω. ◻

Taking the Fourier transform in t rather than x gives the analogous result for the parabolic part.

COROLLARY 3.4. *The unrestricted solution to a regular, first order system with simple forcing depends continuously on its boundary data iff the eigenvalues in the parabolic and hyperbolic parts of the canonical form are of degeneracy zero with strictly real eigenvalues.*

Taken together, this lemma and its corollary give the desired result.

THEOREM 3.5. *The unrestricted solution to a regular, first order system with simple forcing depends continuously on its data iff $\nu_{max} = 0$ and the generalized eigenvalues of the coefficient matrix pair are strictly real.*

Can the canonical form provide information about the number and location of initial and boundary conditions needed to determine a unique solution? This question may be answered separately for each of the three block rows of the canonical form.

First, every Jordan block in the differential part gives rise to a system of the form

$$\mathbf{w}(t,x)_t + \mathbf{N}\mathbf{w}(t,x)_x = \mathbf{g}(t,x). \tag{3.3}$$

Here $\mathbf{N}$ has unity on the first subdiagonal and zero everywhere else.

The index of this system with respect to t is 0, and there are m degrees of freedom on surfaces of the form $t = c$, where m is the dimension of the block. Because information at t_1 cannot be used to determine the solution at $t = t_2$ unless $t_1 \leq t_2$, each of these degrees of freedom must be specified on $t = 0$. The index of the system with respect to x is m, and there are no degrees of freedom on surfaces of the form $x = c$. If n_2 is the dimension of the entire differential part, then n_2 initial conditions and no boundary conditions are required to uniquely determine $\mathbf{v}_2(t,x)$.

Second, every Jordan block in the hyperbolic part gives rise to a system of the form

$$\mathbf{w}(\mathbf{t},\mathbf{x})_t + \mathbf{J}\mathbf{w}(t,x)_x = \mathbf{g}(t,x). \tag{3.4}$$

Here $\mathbf{J}$ has some nonzero scalar λ on the diagonal, unity on the first subdiagonal, and zero everywhere else. Assuming that the original system depends continuously on its data, $\lambda \in \mathbb{R}$.

The index of this system with respect to t is zero, and there are m degrees of freedom on surfaces of the form $t = c$. For the same reason noted for a block of the differential part, all of these degrees of freedom must be specified on $t = 0$.

The index of this system with respect to x is zero, and there are m degrees of freedom on surfaces of the form $x = c$. The constraint involving t also determines where these degrees of freedom must be specified. The first equation in the block is a one-way wave along the characteristic direction $\frac{dx}{dt} = \lambda$, so one boundary condition must be specified at $x = a$ if $\lambda > 0$, or at $x = b$ if $\lambda < 0$. Once w_1 is uniquely specified, the second equation in the block becomes another one-way wave, again along characteristic direction $\frac{dx}{dt} = \lambda$, and another boundary condition must be specified at the same domain boundary.

Let n_1 be the dimension of the hyperbolic part. Let n_{1a} be the sum of the dimensions of all blocks in the hyperbolic part for which $\lambda > 0$, and then let $n_{1b} = n_1 - n_{1a}$. A total of n_{1a} boundary conditions on $x = a$, n_{1b} boundary conditions on $x = b$, and n_1 initial conditions must be specified in order to uniquely determine $\mathbf{v}_1$.

Third, every Jordan block in the parabolic part is a system of the form

$$\mathbf{w}(t,x)_x + \mathbf{N}\mathbf{w}(t,x)_t = \mathbf{g}(t,x). \tag{3.5}$$

Here $\mathbf{N}$ has unity on the first subdiagonal and zero everywhere else.

The index with respect to x is zero, and there are m degrees of freedom on surfaces of the form $x = c$. Now, however, one can integrate either forward or backward in x from either $x = a$ or $x = b$ into the domain interior without violating the constraint on information traveling backward in t. Each of the m degrees of freedom may be specified on either $x = a$ or $x = b$.

The index of the system with respect to t is m. There are no degrees of freedom on surfaces of the form $t = c$. No initial conditions are needed to determine a unique solution. Let n_3 be the dimension of the entire parabolic part. A total of n_3 boundary conditions and no initial conditions must be specified to uniquely determine $\mathbf{v}_3(t,x)$. Any number n_{3a} of these boundary conditions may be enforced on $x = a$, and the remaining $n_{3b} = n_3 - n_{3a}$ are then enforced on $x = b$.

Can the canonical form provide insight into the highest order partial derivatives of forcing functions that might appear in the solution? With the appropriate selection of independent variables, every Jordan block in the canonical form may be written as a system of the form

$$\mathbf{w}(r,s)_r + \mathbf{N}\mathbf{w}(r,s)_s = \mathbf{g}(r,s). \tag{3.6}$$

The first equation in the system is an uncoupled ODE in r involving only $w_1(r,s)$. Once $w_1(r,s)$ is determined, the second equation becomes another uncoupled ODE in r. Integrating each successive equation forward in r gives the solution, which may be written as

$$\begin{aligned} w_1(r,s) &= c_1(s) + \int_0^r g_1(r_1,s)dr_1, \\ w_2(r,s) &= -rc_1'(s) + c_2(s) + \int_0^r \left[g_2(r_2,s) - \int_0^{r_2} g_1(r_1,s)_s dr_1 \right] dr_2, \\ w_3(r,s) &= \frac{r^2}{2}c_1''(s) - rc_2(s) + c_3(s) \\ &\qquad + \int_0^r \left[g_3(r_3,s) - \int_0^{r_3} \left[g_2(r_2,s)_s - \int_0^{r_2} g_1(r_1,s)_{ss} dr_1 \right] dr_2 \right] dr_3 \\ &\;\;\vdots \end{aligned} \tag{3.7}$$

If the block has dimension m, then partial derivatives of the forcing functions of up to order $m-1$ appear in the solution. Inspecting the solution for all Jordan blocks in the canonical form, we see that partial derivatives of up to order ν_{max} of data and forcing functions may appear in the solution. Up to $C^{\nu_{max}+1}$ continuity may therefore be required of data and forcing functions in order for $\mathbf{u} \in C^1$. Note that ν_{max} is unrelated to the differentiation index with respect to either t or x; partial derivatives of data and forcing functions may appear in the solution of linear PDEs of index 0 with respect to both t and x.

Transforming a block row from the hyperbolic part to this standard form requires a change of independent variables. This dependence of the solution on derivatives of the initial data and forcing functions may also be shown in the original independent variables. Consider a system that consists of only a 2×2 hyperbolic part:

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1(t,x) \\ u_2(t,x) \end{bmatrix}_t + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} u_1(t,x) \\ u_2(t,x) \end{bmatrix}_x = \begin{bmatrix} g_1(t,x) \\ g_2(t,x) \end{bmatrix}. \tag{3.8}$$

The first equation in simply a one-way wave with unit velocity. The solution may be given by integrating along the characteristic direction as follows:

$$u_1(t,x) = u_1(t_0, x(t_0)) + \int_{t_0}^t f_1(t, x(t))dt. \tag{3.9}$$

Here $x(t) = x_0 + t - t_0$. Substitution of this expression into the second equation then gives another one-way wave with unit velocity:

$$u_2(t,x)_t + u_2(t,x)_x = f_2(t,x) - u_1(t_0, x(t_0))_x + \int_{t_0}^t f_1(t, x(t))_x dt. \tag{3.10}$$

The first partial derivative of both the initial data and f_1 appears in the solution for u_2, even though $\nu_t = \nu_x = 0$.

The canonical form of a linear system with simple forcing gives several properties of the system directly. The index of the system with respect to t is equal to the dimension of the largest Jordan block within the differential subsystem, while the index with respect to x is equal to the dimension of the largest Jordan block within the parabolic subsystem. The highest order derivative of initial or boundary data that appears in the solution is equal to ν_{max}, the max degeneracy of the system; it is not necessarily given by a differentiation index with respect to either t or x. The solution depends continuously on its data iff $\nu_{max} = 0$ and all generalized eigenvalues are strictly real. The number of independent boundary conditions required equals the dimension of the hyperbolic plus parabolic parts. At least n_{1a} and n_{1b} must be enforced on $x = a$ and $x = b$, respectively, where n_{1a} and n_{1b} are the number of positive and negative generalized eigenvalues, respectively.

4. Linear systems with linear forcing.

When forcing is simple, the canonical form provides the index with respect to both t and x, the dependence of the solution on its data, the number and location of initial and boundary conditions, and bounds on the order of partial derivatives of forcing functions and data that appear in the solution.

Does the canonical form provide the same information when the forcing functions include linear combinations of the dependent variables? Such a system will be referred to as a *linear system with linear forcing* and has the following form:

$$\mathbf{A}\mathbf{u}(t,x)_t + \mathbf{B}\mathbf{u}(t,x)_x = \mathbf{f}(t,x) - \mathbf{G}\mathbf{u}(t,x). \tag{4.1}$$

Here $\mathbf{A}, \mathbf{B}, \mathbf{G} \in \mathbb{R}^{n\times n}$ and $\mathbf{u}, \mathbf{f} : \Omega \to \mathbb{R}^n$.

If the coefficient matrices of a linear system with linear forcing form a regular pencil,[3] the canonical form of the system is

$$\begin{aligned}\begin{bmatrix}\mathbf{I} & & \\ & \mathbf{I} & \\ & & \mathbf{N}_2\end{bmatrix}\begin{bmatrix}\mathbf{v}_1(t,x)\\ \mathbf{v}_2(t,x)\\ \mathbf{v}_3(t,x)\end{bmatrix}_t + \begin{bmatrix}\mathbf{J} & & \\ & \mathbf{N}_1 & \\ & & \mathbf{I}\end{bmatrix}\begin{bmatrix}\mathbf{v}_1(t,x)\\ \mathbf{v}_2(t,x)\\ \mathbf{v}_3(t,x)\end{bmatrix}_x &= \begin{bmatrix}\mathbf{f}_1(t,x)\\ \mathbf{f}_2(t,x)\\ \mathbf{f}_3(t,x)\end{bmatrix} - \mathbf{C}\mathbf{v}(t,x)\\ &= \begin{bmatrix}\mathbf{C}_1\mathbf{v}_1(t,x) + \mathbf{g}_1(\mathbf{v}_2(t,x), \mathbf{v}_3(t,x)) + \mathbf{f}_1(t,x)\\ \mathbf{C}_2\mathbf{v}_2(t,x) + \mathbf{g}_2(\mathbf{v}_1(t,x), \mathbf{v}_3(t,x)) + \mathbf{f}_2(t,x)\\ \mathbf{C}_3\mathbf{v}_3(t,x) + \mathbf{g}_3(\mathbf{v}_1(t,x), \mathbf{v}_2(t,x)) + \mathbf{f}_3(t,x)\end{bmatrix}.\end{aligned} \tag{4.2}$$

Again $\mathbf{J}$ is an invertible lower Jordan matrix, and $\mathbf{N}_1$ and $\mathbf{N}_2$ are lower Jordan matrices of nilpotency ν_1 and ν_2, respectively. The block rows are no longer independent; they may be coupled through the linear forcing terms. Over the entire canonical form, the ith variable is assigned to the ith equation.

Now consider a linear system with linear forcing for which the coefficient matrix pair is singular, and which may be written as

$$\begin{bmatrix}\mathbf{A}_{11} & \mathbf{A}_{12}\\ & \end{bmatrix}\mathbf{u}(t,x)_t + \begin{bmatrix}\mathbf{B}_{11} & \mathbf{B}_{12}\\ & \end{bmatrix}\mathbf{u}(t,x)_x = \mathbf{f}(t,x) - \begin{bmatrix}\mathbf{G}_{11} & \mathbf{G}_{12}\\ \mathbf{G}_{21} & \mathbf{G}_{22}\end{bmatrix}\mathbf{u}(t,x) \tag{4.3}$$

with $\mathbf{G}_{22} \in \mathbb{R}^{m\times m}$ and $|\mathbf{G}_{22}| \neq 0$. Finally, let $\mathbf{A}_{11}$, $\mathbf{B}_{11}$ form a regular pencil. Differentiating the second block row with respect to t produces a system with a regular

[3] Singularity of the coefficient matrix pencil no longer precludes well-posedness; a linear system with linear forcing is regular iff $\exists \rho, \tau \in \mathbb{C}$ such that $|\rho\mathbf{A} + \tau\mathbf{B} + \mathbf{G}| \neq 0$ [2].

coefficient matrix pencil, which may then be transformed to canonical form. This second system admits a larger family of solutions; m additional initial conditions are required to determine a unique member of that family. Analysis of the canonical form of the differentiated system will thus overstate by m the number of initial conditions required to determine a unique solution of the original system. The index of the differentiated system with respect to t may also be one less than the index with respect to t of the original system. The results of all other analyses of the canonical form of the differentiated system will hold for the original system as well.

Can the canonical form give the index with respect to either t or x by inspection? Without restricting the form of $\mathbf{C}$, it is impossible to obtain the index by inspection of the canonical form. However, the canonical form does provide an upper bound on the index.

THEOREM 4.1. *For a linear system with linear forcing, $\nu_t \leq \nu_2$, where ν_2 is the nilpotency of the parabolic part.*

Proof. Rewrite the canonical form as

$$\begin{bmatrix} \mathbf{I} & \\ & \mathbf{N}_2 \end{bmatrix} \mathbf{v}(t,x)_t + \begin{bmatrix} \bar{\mathbf{J}} & \\ & \mathbf{I} \end{bmatrix} \mathbf{v}(t,x)_x = \begin{bmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{bmatrix} \mathbf{v}(t,x) + \mathbf{f}(t,x). \tag{4.4}$$

After one differentiation, the coefficient matrix $\mathcal{A}_2$ in the derivative array has the form

$$\begin{bmatrix} \mathbf{I} & & & \\ & \mathbf{N}_2 & & \\ \bar{\mathbf{J}}\frac{\partial}{\partial x} + \mathbf{C}_{11} & \mathbf{C}_{12} & \mathbf{I} & \\ \mathbf{C}_{21} & \mathbf{I}\frac{\partial}{\partial x} + \mathbf{C}_{22} & & \mathbf{N}_2 \end{bmatrix}. \tag{4.5}$$

Let $\mathbf{B} = \mathbf{I}\frac{\partial}{\partial x} + \mathbf{C}_{22}$, and let $\mathbf{B}^k$ be the diagonal block of $\mathbf{B}$ that corresponds to the kth block of $\mathbf{N}_2$. Also, let $\mathbf{v}^k(t,x)_t$ be the elements of $\mathbf{v}(t,x)_t$ that correspond to the kth block of $\mathbf{N}_2$.

Let n be the dimension of the kth block. Then by successive differentiations with respect to t and construction of determinants within the corresponding derivative array equations, one may show that $\mathbf{v}^k(t,x)_t$ will be given as a continuous function of $\mathbf{v}(t,x)$, t, and x after at most n differentiations with respect to t. The largest block has $n = \nu_2$, so $\nu_t \leq \nu_2$. ∎

COROLLARY 4.2. *For a linear system with linear forcing, $\nu_x \leq \nu_1$, where ν_1 is the nilpotency of the differential part.*

COROLLARY 4.3. *If the coefficient matrix $\mathbf{C}$ in the canonical form of a linear system with linear forcing is strictly lower triangular, then $\nu_t = \nu_2$ and $\nu_x = \nu_1$.*

Proof. If $\mathbf{C}$ is strictly lower triangular, then ν_2 differentiations with respect to t are required to produce a 1-full derivative array in t, and ν_1 differentiations with respect to x are required to produce a 1-full derivative array in x. ∎

Can the canonical form provide insight into the dependence of the unrestricted solution on its data?

THEOREM 4.4. *Assuming that $(\mathbf{A}, \mathbf{B})$ forms a regular pencil, the unrestricted solution to $\mathbf{Au}(t,x)_t + \mathbf{Bu}(t,x)_x + \mathbf{Cu}(t,x) = \mathbf{f}$ depends continuously on its data iff the unrestricted solution to $\mathbf{Au}(t,x)_t + \mathbf{Bu}(t,x)_x = \mathbf{0}$ depends continuously on its data.*

Proof. Assume the systems are already in canonical form, and consider solution of a single block row of one of the three subsystems for the variables $\mathbf{v}(t,x)$ assigned to it, in terms of the remaining variables $\mathbf{w}(t,x)$. With the proper choice of independent variables, the block row has the form

$$\mathbf{v}(r,s)_r + \mathbf{N}\mathbf{v}(r,s)_s + \mathbf{C}\mathbf{v}(r,s) + \mathbf{g}(\mathbf{w}(r,s)) = \mathbf{f}(r,s).$$

At this point, $\mathbf{g}(\mathbf{w}(r,s))$ is simply a vector of unknown functions of r and s, so let $\mathbf{h}(r,s) = \mathbf{f}(r,s) - \mathbf{g}(\mathbf{w}(r,s))$. Taking the Fourier transform of the system produces

$$\hat{\mathbf{v}}(r,\omega)_r + (i\omega\mathbf{N} + \mathbf{C})\,\hat{\mathbf{v}}(r,\omega) = \hat{\mathbf{h}}(r,\omega).$$

The solution is given by

$$\hat{\mathbf{v}}(r,\omega) = e^{-(i\omega\mathbf{N}+\mathbf{C})r}\hat{\mathbf{v}}(0,\omega) + e^{-(i\omega\mathbf{N}+\mathbf{C})r}\int_0^r e^{(i\omega\mathbf{N}+\mathbf{C})r_1}\hat{\mathbf{h}}(r_1,\omega)dr_1.$$

By Duhamel's principle, the forced solution may be thought of as a superposition of solutions to the corresponding homogeneous problem, with

$$\hat{\mathbf{v}}^*(0,\omega) = \int_0^r e^{(i\omega\mathbf{N}+\mathbf{C})r_1}\hat{\mathbf{h}}(r_1,\omega)dr_1.$$

For the homogeneous problem, it then remains to be shown that

$$||e^{-(i\omega\mathbf{N}+\mathbf{C})r}|| \le C(r) \Leftrightarrow ||e^{-(i\omega\mathbf{N})r}|| \le C^*(r)$$

for finite $C(r)$ and $C^*(r)$ independent of ω.

First, assume that $||e^{-(i\omega\mathbf{N})r}|| \le C^*(r)$. Then

$$||e^{-(i\omega\mathbf{N}+\mathbf{C})r}|| = ||e^{-i\omega\mathbf{N}r}e^{-\mathbf{C}r}|| \le ||e^{-i\omega\mathbf{N}r}||\;||e^{-\mathbf{C}r}|| \le C^*(r)||e^{-\mathbf{C}r}||.$$

Because $\mathbf{C}$ is a constant matrix, $C^*(r)||e^{-\mathbf{C}r}|| = C(r)$ is a function of r only.

Second, assume that $||e^{-(i\omega\mathbf{N}+\mathbf{C})r}|| \le C(r)$. Then

$$\begin{aligned} ||e^{-i\omega\mathbf{N}r}e^{-\mathbf{C}r}|| &\le C(r), \\ ||e^{-i\omega\mathbf{N}r}e^{-\mathbf{C}r}||\;||e^{\mathbf{C}r}|| &\le C(r)||e^{\mathbf{C}r}||, \\ ||e^{-i\omega\mathbf{N}r}|| &\le C(r)||e^{\mathbf{C}r}||, \end{aligned}$$

and, by an argument similar to the one presented above, the function $C^*(r) = C(r)||e^{\mathbf{C}r}||$ is a function of r only.

Because the selection of this first block to be solved is arbitrary, bounds on the unrestricted solution independent of ω hold for every block of $\mathbf{A}\mathbf{u}(t,x)_t + \mathbf{B}\mathbf{u}(t,x)_x + \mathbf{C}\mathbf{u}(t,x) = \mathbf{f}(t,x)$ iff they hold for every block of $\mathbf{A}\mathbf{u}(t,x)_t + \mathbf{B}\mathbf{u}(t,x)_x = \mathbf{f}(t,x)$. ∎

Can the canonical form provide insight into the auxiliary conditions required to determine a unique solution? Again, let us consider a single block row from each of the three parts of the canonical form.

First, let $\bar{\mathbf{v}}(t,x)$ be the variables assigned to a particular block row of the hyperbolic part, and let $\mathbf{w}(t,x)$ be the other variables. The block row has the form

$$\bar{\mathbf{v}}(t,x)_t + \mathbf{J}\bar{\mathbf{v}}(t,x)_x = \bar{\mathbf{C}}\bar{\mathbf{v}}(t,x) + \mathbf{g}\left(\mathbf{w}(t,x)\right) + \mathbf{f}(t,x), \tag{4.6}$$

where $\mathbf{J}$ is a lower Jordan matrix with a nonzero scalar λ on the diagonal, unity on the first subdiagonal, and zero everywhere else. Assume that the system depends continuously on its data so that $\lambda \in \mathbb{R}$.

In these equations, $\mathbf{w}(t, x)$ is simply another unknown function of t and x, so forcing terms involving $\mathbf{w}$ are considered together with $\mathbf{f}$. The index of this system with respect to t is zero, so there are m degrees of freedom on $t = 0$.

The index of the system with respect to x is also zero, so m boundary conditions must be specified. The system is equivalent to an ODE along $\frac{dx}{dt} = \lambda$ so, as before, the boundary conditions must be specified on $x = a$ if $\lambda > 0$, or else on $x = b$ if $\lambda < 0$.

Second, let $\bar{\mathbf{v}}(t, x)$ be the variables assigned to a particular block row of the parabolic part, and let $\mathbf{w}(t, x)$ be the other variables. The block row has the form

$$\mathbf{N}\bar{\mathbf{v}}(t,x)_t + \bar{\mathbf{v}}(t,x)_x = \bar{\mathbf{C}}\bar{\mathbf{v}}(t,x) + \mathbf{g}\left(\mathbf{w}(t,x)\right) + \mathbf{f}(t,x), \tag{4.7}$$

where $\mathbf{N}$ is a lower Jordan matrix with unity on the first subdiagonal and zero everywhere else. The index of the system with respect to x is zero, so there are m degrees of freedom on surfaces of the form $x = c$.

The index of the system with respect to t is between 1 and m, and there may be between zero and $m - 1$ degrees of freedom on $t = 0$. If there are no degrees of freedom, the boundary conditions are all data. If there exists one or more degrees of freedom on $t = 0$, some of the boundary conditions must not be used as data in order to admit the additional degrees of freedom on $t = 0$. This is typically done by specification of boundary conditions on the same dependent variable on both $x = a$ and $x = b$.

Example 1. *Consider the system*

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \mathbf{v}(t,x)_t + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{v}(t,x)_x = \begin{bmatrix} 0 & c_{12} \\ 0 & 0 \end{bmatrix} \mathbf{v}(t,x). \tag{4.8}$$

If $c_{12} = 0$, the index with respect to t is two, and there are no degrees of freedom on $t = 0$. The two boundary conditions are data and determine a unique solution, which is weakly ill-posed.

If $c_{12} \neq 0$, the index with respect to t is one, and there is one degree of freedom on $t = 0$. Specification of $v_1(t, a)$ and $v_1(t, b)$ admits a family of sine waves of particular wavelengths, which can represent an arbitrary degree of freedom on $t = 0$. These boundary conditions in this case are not data, but instead restrict the members of the family of sine wave functions from which the solution is constructed. Specification of two boundary conditions on the same domain boundary determines a unique solution that is strongly ill-posed.

Third, let $\bar{\mathbf{v}}(t, x)$ be the variables assigned to a particular block row of the differential part, and let $\mathbf{w}(t, x)$ be the other variables. The block row has the form

$$\mathbf{N}\bar{\mathbf{v}}(t,x)_x + \bar{\mathbf{v}}(t,x)_t = \bar{\mathbf{C}}\bar{\mathbf{v}}(t,x) + \mathbf{g}\left(\mathbf{w}(t,x)\right) + \mathbf{f}(t,x). \tag{4.9}$$

This is analogous to a block in the parabolic part, with the roles of t and x reversed. The index of the system with respect to t is zero, so there are m degrees of freedom on $t = 0$.

The index of the system with respect to x is between 1 and m, and there may be between zero and $m - 1$ degrees of freedom on surfaces of the form $x = c$. If there are no degrees of freedom, the initial conditions are all data. If there exists one or more degrees of freedom, some initial conditions would need to be enforced instead on $t = t_2 > 0$. Under the restriction that information at t_2 cannot be used to determine the solution at $t_1 < t_2$, it is then impossible to form an admissible problem.

Can the canonical form provide insight into the smoothness required of auxiliary conditions and forcing functions? The unrestricted solution consists of blocks of the same basic form as the simple forcing case. However, the forcing functions may depend on dependent variables assigned to other blocks. These other dependent variables may in turn depend on derivatives of other forcing functions and their own data, which may increase the order of derivatives that appear in the solution.

EXAMPLE 2. *Consider the following canonical form of a linear system with linear forcing:*

$$\begin{bmatrix} 1 & 0 & & \\ 0 & 1 & & \\ & & 1 & 0 \\ & & 0 & 1 \end{bmatrix} \mathbf{v}(t,x)_t + \begin{bmatrix} 0 & 0 & & \\ 1 & 0 & & \\ & & 0 & 0 \\ & & 1 & 0 \end{bmatrix} \mathbf{v}(t,x)_x = \begin{bmatrix} f_1(t,x) \\ 0 \\ v_2(t,x) \\ 0 \end{bmatrix}. \tag{4.10}$$

The solution is

(4.11)

$$\begin{aligned} v_1(t,x) &= v_1(0,x) + \int_0^t f_1(t_1,x)dt_1, \\ v_2(t,x) &= v_2(0,x) - tv_1'(0,x) - \int_0^t \int_0^{t_2} f_1(t_1,x)_x dt_1 \ \ dt_2, \\ v_3(t,x) &= v_3(0,x) + tv_2(0,x) - \frac{1}{2}t^2 v_1'(0,x) - \int_0^t \int_0^{t_3} \int_0^{t_2} f_1(t_1,x)_x dt_1 \ \ dt_2 \ \ dt_3, \\ v_4(t,x) &= v_4(0,x) - tv_3'(0,x) - \frac{1}{2}t^2 v_2'(0,x) + \frac{1}{6}t^3 v_1''(0,x) \\ &\quad + \int_0^t \int_0^{t_4} \int_0^{t_3} \int_0^{t_2} f_1(t_1,x)_{xx} dt_1 \ \ dt_2 \ \ dt_3 \ \ dt_4. \end{aligned}$$

The second derivatives of the initial condition $v_1(0,x)$ and forcing function $f_1(t,x)$ appear in the solution, even though $\nu_{max} = 1$.

Each block in the canonical form can potentially increase the order of partial derivatives of data and forcing functions by ν, where ν is the degeneracy of the block. Let the *total degeneracy* be defined as the sum of the degeneracies of all blocks in the canonical form and be denoted by ν_{tot}. For linear systems with linear forcing, the coupling across blocks means that derivatives of up to order ν_{tot} may appear in the solution.

So, for a linear system with linear forcing, the canonical form makes several properties clear by inspection. The index of the system with respect to t is less than or equal to the dimension of the largest Jordan block of the parabolic part, while the index with respect to x is less than or equal to the dimension of the largest Jordan block of the differential part. These bounds become strict equalities if $\mathbf{C}$ is lower triangular. The solution depends continuously on its data iff $\nu_{max} = 0$ and the generalized eigenvalues are all strictly real. If n_1, n_2, and n_3 are the dimensions of the hyperbolic, differential, and parabolic parts of the system, respectively, then at least $n_1 + n_2$ initial and $n_1 + n_3$ boundary conditions are needed. Derivatives of up to order ν_{tot} of forcing functions and data may appear in the solution, so up to $C^{\nu_{tot}+1}$ continuity may be required in order for $\mathbf{u}$ to be C^1.

5. Examples. The generalized eigenvalues of $\mathbf{A}$, $\mathbf{B}$ give all the information needed for the analyses of the previous sections. All eigenvalues with $\rho, \tau \neq 0$ belong to the hyperbolic part. Eigenvalues with $\rho = 0$ belong to the differential part, and eigenvalues with $\tau = 0$ belong to the parabolic part. The maximum degeneracy of any eigenvalue in the differential and parabolic parts gives ν_1 and ν_2, respectively. Generalized eigenvalue pairs of the hyperbolic part for which ρ and τ have the same sign give rise to positive characteristics; those with differing signs correspond to negative characteristics.

Routines that calculate the generalized eigenvalues and their degeneracies for regular coefficient matrix pairs are readily available [4, 7]. For systems that include algebraic equations, symbolic differentiation may be performed automatically. If the coefficient matrix pair is singular, a generalized eigenvalue routine will return a zero eigenvalue pair $\rho = \tau = 0$.

5.1. Telegrapher's equations. Consider an electric circuit, where behavior along a transmission line is described by the telegrapher's equations:

$$\begin{bmatrix} 0 & L \\ C & 0 \end{bmatrix} \begin{bmatrix} u \\ I \end{bmatrix}_t + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ I \end{bmatrix}_x + \begin{bmatrix} 0 & R \\ G & 0 \end{bmatrix} \begin{bmatrix} u \\ I \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{5.1}$$

Here R, L, G, and C are positive constants that give the line resistance, inductance, conductance, and capacitance per unit length. u is the line voltage with respect to ground, and I is the line current.

In field effect transistor (FET) technology [18], line inductance, capacitance, and resistance are often neglected to produce the following simplified model:

$$\begin{bmatrix} 0 & 0 \\ C & 0 \end{bmatrix} \begin{bmatrix} u \\ I \end{bmatrix}_t + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ I \end{bmatrix}_x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{5.2}$$

Multiplying the first equation by C and introducing the new variable $v = Cu$ transform the system to its canonical form:

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} v \\ I \end{bmatrix}_t + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ I \end{bmatrix}_x = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{5.3}$$

The canonical form consists of a single, degenerate parabolic block. The forcing is simple (homogeneous). By Theorem 3.5, an unrestricted solution will not depend continuously on its data; it will in fact be weakly ill-posed. Determination of a unique solution requires specification of two boundary conditions, which may be imposed at either $x = x_1$ or $x = x_2$. Also, the nilpotency of $\mathbf{N}_2$ is two, so the index of the system with respect to t is 2, and there are no dynamic degrees of freedom on $t = 0$. This means that a restricted solution must be uniquely determined by its boundary conditions and must not leave any degrees of freedom for arbitrary specification of an initial condition.

5.2. Euler equations. One might also attempt application of the approach presented here for linear systems to linearizations of more general systems. Generalizing any results back to the original nonlinear system is, in general, not possible. However, performing the analysis on a local linearization may provide valuable insight into difficulties encountered with numerical solution of the original system.

Consider compressible flow through a plant piping network. The engineer might model flow through each pipe run by the well-studied [11, 20] Euler equations with the ideal gas law and might provide the following system of three partial differential and two algebraic equations in five unknowns to a process simulator:

$$(\rho)_t + (\rho u)_x = 0, \tag{5.4}$$
$$(\rho u)_t + \left(p + \rho u^2\right)_x = 0, \tag{5.5}$$
$$(\rho h)_t + (\rho u h - u p)_x = 0, \tag{5.6}$$
$$0 = p - (\gamma - 1)\,\rho i, \tag{5.7}$$
$$0 = i - h + \frac{1}{2}u^2. \tag{5.8}$$

Here ρ is the fluid density, u is the flow velocity, p is pressure, h is the specific total energy, and i is the specific internal energy. The first three model equations are conservation of mass, momentum, and energy, respectively. The fourth is the ideal gas law, with a constant fluid heat capacity ratio of γ. The final equation relates total, internal, and kinetic energy.

Let the initial and boundary conditions for the domain $0 \le x \le 10$, $t \ge 0$ be

$$\rho(0,x) = 79.6 \ \text{kg/m}^3, \tag{5.9}$$
$$u(0,x) = -50.0 \ \text{m/s}, \tag{5.10}$$
$$p(0,x) = 2.76 \ \text{MPa}, \tag{5.11}$$
$$p(t,0) = f_{valve}(t), \tag{5.12}$$
$$p(t,10) = f_{header}(t), \tag{5.13}$$
$$i(t,0) = g_{header}(t). \tag{5.14}$$

In quasi-linear form, the model equations are

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ u & \rho & 0 & 0 & 0 \\ h & 0 & 0 & \rho & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \rho \\ u \\ p \\ h \\ i \end{bmatrix}_t + \begin{bmatrix} u & \rho & 0 & 0 & 0 \\ u^2 & 2\rho u & 1 & 0 & 0 \\ uh & \rho h - p & -u & \rho u & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \rho \\ u \\ p \\ h \\ i \end{bmatrix}_x = \begin{bmatrix} 0 \\ 0 \\ 0 \\ p - (\gamma - 1)\,\rho i \\ i - h + \frac{1}{2}u^2 \end{bmatrix}. \tag{5.15}$$

Differentiating the ideal gas law and the energy equation produces

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ u & \rho & 0 & 0 & 0 \\ h & 0 & 0 & \rho & 0 \\ (1-\gamma)\,i & 0 & 1 & 0 & (1-\gamma)\,\rho \\ 0 & u & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} \rho \\ u \\ p \\ h \\ i \end{bmatrix}_t + \begin{bmatrix} u & \rho & 0 & 0 & 0 \\ u^2 & 2\rho u & 1 & 0 & 0 \\ uh & \rho h + p & u & \rho u & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \rho \\ u \\ p \\ h \\ i \end{bmatrix}_x = \mathbf{0}. \tag{5.16}$$

The system is quasi-linear, so the coefficient matrices must be frozen at a point of interest. Consider the domain boundary at $x = 10$, and let conditions at $x = 10$ be $\rho = 79.6$ kg/m^3, $u = -50.00$ m/s, $p = 2.76$ MPa, $h = 86.6$ kJ, and $i = 86.6$ kJ. The frozen coefficient matrices are submitted to an eigensolver, such as the LAPACK routine `dgegv`. The result is three characteristic directions parallel to the t coordinate axis, and two complex characteristic directions. The system is thus ill-posed in a neighborhood of these nominal values and cannot be solved by a simulator as part of a dynamic simulation.

The engineer is thus advised that the equations, as given to the simulator, are ill-posed. On review of the input, the sign error made in the energy balance (5.6) should be corrected:

$$(\rho h)_t + (\rho u h + u p)_x = 0. \tag{5.17}$$

The analysis may then be repeated for the corrected system:

$$(\rho)_t + (\rho u)_x = 0, \tag{5.18}$$

$$(\rho u)_t + \left(p + \rho u^2\right)_x = 0, \tag{5.19}$$

$$(\rho h)_t + (\rho u h + u p)_x = 0, \tag{5.20}$$

$$0 = p - (\gamma - 1)\, \rho i, \tag{5.21}$$

$$0 = i - h + \frac{1}{2}u^2. \tag{5.22}$$

Now, the degeneracy is found to be zero, so the index with respect to both t and x is zero. There are two characteristic directions parallel to the t coordinate axis and three with slopes -270.32, -50.00, and 170.32 m/s in the (t, x) plane. Two boundary conditions are therefore required at $x = 10$. The boundary conditions as input by the engineer do not determine a unique solution at this domain endpoint; again, there was an equation entry error. Assignment of the header fluid internal energy to the pipe fluid internal energy at $x = 0$ should be corrected to be enforced at $x = 10$. Upon correcting this second error, the resulting problem is well-posed.

6. Conclusion. This paper describes analyses of the well-posedness and index of a fairly general class of linear PDEs of first order that is based on a canonical form. These analyses include determination of proper initial and boundary data required to determine a unique solution, the effect of small errors in that data on the solution, and the index of the system with respect to each independent variable. It was shown that measures of the degeneracy ν_{max} and ν_{tot} provide bounds on the order of derivatives of forcing functions and data that may appear in the solution, rather than a differentiation index. These analyses may be performed automatically as part of a dynamic simulation package.

REFERENCES

[1] K. E. Brenan, S. L. Campbell, and L. R. Petzold, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, North-Holland, New York, 1989.

[2] S. L. Campbell and W. Marszalek, *The index of an infinite-dimensional implicit system*, Math. Comput. Model. Dyn. Syst., 5 (1999), pp. 18–42.

[3] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. 2, Interscience Publishers, New York, 1962.

[4] J. Demmel and B. Kågström, *The generalized Schur decomposition of an arbitrary pencil $A - \lambda B$: Robust software with error bounds and applications. Part* I: *Theory and algorithms*, ACM Trans. Math. Software, 19 (1993), pp. 160–174.

[5] F. R. Gantmacher, *The Theory of Matrices*, Chelsea Publishing Company, New York, 1959.

[6] C. W. Gear, *Differential-algebraic equation index transformations*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 39–47.

[7] G. Golub and C. van Loan, *Matrix Computations*, 2nd ed., The Johns Hopkins University Press, Baltimore, 1989.

[8] V. G. Grassi, *Dynamic simulation as a tool to integrate process design and process control*, in Advances in Instrumentation and Control, Vol. 48, Instrument Society of America, Chicago, 1993, pp. 1345–1365.

[9] M. Günther and Y. Wagner, *Index concepts for linear mixed systems of differential-algebraic and hyperbolic-type equations*, SIAM J. Sci. Comput., 22 (2000), pp. 1610–1629.

[10] E. Hairer, C. Lubich, and M. Roche, *The Numerical Solution of Differential-Algebraic Systems by Runge-Kutta Methods*, Springer-Verlag, New York, 1989.
[11] A. Jeffrey, *Quasilinear Hyperbolic Systems and Waves*, Pitman, London, 1976.
[12] H.-O. Kreiss and J. Lorenz, *Initial-Boundary Value Problems and the Navier-Stokes Equations*, Academic Press, New York, 1989.
[13] P. Kunkel and V. Mehrmann, *A new class of discretization methods for the solution of linear differential-algebraic equations with variable coefficients*, SIAM J. Numer. Anal., 33 (1996), pp. 1941–1961.
[14] E. J. Longwell, *Dynamic modeling for process control and operability*, in Advances in Instrumentation and Control, Vol. 48, Instrument Society of America, Chicago, 1993, pp. 1323–1336.
[15] W. Lucht and K. Debrabant, *On quasi-linear PDAEs with convection: Applications, indexes, numerical solution*, Appl. Numer. Math., 42 (2002), pp. 297–314.
[16] W. Lucht, K. Strehmel, and C. Eichler-Liebenow, *Indexes and special discretization methods for linear partial differential equations*, BIT, 39 (1999), pp. 484–512.
[17] W. S. Martinson and P. I. Barton, *A differentiation index for partial differential-algebraic equations*, SIAM J. Sci. Comput., 21 (2000), pp. 2295–2315.
[18] G. Massobrio and P. Antognetti, *Semiconductor Device Modeling with SPICE*, 2nd ed., McGraw-Hill, New York, 1993.
[19] L. Petzold, *Differential/algebraic equations are not ODE's*, SIAM J. Sci. Statist. Comput., 3 (1982), pp. 367–384.
[20] P. L. Roe, *Characteristic-based schemes for the Euler equations*, in Annual Review of Fluid Mechanics, Vol. 18, Annual Reviews, Palo Alto, CA, 1986, pp. 337–365.
[21] R. F. Sincovec, A. M. Erisman, E. L. Yip, and M. A. Epton, *Analysis of descriptor systems using numerical algorithms*, IEEE Trans. Automat. Control, 26 (1981), pp. 139–147.
[22] J. C. Strikwerda, *Finite Difference Schemes and Partial Differential Equations*, Wadsworth and Brooks/Cole Advanced Books and Software, Pacific Grove, CA, 1989.