A Geometric Approach to Linear Ordinary Differential Equations

R.C. Churchill
Hunter College and the Graduate Center of CUNY,
and the University of Calgary

Address for Correspondence
Department of Mathematics
Hunter College
695 Park Avenue, New York, NY 10021, USA

October 19, 2006

A Geometric Approach to Linear Ordinary Differential Equations

R.C. Churchill

Prepared for the

Kolchin Seminar on Differential Algebra

Graduate Center, City University of New York
13 October 2006

Amended and Corrected, 19 October 2006

Abstract

We offer a formulation of linear ordinary differential equations midway between what one encounters in a first undergraduate ODE course and what one encounters in a graduate Differential Geometry course (in the latter instance under the heading of "connections"). Analogies with elementary linear algebra are emphasized; no familiarity with Differential Geometry is assumed.

Contents

Introduction

Let V be a finite-dimensional vector space over a field K and let $T : V \to V$ be a K-linear operator. In a first course in linear algebra one learns that T can be studied via matrices as follows.

- Select an ordered basis $\mathbf{e} = (e_1, e_2, \ldots, e_n)$ of V and let $B = (b_{ij})$ be the $\mathbf{e}$-matrix of T, i.e., the $n \times n$ matrix with entries $b_{ij} \in K$ defined by

  (i) $$Te_j = \textstyle\sum_{j=1}^{n} b_{ij} e_i, \qquad j = 1, 2, \ldots, n.$$

- Let $\text{col}_n(K) \simeq K^n$ denote the K-space of $n \times 1$ column vectors of elements of K i.e., $n \times 1$ matrices, and let $\beta_{\mathbf{e}} : V \to \text{col}_n(K)$ be the isomorphism defined by

  (ii) $$v = \textstyle\sum_j k_j e_j \in V \qquad \mapsto \qquad v_e := \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} \in \text{col}_n(K).$$

- Let $T_B : \text{col}_n(K) \to \text{col}_n(K)$ denote left multiplication by B, i.e., the function $x \in \text{col}_n(K) \mapsto Bx \in \text{col}_n(K)$. Then T_B is K-linear and one has a commutative diagram

  (iii) $$\begin{array}{ccc} V & \xrightarrow{T} & V \\ {\scriptstyle \beta_{\mathbf{e}}}\downarrow & & \downarrow{\scriptstyle \beta_{\mathbf{e}}} \\ \text{col}_n(K) & \xrightarrow{T_B} & \text{col}_n(K) \end{array}$$

The commutativity of the diagram suggests that *anything one wants to know about T should somehow be reflected in the K-linear mapping T_B or, better yet, in the matrix B alone.* The hope is that information desired about T is more easily extracted from B.

Example : An ordered pair $(\lambda, v) \in F \times (V \setminus \{0\})$ is an *eigenpair* for T if

(iv) $$Tv = \lambda v;$$

one calls λ an *eigenvalue* of T, v an *eigenvector*, and the two entities are said to be "associated."

We interpret (iv) geometrically: the effect of T on v is to "scale" this vector by a factor of λ. The idea is easy enough to comprehend, but suppose one actually needs to compute such pairs?

In that case one switches from T to the matrix B: one learns that the eigenvalues of T are the roots of the characteristic polynomial

$$\mathrm{char}(B) := \det(xI - B)$$

of B, and once these have been determined one can begin searching for associated eigenvectors.

Of course the matrix B introduced above is not the unique "matrix representation" of T: a different choice of basis results in a different matrix, say A. One learns that A and B must be "similar," i.e., that $A = PBP^{-1}$ for some non-singular $n \times n$ matrix P, and that similar matrices have the same characteristic polynomial. This shows that *the characteristic polynomial is something intrinsically associated with T, not simply with matrix representations of T.* In particular, one can unambiguously define the *characteristic polynomial* $\mathrm{char}(T)$ *of* T by $\mathrm{char}(T) := \mathrm{char}(B)$. Since the determinant (up to sign) and the negative of the trace of B appear as coefficients of this polynomial, these entities are also directly associated with T, not simply with matrix representations thereof. One can therefore define $\det(T) := \det(B)$ and $\mathrm{tr}(T) := \mathrm{tr}(B)$, again without ambiguity.

To summarize: there are entities intrinsically associated with T which can easily be computed from matrix representations of this operator.

In these notes we will view linear ordinary differential equations in an analogous manner. Specifically, we will explain how to think of a first order system

$$\text{(v)} \qquad \begin{array}{rcccl} \dot{x}_1 & + & b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n & = & 0 \\ \dot{x}_2 & + & b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n & = & 0 \\ & \vdots & & & \\ \dot{x}_n & + & b_{n1}x_1 + b_{n2}x_2 + \cdots + b_{nn}x_n & = & 0 \end{array}$$

as a "basis representation" of a geometric entity which we call a "differential module." For our purposes the fundamental object intrinsically associated with such an entity is the "differential Galois group," a group traditionally associated with basis representations as in (v).

We need some preliminaries on derivations.

1. Preliminaries on Derivations

In this section R *is a* (*not necessarily commutative*) *ring with multiplicative identity* 1*, with* $1 = 0$ *allowed* (in which case $R = 0$).

An additive group endomorphism $\delta : r \in R \mapsto r' \in R$ is a *derivation* if the *product* or *Leibniz rule*

$$(rs)' = rs' + r's \tag{1.1}$$

holds for all $r, s \in R$. One also writes r' as $r^{(1)}$ and defines $r^{(n)} := (r^{(n-1)})'$ for $n \geq 2$. The notation $r^{(0)} := r$ proves convenient.

The usual differentiation operators $\frac{d}{dz}$ on the polynomial ring $\mathbb{C}[z]$ and the quotient field $\mathbb{C}(z)$ are the basic examples of derivations. The second can be viewed as a derivation on the field $\mathcal{M}(\mathbb{P}^1)$ of meromorphic functions on the Riemann sphere by means of the usual identification $\mathbb{C}(z) \simeq \mathcal{M}(\mathbb{P}^1)$. The same derivation on $\mathcal{M}(\mathbb{P}^1)$ is described in terms of the usual coordinate $t = 1/z$ at ∞ by $-t^2 \frac{d}{dt}$.

Another example of a derivation is provided by the zero mapping $r \in R \mapsto 0 \in R$; this is the *trivial derivation*.

For an example involving a non-commutative ring choose an integer $n > 1$, let R be the collection of $n \times n$ matrices with entries in a commutative ring A with a derivation $a \mapsto a'$, and for $r = (a_{ij}) \in R$ define $r' := (a'_{ij})$.

When $r \mapsto r'$ is a derivation on R one sees from (1.1) that $1' = (1 \cdot 1)' = 1 \cdot 1' + 1' \cdot 1 = 1' + 1'$, and as a result that

$$1' = 0\,. \tag{1.2}$$

When $r \in R$ is a unit it then follows from $1 = rr^{-1}$ and (1.1) that

$$0 = (rr^{-1})' = r \cdot (r^{-1})' + r' \cdot r^{-1},$$

whence

$$(r^{-1})' = -r^{-1} \cdot r' \cdot r^{-1}\,. \tag{1.3}$$

This formula is particularly useful in the matrix example given above. When R is commutative it assumes the more familiar form

$$(r^{-1})' = -r'r^{-2}\,. \tag{1.4}$$

The ring assumption suggests the generality of the concept of a derivation, but our main interest will be in derivations on fields. In this regard we note that any derivation on an integral domain extends uniquely, via the quotient rule, to the quotient field.

Henceforth K denotes a *differential field* (of characteristic 0), i.e., a field K equipped with a non-trivial derivation $k \mapsto k'$. By a *constant* we mean an element $k \in K$ satisfying $k' = 0$, e.g., we see from (1.2) that $1 \in R$ has this property. Indeed, the collection $K_C \subset K$ of constants is easily seen to be a subfield containing $\mathbb{Q}$; this is the *field of constants* (of $K = (K, \delta)$).

When $K = \mathbb{C}(z)$ with derivation $\frac{d}{dz}$ we have $K_C = \mathbb{C}$. The constants of the differential field $\mathcal{M}(\mathbb{P}^1)$ defined above are the constant functions $f : \mathbb{P}^1 \to \mathbb{C}$.

The determinant

$$(1.5) \qquad W := W(k_1, \ldots, k_n) := \det \begin{pmatrix} k_1 & k_2 & \cdots & & k_n \\ k_1' & k_2' & & \vdots & k_n' \\ k_1^{(2)} & k_2^{(2)} & & & \vdots \\ \vdots & & & & \\ k_1^{(n-1)} & k_2^{(n-1)} & \cdots & \cdots & k_n^{(n-1)} \end{pmatrix}$$

is the *Wronskian* of the elements $k_1, \ldots, k_n \in K$. This entity is useful for determining linear (in)dependence over K_C.

Proposition 1.6 : *Elements $k_1, \ldots, k_n$ of a differential field K are linearly dependent over the field of constants K_C if and only if their Wronskian is 0.*

Proof :

$\Rightarrow$ For any $c_1, \ldots, c_n \in K_C$ and any $0 \leq m \leq n$ we have $(\sum_j c_j k_j)^{(m)} = \sum_j c_j k_j^{(m)}$. In particular, when $\sum_j c_j k_j = 0$ the same equality holds when k_j is replaced by the j^{th} column of the Wronskian and 0 is replaced by a column of zeros. The forward assertion follows.

$\Leftarrow$ The vanishing of the Wronskian implies a dependence relation (over K) among columns, and as a result there must be elements $c_1, \ldots, c_n \in K$, not all 0, such that

$$\text{(i)} \qquad \sum_{j=1}^{n} c_j k_j^{(m)} = 0 \qquad \text{for} \qquad m = 0, \ldots, n-1.$$

What requires proof is that the c_j may be chosen in K_C, and this we establish by induction on n. As the case $n = 1$ is trivial we assume $n > 1$ and that the result holds for any subset of K with at most $n - 1$ elements.

If there is also a dependence relation (over K) among the columns of the Wronskian of $y_2, \ldots, y_n$, e.g., if $c_1 = 0$, then by the induction hypothesis the elements $y_2, \ldots, y_n \in K$ must be linearly dependent over K_C. But the same then holds for $y_1, \ldots, y_n$, which is precisely what we want to prove. We therefore assume (w.l.o.g.) that $c_1 = 1$ and that the columns of the Wronskian of $y_2, \ldots, y_k$ are linearly independent over K. From (i) we then have

$$0 = (\sum_{j=1}^{n} c_j k_j^{(m)})' = \sum_{j=1}^{n} c_j k_j^{(m+1)} + \sum_{j=2}^{n} c_j' k_j^{(m)} = 0 + \sum_{j=2}^{n} c_j' k_j^{(m)} = \sum_{j=2}^{n} c_j' k_j^{(m)}$$

for $m = 0, \ldots, n-2$, thereby forcing $c_2' = \cdots = c_n' = 0$. But this means $c_j \in K_C$ for $j = 1, \ldots, n$, and the proof is complete. **q.e.d.**

2. Differential Modules

Throughout this section K *denotes a differential field with derivation* $k \mapsto k'$ *and* V *is a* K*-space* (i.e., a vector space over K). *The collection of* $n \times n$ *matrices with entries in* K *is denoted* $\mathfrak{gl}(n, K)$*; the group of invertible matrices in* $\mathfrak{gl}(n, K)$ *is denoted* $\mathrm{GL}(n, k)$.

A *differential structure* on V is an additive group homomorphism $D : V \to V$ satisfying

$$D(kv) = k'v + kDv, \qquad k \in K, \qquad v \in V, \tag{2.1}$$

where Dv abbreviates $D(v)$. The *Leibniz rule* terminology is also used with (2.1). Vectors $v \in V$ satisfying $Dv = 0$ are said to be *horizontal*[1]. The zero vector $0 \in V$ is always has this property; other such vectors need not exist.

When $D : V \to V$ is a differential structure the pair (V, D) is called a *differential* K*-module*, or simply a *differential module* when K is clear from context. When $\dim_K(V) = n < \infty$ the integer n is the *dimension* of the differential module.

For an example of a differential structure/module take $V := \mathrm{col}_n(K)$ and define $D : \mathrm{col}_n(K) \to \mathrm{col}_n(K)$ by

$$D \begin{pmatrix} k_1 \\ k_2 \\ \vdots \\ k_n \end{pmatrix} := \begin{pmatrix} k_1' \\ k_2' \\ \vdots \\ k_n' \end{pmatrix}. \tag{2.2}$$

Further examples will be evident from Proposition 2.12.

Since K_C is a subfield of K we can regard V as a vector space over K_C by restricting scalar multiplication to $K_C \times V$.

Proposition 2.3 :

(a) *Any differential structure* $D : V \to V$ *is* K_C*-linear.*

(b) *The collection of horizontal vectors of a differential structure* $D : V \to V$ *coincides with the kernel* $\ker D$ *of* D *when* D *is considered as a* K_C*-linear mapping.*

[1] The terminology is borrowed from Differential Geometry.

(c) *The horizontal vectors of a differential structure* $D : V \to V$ *constitute a* K_C-*subspace of* (the K_C-space) V.

Proof :

(a) Immediate from (2.1).

(b) Obvious from the definition of horizontal.

(c) Immediate from (b).

q.e.d.

To obtain a basis description of a differential structure $D : V \to V$ let $\mathbf{e} = (e_j)_{j=1}^n \subset V^n$ be a(n ordered) basis of V let and $B = (b_{ij}) \in \mathfrak{gl}(n, K)$ be defined by

$$De_j := \sum_{j=1}^n b_{ij} e_i, \qquad j = 1, \dots, n. \tag{2.4}$$

(Example: For D as in (2.2) and[2] $e_j = (0, \dots, 0, 1, 0, \dots, 1)^\tau$ [1 in slot j] for $j = 1, \dots, n$ we have $B = (0)$ [the zero matrix].) We refer to B as the *defining* ($\mathbf{e}$)-*matrix* of D, or as the *defining matrix of* D *relative to the basis* $\mathbf{e}$. Note that, for any $v = \sum_{j=1}^n v_j e_j \in V$, additivity and the Leibniz rule (2.1) give

$$Dv = \sum_{i=1}^n (v_i' + \sum_{j=1}^n b_{ij} v_j) e_i. \tag{2.5}$$

This is better expressed in the matrix form

$$(Dv)_{\mathbf{e}} = v_{\mathbf{e}}' + Bv_{\mathbf{e}}, \tag{2.6}$$

wherein $v_{\mathbf{e}}$ is as in (ii) of the introduction, i.e., the image of v under the isomorphism

$$\beta_{\mathbf{e}} : v = \sum_j v_j e_j \in V \qquad \mapsto \qquad \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathrm{col}_n(K),$$

and

$$v_{\mathbf{e}}' := \begin{pmatrix} v_1' \\ v_2' \\ \vdots \\ v_n' \end{pmatrix}.$$

[2] The superscript τ ("tau") denotes transposition.

Where all this is leading should hardly be a surprise: the mapping $D_B : x \in \mathrm{col}_n(K) \to x' + Bx \in \mathrm{col}_n(K)$ defines a differential structure on the K-space $\mathrm{col}_n(K)$, and one has a commutative diagram

$$\begin{array}{ccc} V & \xrightarrow{D} & V \\ {\scriptstyle \beta_{\mathbf{e}}}\downarrow & & \downarrow{\scriptstyle \beta_{\mathbf{e}}} \\ \mathrm{col}_n(K) & \xrightarrow{D_B} & \mathrm{col}_n(K). \end{array} \tag{2.7}$$

An immediate connection with linear ordinary differential equations is seen from (2.6): *a vector* $v \in V$ *is horizontal if and only if* $v_{\mathbf{e}} \in \mathrm{col}_n(K)$ *is a solution of the first-order linear system*

$$x' + Bx = 0. \tag{2.8}$$

This is the *defining* ($\mathbf{e}$)-*equation* of D.

Linear systems of ordinary differential equations of the form

$$x' + Bx = 0 \tag{2.9}$$

are called *homogeneous*. One can also ask for solutions of *inhomogeneous* systems, i.e., systems of the form

$$x' + Bx = b, \tag{2.10}$$

wherein $0 \neq b \in \mathrm{col}_n(K)$ is given. For $b = w_{\mathbf{e}}$ this is equivalent to the search for a vector $v \in V$ satisfying

$$Dv = w\,. \tag{2.11}$$

Equation (2.9) is the *homogeneous equation corresponding to* (2.10).

Proposition 2.12 : *When* $\dim_K V < \infty$ *and* $\mathbf{e}$ *is a basis the correspondence between differential structures* $D : V \to V$ *and* $n \times n$ *matrices* B *defined by* (2.4) *is bijective; the inverse assigns to a matrix* $B \in \mathfrak{gl}(n, K)$ *the differential structure* $D : V \to V$ *defined by* (2.6).

Proof : The proof is by routine verification. **q.e.d.**

Proposition 2.13 :

(a) *The solutions of* (2.9) *within* $\mathrm{col}_n(K)$ *form a vector space over* K_C.

(b) *When* $\dim_K V = n < \infty$ *and* $\mathbf{e}$ *is a basis of* V *the* K*-linear isomorphism* $v \in V \mapsto v_{\mathbf{e}} \in \mathrm{col}_n(K)$ *restricts to a* K_C*-linear isomorphism between the* K_C*-subspace of* V *consisting of horizontal vectors and the* K_C*-subspace of* $\mathrm{col}_n(K)$ *consisting of solutions of* (2.9).

Proof :

(a) When $y_1, y_2 \in \mathrm{col}_n(K)$ are solutions and $c_1, c_2 \in K_C$ we have

$$\begin{aligned}(c_1y_1 + c_2y_2)' &= (c_1y_1)' + (c_2y_2)' \\ &= c_1'y_1 + c_1y_1' + c_2'y_2 + c_2y_2' \\ &= 0 \cdot y_1 + c_1(-By_1) + 0 \cdot y_2 + c_2(-By_2) \\ &= -Bc_1y_1 - Bc_2y_2 \\ &= -B(c_1y_1 + c_2y_2)\,.\end{aligned}$$

(b) That the mapping restricts to a bijection between horizontal vectors and solutions was already noted immediately before (2.9), and since the correspondence $v \mapsto v_{\mathbf{e}}$ is K-linear and K_C is a subfield of K any restriction to a K_C-subspace must be K_C-linear.

q.e.d.

Suppose $\hat{\mathbf{e}} = (\hat{e}_j)_{j=1}^n \subset V^n$ is a second basis and $P = (p_{ij})$ is the transition matrix, i.e., $\hat{e}_j = \sum_{i=1}^n p_{ij}e_i$. Then the defining $\mathbf{e}$ and $\hat{\mathbf{e}}$-matrices B and A of D are easily seen to be related by

(2.14) $$A := PBP^{-1} - P'P^{-1},$$

where $P' := (p_{ij}')$. The transition from B to A is viewed classically as a change of variables: substitute $w = Px$ in (2.9); then note from

$$w' = Px' + P'x = P(-Bx) + P'P^{-1}w = -PBP^{-1}w + P'P^{-1}w$$

that

$$w' + (PBP^{-1} - P'P^{-1})w = 0\,.$$

The modern viewpoint is to regard $(P, B) \mapsto PBP^{-1} - P'P^{-1}$ as defining a left action of $GL(n, K)$ on $\mathfrak{gl}(n, K)$; this is the action by *gauge transformations*.

Example 2.15 : Assume $K = \mathbb{C}(z)$ with derivation $\frac{d}{dz}$ and consider the first-order system

$$\text{(i)} \qquad x' + \begin{pmatrix} \left(\frac{10z^4-(2\nu^2-1)z^2-2}{z(2z^4-1)}\right) & \frac{4z^6-4\nu^2z^4-4z^2+1}{z^4(2z^4-1)} \\ -\left(\frac{z^2(z^4+3z^2-\nu^2)}{2z^4-1}\right) & \frac{(2\nu^2-1)z^2+1}{z(2z^4-1)} \end{pmatrix} x = 0\,,$$

i.e.,

$$\begin{array}{lclclcl} x_1' & + & \left(\frac{10z^4-(2\nu^2-1)z^2-2}{z(2z^4-1)}\right) x_1 & + & \left(\frac{4z^6-4\nu^2z^4-4z^2+1}{z^4(2z^4-1)}\right) x_2 & = & 0 \\ x_2' & - & \left(\frac{z^2(z^4+3z^2-\nu^2)}{2z^4-1}\right) x_1 & + & \left(\frac{(2\nu^2-1)z^2+1}{z(2z^4-1)}\right) x_2 & = & 0 \end{array} ,$$

wherein ν is a complex parameter. This has the form (2.9) with

$$B := \begin{pmatrix} \left(\frac{10z^4-(2\nu^2-1)z^2-2}{z(2z^4-1)}\right) & \frac{4z^6-4\nu^2z^4-4z^2+1}{z^4(2z^4-1)} \\ -\left(\frac{z^2(z^4+3z^2-\nu^2)}{2z^4-1}\right) & \frac{(2\nu^2-1)z^2+1}{z(2z^4-1)} \end{pmatrix} ,$$

and with the choice[3]

$$P := \begin{pmatrix} z^2 & 2z \\ z^3 & z^{-2} \end{pmatrix}$$

one sees that the transformed system is

$$\text{(ii)} \qquad x' + Ax = 0, \qquad \text{where} \qquad A := PBP^{-1} - P'P^{-1} = \begin{pmatrix} 0 & -1 \\ 1 - \frac{\nu^2}{z^2} & \frac{1}{z} \end{pmatrix} .$$

We regard (i)-(ii) as distinct basis descriptions of the same differential structure.

[3] At this point readers should not be concerned with how this particular P was constructed.

3. n^{th}-Order Linear Differential Equations

The concept of an n^{th}-order linear homogeneous equation in the context of a differential field K is formulated in the obvious way: an element $k \in K$ is a solution of

$$y^{(n)} + \ell_1 y^{(n-1)} + \cdots + \ell_{n-1} y' + \ell_n y = 0\,, \tag{3.1}$$

where $\ell_1, \ldots, \ell_n \in K$, if and only if

$$k^{(n)} + \ell_1 k^{(n-1)} + \cdots + \ell_{n-1} k' + \ell_n k = 0\,, \tag{3.2}$$

where $k^{(2)} := k'' := (k')'$ and $k^{(j)} := (k^{(j-1)})'$ for $j > 2$. Using a Wronskian argument one can easily prove that (3.1) has at most n solutions (in K) linearly independent over K_C.

As in the classical case $k \in K$ is a solution of (3.1) if and only if the column vector $(k, k', \ldots, k^{(n-1)})^\tau$ is a solution of

$$x' + Bx = 0, \qquad B = \begin{pmatrix} 0 & -1 & 0 & \cdots & & 0 \\ \vdots & 0 & -1 & & & \vdots \\ & & 0 & \ddots & & \\ & & & \ddots & -1 & 0 \\ & & & & 0 & -1 \\ \ell_n & \ell_{n-1} & \cdots & \cdots & \ell_2 & \ell_1 \end{pmatrix}. \tag{3.3}$$

Indeed, one has the following analogue of Proposition 2.13.

Proposition 3.4 :

(a) *The solutions of* (3.1) *within* K *form a vector space over* K_C.

(b) *The* K_C*-linear mapping* $(y, y', \ldots, y^{(n-1)})^\tau \in \mathrm{col}_n(K) \mapsto y \in K$ *restricts to a* K_C*-linear isomorphism between the* K_C*-subspace of* V *consisting of horizontal vectors and the* K_C*-subspace of* K *described in* (a).

Proof : The proof is a routine verification. **q.e.d.**

Example 3.5 : Equation (ii) of Example 2.15, i.e.,

$$\text{(i)} \qquad x' + \begin{pmatrix} 0 & -1 \\ 1 - \frac{\nu^2}{z^2} & \frac{1}{z} \end{pmatrix} x = 0,$$

has the form seen in (3.3). This linear differential equation would commonly be written as

$$\text{(ii)} \qquad y'' + \frac{1}{z}\, y' + \left(1 - \frac{\nu^2}{z^2}\right) y = 0,$$

or as

$$\text{(iii)} \qquad z^2 y'' + z\, y + (z^2 - \nu^2)\, y = 0.$$

Either form can be regarded as a basis description of the differential module of that example. Some readers will immediately recognize (iii): it is *Bessel's equation*.

Converting the n^{th}-order equation (3.1) to the first-order system (3.3) is standard practice. Less well-known is the fact that any first-order system of n equations can be converted to the form (3.3), and as a consequence can be expressed n^{th}-order form. We will prove this in §7. For many purposes n^{th}-order form has distinct advantages, e.g., explicit solutions are often easily constructed with series expansions, e.g., equation (iii) of Example 3.5 can be solved explicitly in terms of Bessel functions.

4. Dimension Considerations

In this section K denotes a differential field and (V, D) is a differential K-module of dimension $n \geq 1$.

Proposition 4.1 : *When V is a K-space with differential structure $D : V \to V$ the following assertions hold.*

(a) *A collection of horizontal vectors within V is linearly independent over K if and only if it is linearly independent over K_C.*

(b) *The collection of horizontal vectors of V is a vector space over K_C of dimension at most n.*

Proof :

(a) $\Rightarrow$ Immediate from the inclusion $K_C \subset K$. (In this direction the horizontal assumption is unnecessary.)

$\Leftarrow$ If the implication is false there is a collection of horizontal vectors in V which is K_C-(linearly) independent but K-dependent, and from this collection we can choose vectors $v_1, \ldots, v_m$ which are K-dependent with $m > 1$ minimal w.r.t. this property. We can then write $v_m = \sum_{j=1}^{m-1} k_j v_j$, with $k_j \in K$, whereupon applying D and the hypotheses $Dv_j = 0$ results in the identity $0 = \sum^{m-1} k_j' v_j$. By the minimality of m this forces $k_j' = 0$, $j = 1, \ldots, m-1$, i.e., $k_j \in K_C$, and this contradicts linear independence over K_C.

(b) This is immediate from (a) and the fact that any K-linearly independent subset of V can be extended to a basis.

q.e.d.

Suppose $\dim_K V = n < \infty$, $\mathbf{e}$ is a basis of V, and $x' + Ax = 0$ is the defining $\mathbf{e}$-equation of D. Then assertion (b) of the preceding result has the following standard formulation.

Corollary 4.2 : *For any matrix $B \in \mathfrak{gl}(n, K)$ a collection of solutions of*

(i) $$x' + Bx = 0$$

within $\mathrm{col}_n(K)$ is linearly independent over K if and only if the collection is linearly independent over K_C. In particular, the K_C-subspace of $\mathrm{col}_n(K)$ consisting of solutions of (i) has dimension at most n.

Proof : By Proposition 2.13. **q.e.d.**

Equation (i) of Corollary 4.2 is always satisfied by the column vector $x = (0, 0, \ldots, 0)^\tau$; this is the *trivial solution*, and any other is *non-trivial*. Unfortunately, non-trivial solutions (with entries in K) need not exist. For example, the linear differential equation $y' - y = 0$ admits only the trivial solution in the field $\mathbb{C}(z)$: for non-trivial solutions one must recast the problem so as to include the extension field $(\mathbb{C}(z))(\exp(z))$.

Corollary 4.3 : *For any elements* $\ell_1, \ldots, \ell_{n-1} \in K$ *a collection of solutions* $\{y_j\}_{j=1}^m \subset K$ *of*

$$\text{(i)} \qquad y^{(n)} + \ell_1 y^{(n-1)} + \cdots + \ell_{n-1} y' + \ell_n y = 0$$

is linearly independent over K_C *if and only if the collection* $\{(y_j, y_j', \ldots, y_j^{(n-1)}\}_{j=1}^m$ *is linearly independent over* K*. In particular, the* K_C*-subspace of* K *consisting of solutions of* (i) *has dimension at most* n*.*

Proof : Use Proposition 3.4(b) and Corollary 4.2. **q.e.d.**

5. Fundamental Matrix Solutions

Throughout the section K is a differential field and $D : V \to V$ is a differential module of positive dimension n.

Choose any basis $\mathbf{e}$ of V and let

$$x' + Bx = 0 \tag{5.1}$$

be the defining $\mathbf{e}$-equation of D. A non-singular matrix $M \in \mathfrak{gl}(n, K)$ is a *fundamental matrix solution* of (5.1) if M satisfies this equation, i.e., if and only if

$$M' + BM = 0. \tag{5.2}$$

Example: Take $K = \mathbb{C}(z)$ with the usual derivation $\delta = d/dz$; then

$$M := \begin{pmatrix} z^2 & 1+z \\ 0 & z^3 \end{pmatrix}$$

is a fundamental matrix solution of the equation

$$x' + \begin{pmatrix} -\frac{2}{z} & \frac{z+2}{z^4} \\ 0 & -\frac{3}{z} \end{pmatrix} x = 0.$$

Proposition 5.3 : *A matrix $M \in \mathfrak{gl}(n, K)$ is a fundamental matrix solution of* (5.1) *if and only if the columns of M constitute n solutions of that equation linearly independent over K_C.*

Of course linear independence over K is equivalent to the non-vanishing of the Wronskian $W(y_1, \ldots, y_n)$.

Proof : First note that (5.2) holds if and only if the columns of M are solutions of (5.1). Next observe that M is non-singular if and only if these columns are linearly independent over K. Finally, note from Propositions 2.13(b) and 4.1(a) that this will be the case if and only if these columns are linearly independent over K_C. **q.e.d.**

Corollary 5.4 : *Equation* (5.1) *admits a fundamental matrix solution in* $\mathrm{GL}(n, K)$ *if and only if V admits a basis consisting of D-horizontal vectors.*

Proof : By Proposition 5.3 the existence of a fundamental matrix solution in $\mathrm{GL}(n, K)$ is equivalent to the existence of n D-horizontal vectors linearly independent over K_C. Now recall Proposition 4.1. **q.e.d.**

Proposition 5.5 : *Suppose $M, N \in \mathfrak{gl}(n, K)$ and M is a fundamental matrix solution of* (5.1). *Then N is a fundamental matrix solution if and only if $N = MC$ for some matrix $C \in \mathrm{GL}(n, K_C)$.*

Proof : B

$\Rightarrow$: By (1.3) we have

$$\begin{aligned}(M^{-1}N)' &= M^{-1} \cdot N' = (M^{-1})' \cdot N \\ &= M^{-1} \cdot (-BN) + (-M^{-1}M'M^{-1}) \cdot N \\ &= -M^{-1}BN + (-M^{-1})(-BM)(-M^{-1})N \\ &= -M^{-1}BN + M^{-1}BN \\ &= 0\,.\end{aligned}$$

$\Leftarrow$: We have $N' = (MC)' = M'C = -BM \cdot C = -B \cdot MC = -BN$.

q.e.d.

Fundamental matrix solutions have an interesting geometric characterization. To explain that we need to introduce the dual of a differential module.

6. Dual Structures and Adjoint Equations

Differential structures allow for a simple conceptual formulation of the "adjoint equation" of a linear ordinary differential equation.

In this section K is a differential field and (V, D) is a differential K-module of dimension $n \geq 1$. Recall that the dual space V^* of V is defined as the K-space of linear functionals $v^* : V \to K$, and the dual basis $\mathbf{e}^*$ of a basis $\mathbf{e} = \{e_\alpha\}$ of V is the basis $\{e_\alpha^*\}$ of V^* satisfying $e_\beta^* e_\alpha = \delta_{\alpha\beta}$ (wherein $\delta_{\alpha\beta}$ is the usual Kronecker delta, i.e., $\delta_{\alpha\beta} := 1$ if and only if $\alpha = \beta$; otherwise $\delta_{\alpha\beta} := 0$).

There is a dual differential structure $D^* : V^* \to V^*$ on the dual space V^* naturally associated with D: the definition is

$$(D^*u^*)v = \delta(u^*v) - u^*(Dv), \qquad u^* \in V^*,\ v \in V. \tag{6.1}$$

The verification that this is a differential structure is straightforward, and is left to the reader. One often sees u^*v written as $\langle v, u^* \rangle$, and when this notation is used (6.1) becomes

$$\delta\langle v, u^*\rangle = \langle Dv, u^*\rangle + \langle v, D^*u^*\rangle. \tag{6.2}$$

This is the *Lagrange identity*; it implies that $u^*v \in K_C$ whenever v and u^* are horizontal.

Proposition 6.3 : *Suppose $\mathbf{e} \subset V^n$ is a basis of V and $B = (b_{ij}) \in \mathfrak{gl}(n, K)$ is the defining $\mathbf{e}$-matrix of D. Then the defining $\mathbf{e}^*$-matrix of D^* is $-B^\tau$.*

The proof, as are most of the proofs in this section, is a simple application of the Lagrange identity.

Proof : First note from (6.2) that for any $1 \leq i, j \leq n$ we have

$$\begin{aligned} 0 &= \delta_{ij}' \\ &= \delta\langle e_i, e_j^*\rangle \\ &= \langle De_i, e_j^*\rangle + \langle e_i, D^*e_j^*\rangle \\ &= \langle \textstyle\sum_k b_{ki}e_k, e_j^*\rangle + \langle e_i, D^*e_j^*\rangle \\ &= \textstyle\sum_k b_{ki}\langle e_k, e_j^*\rangle + \langle e_i, D^*e_j^*\rangle \\ &= b_{ji} + \langle e_i, D^*e_j^*\rangle, \end{aligned}$$

from which we see that

(i) $$\langle e_i, D^* e_j^* \rangle = -b_{ji}.$$

However, for the defining $\mathbf{e}^*$-matrix $C = (c_{ij})$ of D^* we have

$$D^* e_i^* = \sum_k c_{ki} e_k^*,$$

and therefore

$$\langle e_i, D^* e_j^* \rangle = \langle e_i, \sum_k c_{kj} e_k^* \rangle = \sum_k c_{kj} \langle e_i, e_k^* \rangle = c_{ij}$$

for any $1 \leq i, j \leq n$. The result now follows from (i). **q.e.d.**

As an immediate consequence of Proposition 6.3 we see that the defining $\mathbf{e}^*$-equation of D^* is

(6.4) $$y' - B^\tau y = 0;$$

this is the *adjoint equation* of

(6.5) $$x' + Bx = 0.$$

Note from the usual identification $V^{**} \simeq V$ and $-(-B^\tau)^\tau = B$ that (6.5) can be viewed as the adjoint equation of (6.4). Intrinsically: the identification $V \simeq V^{**}$ induces the additional identification $D^{**} := (D^*)^* \simeq D$.

Equations (6.5) and (6.4) are interchangeable in the sense that information about either one can always be obtained from information about the other. In particular, there is fundamental relationship between solutions of a linear differential equation and the solutions of the adjoint equation. This is explained in the following result, which also contains the promised geometric characterization of a fundamental matrix solution.

Proposition 6.6 : *Suppose* $\mathbf{e}$ *is a basis of* V *and*

(i) $$x' + Bx = 0$$

is the defining $\mathbf{e}$*-matrix of* D*. Then for any* $M \in \mathfrak{gl}(n, K)$ *the following statements are equivalent:*

(a) M *is a fundamental matrix solution of* (i)*;*

(b) $(M^\tau)^{-1}$ *is a fundamental matrix solution of the adjoint equation*

(ii) $$x' - B^\tau x = 0$$

of (i)*; and*

(c) M^τ *is the transition matrix from the dual basis* $\mathbf{e}^*$ *of* $\mathbf{e}$ *to a basis of* V^* *consisting of* D^**-horizontal vectors.*

Proof : First note that $(M^\tau)' = (M')^\tau$; hence that

(iii) $$M' + BM = 0 \qquad \Leftrightarrow \qquad (M^\tau)' + M^\tau B^\tau = 0.$$

(a) $\Leftrightarrow$ (b) : From (iii), $(M^\tau)^{-1} = (M^{-1})^\tau$ and $(M^\tau)' = (M')^\tau$ we have

$$\begin{aligned} M' + MB = 0 \;&\Leftrightarrow\; (M')^\tau + M^\tau B^\tau = 0 \\ &\Leftrightarrow\; -(M^\tau)^{-1}(M')^\tau(M^\tau)^{-1} + (M^\tau)^{-1}M^\tau B^\tau(M^\tau)^{-1} = 0 \\ &\Leftrightarrow\; -(M^{-1})^\tau(M')^\tau(M^{-1})^\tau + B^\tau(M^{-1})^\tau = 0 \\ &\Leftrightarrow\; -(M^{-1}M'M^{-1})^\tau + B^\tau(M^{-1})^\tau = 0 \\ &\Leftrightarrow\; ((M^{-1})')^\tau - B^\tau(M^\tau)^{-1} = 0 \\ &\Leftrightarrow\; ((M^\tau)^{-1})' - B^\tau(M^\tau)^{-1} = 0\,. \end{aligned}$$

(a) $\Leftrightarrow$ (c) : One has

$$\begin{aligned} M' + BM = 0 \;&\Leftrightarrow\; (M^\tau)' + M^\tau B^\tau = 0 \\ &\Leftrightarrow\; (M^\tau)'(M^\tau)^{-1} + M^\tau B^\tau(M^\tau)^{-1} = 0 \\ &\Leftrightarrow\; (M^\tau)(-B^\tau)(M^\tau)^{-1} - (M^\tau)'(M^\tau)^{-1} = 0. \end{aligned}$$

This last line is precisely what one obtains when A, B and P in (2.14) are replaced by 0, $-B^\tau$ and M^τ respectively, i.e., it asserts that when M^τ is regarded as a transition matrix from $\mathbf{e}^*$ to some other basis $\hat{\mathbf{e}}^*$ of V^*, the defining $\hat{\mathbf{e}}^*$-matrix of D^* must vanish. Now simply observe from (2.4) that that the defining matrix of a basis vanishes if and only if that basis consists of horizontal vectors.

q.e.d.

Dual differential structures are useful for solving equations of the form $Du = w$, wherein $w \in V$ is given. The fundamental result in this direction is the following.

Proposition 6.7 : *Suppose* $(v_m^*)_{j=1}^n$ *is a basis of* V^* *consisting of horizontal vectors and* $(v_j)_{j=1}^n$ *is the dual basis of* $V \simeq V^{**}$*. Then the following statements hold.*

(a) *All* v_j *are horizontal.*

(b) *Suppose* $w \in V$ *and there are elements* $k_j \in K$ *such that* $k_j' = \langle w, v_j^* \rangle$ *for* $j = 1, \ldots, n$*. Then the vector*

(i) $$u := \textstyle\sum_j k_j v_j \in V$$

satisfies

(ii) $$Du = w\,.$$

Te vector u introduced in (i) is not the unique solution of (ii): the sum of u and any horizontal vector will also satisfy that equation.

Proof :

(a) This can be seen as a corollary of Proposition 6.6, but a direct proof is quite easy: from $\langle v_i, v_j^* \rangle \in \{0,1\} \subset K_C$, Lagrange's identity (6.2) and the hypothesis $D^*v_j^* = 0$ we have

$$\begin{aligned} 0 &= \langle v_i, v_j^* \rangle' \\ &= \langle Dv_i, v_j^* \rangle + \langle v_i, D^*v_j^* \rangle \\ &= \langle Dv_i, v_j^* \rangle\,, \end{aligned}$$

and since (v_j^*) is a basis this forces $Dv_i = 0$ for $i = 1, \ldots, n$.

(b) First note that for any $v \in V$ we have

(iii) $$v = \textstyle\sum_j \langle v, v_j^* \rangle v_j\,.$$

Indeed, we can always write v in the form $v = \sum_i c_i v_i$, where $c_i \in K$, and applying v_j^* to this identity gives $\langle v, v_j^* \rangle = \sum_i c_i \langle v_i, v_j^* \rangle = c_j$.

From (a) and (iii) we then have

$$\begin{aligned} Du &= \textstyle\sum_j D(k_j v_j) \\ &= \textstyle\sum_j (k_j' v_j + k_j D v_j) \\ &= \textstyle\sum_j (\langle w, v_j^* \rangle v_j + k_j \cdot 0) \\ &= \textstyle\sum_j \langle w, v_j^* \rangle v_j \\ &= w\,. \end{aligned}$$

q.e.d.

The following corollary explains the relevance of the adjoint equation for solving inhomogeneous systems. In the statement we adopt more classical notation: when $k, \ell \in K$ satisfy $\ell' = k$ we write ℓ as $\int k$, omit specific reference to ℓ, and simply assert that $\int k \in K$. Moreover, we use the usual inner product $\langle y, z\rangle := \sum_j y_j z_j$ to identify $\mathrm{col}_n(K)$ with $(\mathrm{col}_n(K))^*$, i.e., we identify the two spaces by means of the K-isomorphism $v \in \mathrm{col}_n(K) \mapsto (w \in \mathrm{col}_n(K) \mapsto \langle w, v\rangle \in K) \in (\mathrm{col}_n(K))^*$.

Corollary 6.8 [4]**:** *Suppose:*

(a) $B \in \mathfrak{gl}(n, K)$;

(b) $b \in \mathrm{col}_n(K)$;

(c) $(z_j)_{j=1}^n$ *is a basis of* $\mathrm{col}_n(K)$ *consisting of solutions of the adjoint equation*

(i) $$x' - B^\tau x = 0$$

of

(ii) $$x' + Bx = 0\,;$$

(d) $\int \langle b, z_j\rangle \in K$ *for* $j = 1, \ldots, n$; *and*

(e) $(y_i)_{i=1}^n$ *is a basis of* $\mathrm{col}_n(K)$ *satisfying*

(iii) $$\langle y_i, z_j\rangle = \begin{cases} 1 & \text{if } \; i = j \\ 0 & \text{otherwise.} \end{cases}$$

Then $(y_j)_{j=1}^n$ *is a basis of solutions of the homogeneous equation* (ii) *and the vector*

(iv) $$y := \textstyle\sum_j (\int \langle b, z_j\rangle) \cdot y_j$$

is a solution of the inhomogeneous system

(v) $$x' + Bx = b\,.$$

[4] For a classical account of this result see, e.g., [Poole, Chapter III, §10, pp. 36-39]. In fact the treatment in this reference was the inspiration for our formulation of this corollary.

The appearance of the integrals in (iv) explains why solutions of (i) are called *integrating factors* of (ii) (and vice-versa, since, as already noted, (i) may be regarded as the adjoint equation of (ii)).

The result is immediate from Proposition 6.7. However, it is a simple enough matter to give a direct proof, and we therefore do so.

Proof : Hypothesis (d) identifies $(z_j)_{j=1}^n$ with the dual basis of $(y_i)_{i=1}^n$. In particular, it allows us to view $(z_j)_{j=1}^n$ as a basis of $(\mathrm{col}_n(K))^*$.

To prove that the x_j satisfy (ii) simply note from (iii) and (i) that

$$\begin{aligned} 0 &= \langle y_i, z_j\rangle' \\ &= \langle y_i', z_j\rangle + \langle y_i, z_j'\rangle \\ &= \langle y_i', z_j\rangle + \langle y_i, B^\tau z_j\rangle \\ &= \langle y_i', z_j\rangle + \langle By_i, z_j\rangle \\ &= \langle y_i' + By_i, z_j\rangle. \end{aligned}$$

Since $(z_j)_{j=1}^n$ is a basis (of $(\mathrm{col}_n(K))^*$) it follows that

(vi) $$y_j' + By_j = 0\,, \qquad j = 1, \ldots, n.$$

Next observe, as in (i) of the proof of Proposition 6.7, that for any $b \in \mathrm{col}_n(K)$ condition (iii) implies

$$b = \sum_j \langle b, z_j\rangle y_j\,.$$

It then follows from (vi) that

$$\begin{aligned} y' &= \textstyle\sum_j \big((\int\langle b, z_j\rangle)\cdot y_j' + \langle b, z_j\rangle y_j\big) \\ &= \textstyle\sum_j \big(-(\int\langle b, z_j\rangle)\cdot By_j + \langle b, z_j\rangle y_j\big) \\ &= -B\textstyle\sum_j(\int\langle b, z_j\rangle)\cdot y_j + \sum_j\langle b, z_j\rangle y_j \\ &= -By + b\,. \end{aligned}$$

q.e.d.

Corollary 6.8 was formulated so as to make the role of the adjoint equation evident. The following alternate formulation is easier to apply in practice.

Corollary 6.9 : *Suppose $B \in \mathfrak{gl}(n, K)$ and $M \in \mathrm{GL}(n, K)$ is a fundamental matrix solution of*

(i) $$x' + Bx = 0.$$

Denote the j^{th}-columns of M and $(M^\tau)^{-1}$ by y_j and z_j respectively, and suppose $b \in \mathrm{col}_n(K)$ and $\int \langle b, z_j \rangle \in K$ for $j = 1, \ldots, n$. Then

(ii) $$y := \sum_j (\int \langle b, z_j \rangle) \cdot y_j$$

is a solution of the inhomogeneous system

(iii) $$x' + Bx = b\,.$$

Proof : By Proposition 5.3 the z_j form a basis of $\mathrm{col}_n(K)$, and from $(M^\tau)^{-1} = (M^{-1})^\tau$ and $M^{-1}M = I$ we see that

$$\langle y_i, z_j \rangle = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise.} \end{cases}$$

The result is now evident from Corollary 6.8. **q.e.d.**

Finally, consider the case of an n^{th}-order linear equation

(6.10) $$u^{(n)} + \ell_1 u^{(n-1)} + \cdots + \ell_{n-1} u' + \ell_n u = 0\,.$$

In this instance the *adjoint equation* generally refers to the n^{th}-order linear equation

(6.11) $$(-1)^n v^{(n)} + (-1)^{n-1}(\ell_1 v)^{(n-1)} + \cdots + (-1)(\ell_{n-1} v)' + \ell_n v = 0\,,$$

e.g., the adjoint equation of

(6.12) $$u'' + \ell_1 u' + \ell_2 u = 0$$

is

(6.13) $$v'' - \ell_1 v' + (\ell_2 - \ell_1')v = 0\,.$$

Examples 6.14 :

(a) The adjoint equation of Bessel's equation

$$y'' + \frac{1}{x}\,y' + (1 - \frac{\nu^2}{x^2})\,y = 0$$

is

$$z'' - \frac{1}{x}\,z' + (1 - \frac{\nu^2 - 1}{x^2})\,z = 0\,.$$

(b) The adjoint equation of any second-order equation of the form

$$y'' + \ell_2 y = 0$$

is the identical equation (despite the fact that they describe differential structures on spaces dual to one another).

To understand why the "adjoint" terminology is used with (6.11) first convert (6.10) to the first order form (3.3) and write the corresponding adjoint equation accordingly, i.e., as

$$\text{(6.15)} \qquad x' - B^\tau x = 0, \qquad -B^\tau = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -\ell_n \\ 1 & 0 & 0 & & 0 & -\ell_{n-1} \\ 0 & 1 & 0 & & \vdots & \vdots \\ & & \ddots & \ddots & & \\ \vdots & & & 1 & 0 & -\ell_2 \\ 0 & & & 0 & 1 & -\ell_1 \end{pmatrix}.$$

The use of the terminology is then explained by the following result.

Proposition 6.16 : *A column vector $x = (x_1, \ldots, x_n)^\tau \in \mathrm{col}_n(K)$ is a solution of (6.15) if any only if x_n is a solution of (6.11) and*

$$\text{(i)} \qquad x_{n-j} = (-1)^j x_n^{(j)} + \sum_{i=0}^{j-1} (-1)^i (\ell_{j-i} x_n)^{(i)} \qquad \textit{for} \qquad j = 1, \ldots, n-1\,.$$

Proof :

$\Rightarrow$ If $x = (x_1, \ldots, x_n)^\tau$ satisfies (6.15) then

(ii) $$x_j' = -x_{j-1} + \ell_{n+1-j} x_n \qquad \text{for} \qquad j = 1, \ldots, n\,,$$

where $x_0 := 0$. It follows that

$$\begin{aligned}
x_n' &= -x_{n-1} + \ell_1 x_n, \\
\\
x_n'' &= -x_{n-1}' + (\ell_1 x_n)' \\
&= -(-x_{n-2} + \ell_2 x_n) + (\ell_1 x_n)' \\
&= (-1)^2 x_{n-2} + (-1)\ell_2 x_n + (\ell_1 x_n)', \\
\\
x_n^{(3)} &= (-1)^2 x_{n-2}' + (-1)(\ell_2 x_n)' + (\ell_1 x_n)'' \\
&= (-1)^2(-x_{n-3} + \ell_3 x_n) + (-1)(\ell_2 x_n)' + (\ell_1 x_n)'' \\
&= (-1)^3 x_{n-3} + (-1)^2 \ell_3 x_n + (-1)(\ell_2 x_n)' + (\ell_1 x_n)'',
\end{aligned}$$

and by induction (on j) that

$$x_n^{(j)} = (-1)^j x_{n-j} + \sum_{i=0}^{j-1} (-1)^{j-1-i} (\ell_{j-i} x_n)^{(i)}\,, \qquad j = 1, \ldots, n\,.$$

This is equivalent to (i), and equation (6.11) amounts to the case $j = n$.

$\Leftarrow$ Conversely, suppose x_n is a solution of (6.11) and that (i) holds. We must show that (iii) holds or, equivalently, that

$$x_{n-j}' = -x_{n-(j+1)} + \ell_{j+1} x_n \qquad \text{for} \qquad j = 1, \ldots, n.$$

This, however, is immediate from (i). Indeed, we have

$$\begin{aligned}
x_{n-j}' &= (-1)^j x_n^{(j+1)} + \textstyle\sum_{i=0}^{j-1} (-1)^i (\ell_{j-i} x_n)^{i+1} \\
&= (-1)^j x_n^{(j+1)} + \textstyle\sum_{i=1}^{j} (-1)^{i+1} (\ell_{j+1-i} x_n)^{(i)} \\
&= (-1)^j x_n^{(j+1)} + \textstyle\sum_{i=0}^{j} (-1)^{i+1} (\ell_{j+1-i} x_n)^{(i)} + \ell_{j+1} x_n \\
&= -\Big((-1)^{j+1} x_n^{(j+1)} + \textstyle\sum_{i=0}^{j} (\ell_{j+1-i} x_n)^{(i)}\Big) + \ell_{j+1} x_n \\
&= -x_{n-(j+1)} + \ell_{j+1} x_n\,.
\end{aligned}$$

q.e.d.

For completeness we record the n^{th}-order formulation of Corollary 6.9.

Proposition 6.17 ("Variation of Constants") : *Suppose* $\ell_1, \ldots, \ell_n \in K$ *and* $\{y_1, \ldots, y_n\} \subset \mathrm{col}_n(K)$ *is a collection of solutions of the* n^{th}*-order equation*

$$\text{(i)} \qquad u^{(n)} + \ell_1 u^{(n-1)} + \cdots + \ell_{n-1} u' + \ell_n u = 0$$

linearly independent over K_C. *Let*

$$M := \begin{pmatrix} y_1 & y_2 & \cdots & & y_n \\ y_1' & y_2' & & \vdots & y_n' \\ y_1^{(2)} & y_2^{(2)} & & & \vdots \\ \vdots & & & & \\ y_1^{(n-1)} & y_2^{(n-1)} & \cdots & \cdots & y_n^{(n-1)} \end{pmatrix}$$

and let $(z_1, \ldots, z_n)$ *denote the* n^{th}*-row of the matrix* $(M^\tau)^{-1}$. *Suppose* $k \in K$ *is such that* $\int k z_j \in K$ *for* $j = 1, \ldots, n$. *Then*

$$\text{(ii)} \qquad y := \sum_j (\int k z_j) \cdot y_j$$

is a solution of the inhomogeneous equation

$$\text{(iii)} \qquad u^{(n)} + \ell_1 u^{(n-1)} + \cdots + \ell_{n-1} u' + \ell_n u = k\,.$$

Proof : Convert (i) to a first-order system as in (3.3) and note from Propositions 1.6 and 3.4 that M is a fundamental matrix solution. Denote the j^{th}-column of M by $\hat{y}_j$ and apply Corollary 6.9 with $b = (0, \ldots, 0, k)^\tau$ and y_j (in that statement) replaced by $\hat{y}_j$ so as to achieve

$$\hat{y}' + B\hat{y} = b.$$

Now write $\hat{y} = (y, \hat{y}_2, \ldots, \hat{y}_n)^\tau$ and eliminate $\hat{y}_2, \ldots, \hat{y}_n$ in the final row of (iii) by expressing these entities in terms of y and derivatives thereof: the result is (iii).

q.e.d.

Since our formulation of Proposition 6.17 is not quite standard[5], a simple example seems warranted.

[5] Cf. [C-L, Chapter 3, §6, Theorem 6.4, p. 87].

Example 6.18 : For $K = \mathbb{C}(x)$ with derivation $\frac{d}{dx}$ we consider the inhomogeneous second-order equation

$$\text{(i)} \qquad u'' + \frac{2}{x}\,u' - \frac{6}{x^2}\,u = x^3 + 4x\,,$$

and for the associated homogeneous equation

$$u'' + \frac{2}{x}\,u' - \frac{6}{x^2}\,u = 0$$

take $y_1 = x^2$ and $y_2 = 1/x^3$ so as to satisfy the hypothesis of Proposition 6.17. In the notation of that proposition we have

$$M = \begin{pmatrix} x^2 & \frac{1}{x^3} \\ 2x & -\frac{3}{x^4} \end{pmatrix}, \qquad (M^\tau)^{-1} = \begin{pmatrix} \frac{3}{5x^2} & \frac{2x^3}{5} \\ \frac{1}{5x} & -\frac{x^4}{5} \end{pmatrix},$$

and from the second matrix we see that $z_1 = 1/5x$, $z_2 = -x^4/5$. A solution to (i) is therefore given by

$$\begin{aligned} y &= \left(\int \tfrac{1}{5x} \cdot (x^3 + 4x)\right) \cdot x^2 + \left((-1) \int \tfrac{x^4}{5} \cdot (x^3 + 4x)\right) \cdot \frac{1}{x^3} \\ &= \frac{x^5}{24} + \frac{2x^3}{3}\,, \end{aligned}$$

as is easily checked directly.

7. Cyclic Vectors

Throughout the section K *is a differential field and* (V, D) *is a differential* K*-module of dimension* $n \geq 1$.

Define $D^0 := \mathrm{id}_V$ (the identity operator on V), $D^1 := D$, and $D^k := D \circ D^{k-1}$ for $k > 1$. Using (1.1) and induction on $n \geq 1$ we see that for any $k \in K$ and $v \in V$ we have the *Leibniz rule*

$$D^n(kv) = \sum_{j=0}^{n} \binom{n}{j} k^{(j)} D^{n-j} v \,. \tag{7.1}$$

A vector $v \in V$ is *cyclic* (w.r.t. D), or is D*-cyclic*, if $\{v, Dv, \ldots, D^{n-1}v\}$ is a basis of V.

Example 7.2 : *Suppose* K *has characteristic* 0 *and contains an element* k *such that* $k' = 1$. *Moreover, assume* V *admits a basis* $\{e_1, \ldots, e_n\}$ *of horizontal vectors. Then the vector*

$$v := \sum_{j=0}^{n-1} \frac{k^j}{j!} e_{j+1}$$

is cyclic. Indeed, by induction one sees that

$$D^\ell v = \sum_{j=\ell}^{n-1} \frac{k^{j-\ell}}{(j-\ell)!} e_{j+1}$$

for $\ell = 1, \ldots, n-1$, and the claim follows easily.

The hypotheses of Example 7.2 are too restrictive for most applications, but we will see that cyclic vectors always exist when K has characteristic 0, the inclusion $K_C \subset K$ is proper (i.e., the derivation on K is nontrivial), and the field of constants K_C is algebraically closed. The goal of the section is a proof of this assertion,[6] but it seems preferable to first indicate why such vectors might be of interest.

[6]The proof we offer is due to J. Kovacic (unpublished). However, any errors that appear are the responsibility of R. Churchill. For a constructive proof, and references to alternate proofs, see [C-K].

Proposition 7.3 : *The following assertions are equivalent:*

(a) $D^* : V^* \to V^*$ *admits a cyclic vector;*

(b) *there is a basis* $\mathbf{e}$ *of* V *such that the defining* $\mathbf{e}$*-equation of* D *has the form*

$$\text{(i)} \qquad x' + \begin{pmatrix} 0 & -1 & 0 & \cdots & & 0 \\ 0 & 0 & -1 & 0 & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \\ & & & & & \\ 0 & & & & & -1 \\ p_0 & p_1 & \cdots & & p_{n-2} & p_{n-1} \end{pmatrix} x = 0,$$

and

(c) *there is a basis representation of* D *which can be converted to the* n^{th}*-order form*

$$\text{(ii)} \qquad y^{(n)} + p_{n-1}y^{(n-1)} + \cdots + p_1 y' + p_0 y = 0\,.$$

Less formally: *finding a cyclic vector for* D^* *is equivalent to expressing* D *in terms of an* n^{th}*-order homogeneous linear differential equation.*

Proof :

(a) $\Leftrightarrow$ (b) : From the definition of a cyclic vector one sees that D^* admits such a vector v^* if and only if there is a basis $(v^*, e_2^*, \ldots, e_n^*)$ of V^* such that the associated defining matrix has the form

$$\begin{pmatrix} 0 & 0 & \cdots & & 0 & -p_0 \\ 1 & 0 & 0 & & 0 & -p_1 \\ 0 & 1 & 0 & & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & & \\ & & & & 0 & -p_{n-2} \\ 0 & 0 & \cdots & 0 & 1 & -p_{n-1} \end{pmatrix}.$$

Since this matrix is the negative transpose of that in (i), the asserted equivalence is evident from Proposition 6.3.

(b) $\Leftrightarrow$ (c) : This equivalence has already been discussed: see the paragraph surrounding (3.3).

q.e.d.

We turn our attention to the existence problem. Although Proposition 7.3 is phrased in terms of cyclic vectors for D^*, it suffices, since $D^{**} \simeq D$, to concentrate on the existence of cyclic vectors for D.

A D-invariant subspace $W \subset V$ is a *differential subspace*, or *D-subspace*, of V. As the reader can easily check, the intersection of any family of such subspaces is again such a subspace. In particular, the intersection $\langle S \rangle$ of those differential subspaces contained some nonempty subset $S \subset V$ is a D-subspace; this is the subspace *differentially generated* by S. When $S = \{s_1, \ldots, s_r\} \subset V$ is finite the notation $\langle \{s_1, \ldots, s_r\} \rangle$ is abbreviated to $\langle s_1, \ldots, s_r \rangle$. In particular, $\langle v \rangle$ denotes the subspace differentially generated by a single vector $v \in V$.

Proposition 7.4 : *When $1 \leq m \leq n$ and $W \subset V$ is an m-dimensional D-space and $v \in W$ the following statements are equivalent:*

(a) $\langle v \rangle = W$ *;*

(b) $\{v, Dv, \ldots, D^{m-1}v\}$ *is a basis of W.*

Proof :

(a) $\Rightarrow$ (b) : If (b) is false there must be scalars $k_j \in K$ such that $\sum_{j=0}^{m-1} k_j D^j v = 0$. Letting p denote the maximal $j \leq m-1$ satisfying $k_j \neq 0$ we can write this dependence relation in the form $D^p v = \sum_{j=0}^{p-1} \hat{k}_j D_j v$, from which we easily see that $D^{p+s}v$ must be in the span of $\{v, Dv, \ldots, D^{p-1}v\}$ for all $s \geq 0$. We conclude that $\dim_K(\langle v \rangle) \leq p < m = \dim_K(W)$, hence that $\langle v \rangle \neq W$, and we have a contradiction.

(b) $\Rightarrow$ (a) : Obvious.

q.e.d.

Corollary 7.5 : *A vector $v \in V$ is cyclic if and only if $V = \langle v \rangle$.*

Our proof of the existence of cyclic vectors requires four lemmas.

Lemma 7.6 : *Suppose $w, v \in V$ are non zero vectors satisfying both $V = \langle w, v \rangle$ and $\langle v \rangle \neq V$. Then :*

(a) $V = \langle w \rangle + \langle v \rangle$;

(b) $\langle w \rangle \cap \langle v \rangle = \{0\}$ *if and only if* $\dim_K(V) = \dim_K(\langle v \rangle) + \dim_K(\langle w \rangle)$; *and*

(c) *the inclusion* $\langle w\rangle \cap \langle v\rangle \subset \langle w\rangle$ *is proper.*

Proof :
(a) : $\langle w\rangle + \langle v\rangle$ is a differential subspace of V containing the set $\{w, v\}$, and $\langle w, v\rangle = V$ is the intersection of all such subspaces. The equality follows.

(b) and (c) : By (a) (and elementary linear algebra) we have $\dim_K(V) = \dim_K(\langle w\rangle) + \dim_K(\langle v\rangle) - \dim_K(\langle w\rangle \cap \langle v\rangle)$. Assertion (b) follows immediately, and if (c) fails the equality reduces to $\dim_K(V) = \dim_K(\langle v\rangle)$, contradicting $V \neq \langle v\rangle$.

q.e.d.

Lemma 7.7 : *Suppose* K_C *is algebraically closed, of characteristic* 0*, and the inclusion* $K_C \subset K$ *is proper. Let* $W \subset V$ *be a nontrivial differential subspace of* V*, let* $1 \leq r \in \mathbb{Z}$*, and let* $w_1, \ldots, w_r \in W$ *be subject only to the restriction* $w_r \neq 0$*. Then there is an element* $\ell \in K$ *such that* $\sum_{j=0}^{r} \ell^{(j)} w_j \neq 0$*.*

Without the characteristic 0 assumption the proper inclusion hypothesis on $K_C \subset K$ can fail, e.g., when K is a finite field one has $K_C = K$.

Proof : Choose any $k \in K \backslash K_C$. It is not difficult to see that when k is algebraic over K_C one has $k \in K_C$, contrary to hypothesis, and we conclude that k must be transcendental over K_C. In particular, the collection $\{1, k, \ldots, k^r\}$ must be linearly independent over K_C.

Fix a basis $(e_1, e_2, \ldots, e_n)$ of V, write $w_j = \sum_{i=1}^{n} a_{ij} e_i$, and choose $\ell \in K$ at random. Then from

$$\sum_{j=0}^{r} \ell^{(j)} w_j = \sum_j \ell^{(j)} \sum_i a_{ij} e_i = \sum_i (\sum_j a_{ij} \ell^{(j)}) e_i$$

we see that

$$\sum_{j=0}^{r} \ell^{(j)} w_j = 0 \Leftrightarrow \sum_j a_{ij} \ell^{(j)} = 0 \quad \text{for} \quad i = 1, \ldots, n\,.$$

In other words, $\sum_{j=0}^{r} \ell^{(j)} w_j = 0$ if and only if ℓ is a solution of the system

$$\begin{aligned} a_{1r} y^{(r)} + a_{1,r-1} y^{(r-1)} + \cdots + a_{10} y &= 0 \\ a_{2r} y^{(r)} + a_{2,r-1} y^{(r-1)} + \cdots + a_{20} y &= 0 \\ \vdots \qquad\qquad\qquad & \quad \vdots \\ a_{nr} y^{(r)} + a_{n,r-1} y^{(r-1)} + \cdots + a_{n0} y &= 0\,. \end{aligned}$$

But from Corollary 4.3 we know that the solution space of any one (and therefore all) of these r^{th}-order homogeneous linear differential equations has K_C-dimension at most r. From the previous paragraph we conclude that for any $k \in K \backslash K_C$ at least one of $1, k, \ldots, k^r$ cannot be a solution, and the lemma is thereby established.

q.e.d.

Lemma 7.8 : *Assume K_C is algebraically closed of characteristic 0 and the inclusion $K_C \subset K$ is proper. Suppose $w, v \in V$ are non-zero and satisfy $\langle w \rangle \cap \langle v \rangle = \{0\}$. Then there is an element $\ell \in K$ with the property that for $\hat{v} := v + \ell w$ one has*

(a) $\langle w, \hat{v} \rangle = \langle w, v \rangle$ *and*

(b) $\dim_K(\langle \hat{v} \rangle) > \dim_K(\langle v \rangle)$.

Proof :

(a) The equality holds for any $\ell \in K$.

(b) Let $r := \dim_K(\langle v \rangle)$. Then $\{v, Dv, \ldots, D^{r-1}v\}$ is a basis of this space (by Proposition 7.4(b)) and the collection $\{v, Dv, \ldots, D^{r-1}v, D^r v\}$ is therefore linearly dependent over K. In particular we can find scalars $a_0, \ldots, a_r \in K$, where w.l.o.g. $a_r = 1$, such that $\sum_{j=0}^r a_j D^j v = 0$.

Now set

$$w_j := \sum_{i=j}^r \binom{i}{j} a_i D^{i-j} w, \quad j = 0, 1, \ldots, r\,,$$

and note that $w_r = w \neq 0$. It follows from Lemma 7.7 that we can find an element $\ell \in K$ such that

(i) $$0 \neq \sum_{j=0}^r \ell^{(j)} w_j = \sum_{j=0}^r \ell^{(j)} \sum_{i=j}^r a_i D^{i-j} w.$$

Define $\hat{v} := v + \ell w$.

Suppose (b) is false. Then there is an integer $1 \leq s \leq r$, and elements $b_i \in K$ for $i = 0, \ldots, s$, where w.l.o.g. $b_s = 1$, such that $\sum_{i=0}^s b_i D^i \hat{v} = 0$, i.e., such that

$$0 = \sum_{i=0}^s b_i D^i \hat{v} = \sum_{i=0}^s b_i D^i v + \sum_{i=0}^s b_i \sum_{j=0}^i \binom{i}{j} \ell^{(j)} D^{i-j} w\,,$$

where in computing the final term we have used (7.1). The first term on the right is in $\langle v \rangle$ while the second is in $\langle w \rangle$, and since $\langle w \rangle \cap \langle v \rangle = \{0\}$ it follows that both must vanish. Immediate consequences of this vanishing are: $s = r$ (because $\sum_{i=0}^s b_i D^i v = 0$ and $b_s = 1$); $a_i = b_i$ for $i = 0, 1, \ldots, r$; and

$$0 = \sum_{i=0}^r a_i \sum_{j=0}^i \binom{i}{j} \ell^{(j)} D^{i-j} w = \sum_{j=0}^r \ell^{(j)} \sum_{i=j}^r \binom{i}{j} a_i D^{i-j} w = \sum_{j=0}^r \ell^{(j)} w_j\,.$$

This contradicts (i), and (b) is thereby established.

q.e.d.

When $W \subset V$ is a D-subspace the quotient space V/W inherits a differential structure in the expected way, i.e., $[v] := v + W \mapsto [Dv] := Dv + W$. This "quotient differential structure" is useful for induction arguments, such as the proof of the following result.

Lemma 7.9 : *Suppose $n = \dim_K(V) > 1$ and every differential K-space of lower positive dimension admits a cyclic vector. Then for any $0 \neq w \in V$ there is a $v \in V$ such that $\langle w, v\rangle = V$.*

Proof : If $\langle w\rangle = V$ take $v = 0$; if not give $V/\langle w\rangle$ the quotient differential structure and let $\pi : V \to V/\langle w\rangle$ denote the quotient mapping. By assumption $V/\langle w\rangle$ contains a cyclic vector $[\hat{v}]$, and for any $v \in \pi^{-1}([\hat{v}])$ we then have $V = \langle w, v\rangle$. **q.e.d.**

Theorem 7.10 (The Cyclic Vector Theorem) : *Suppose K_C is algebraically closed of characteristic 0 and the inclusion $K_C \subset K$ is proper. Then V contains a cyclic vector.*

Proof : We argue by induction on $n = \dim_K(V)$, omitting the trivial case $n = 1$. We assume $n > 1$, and that the result holds for all non trivial differential K-spaces of dimension strictly less than n.

Choose $0 \neq w_1 \in V$ at random. If $\langle w_1\rangle = V$ we are done; otherwise we invoke Lemma 7.9 (and the induction hypothesis) to guarantee the existence of a vector $v_1 \in V$ such that $\langle w_1, v_1\rangle = V$.

We may assume that

(i) $$\dim_K(\langle v_1\rangle) > \dim_K(V) - \dim_K(\langle w_1\rangle)\,,$$

and (as a consequence) that

(ii) $$\langle w_1\rangle \cap \langle v_1\rangle \neq \{0\}.$$

Indeed, if (i) fails then $\langle w_1\rangle \cap \langle v_1\rangle = \{0\}$ by Lemma 7.6(b), and we can use Lemma 7.8 to replace v_1 with a vector $\hat{v}_1$ such that $\langle w_1, \hat{v}_1\rangle = \langle w_1, v_1\rangle = V$ and $\dim_K(\langle \hat{v}_1\rangle) > \dim_K(\langle v_1\rangle)$. Inequality (i) is then evident from $\dim(\langle v_1\rangle) = \dim_K(V) - \dim_K(\langle w_1\rangle)$.

If $\langle v_1\rangle = V$ we are done. Otherwise we use (ii) to find some $0 \neq w_2 \in \langle w_1\rangle \cap \langle v_1\rangle$, noting from Lemma 7.6(c) that $\dim_K(\langle w_2\rangle) < \dim_K(\langle w_1\rangle)$ (which implies that w_2

cannot be cyclic). Repeating the argument of the previous two paragraphs with w_2 replacing w_1 we then produce a vector $v_2 \in V$ such that $V = \langle w_2, v_2 \rangle$,

$$\dim_K(\langle v_2 \rangle) > \dim_K(V) - \dim_K(\langle w_2 \rangle)\,,$$

and

$$\langle w_2 \rangle \cap \langle v_2 \rangle \neq \{0\}\,.$$

If $\langle v_2 \rangle = V$ we are done; otherwise we repeat the construction once again, etc. The result is a sequence of subspaces $\langle w_j \rangle$ with strictly decreasing dimensions and a sequence of vectors v_j satisfying

$$\dim_K(\langle v_j \rangle) > \dim_K(V) - \dim_K(\langle w_j \rangle)\,.$$

But this inequality is impossible if we reach $\dim_K(\langle w_j \rangle) = 0$ (because $\langle v_j \rangle \subset V$), and we conclude that the iteration terminates after finitely many steps. Since the only requirement for continuing is that v_j not be cyclic, this proves the theorem. **q.e.d.**

8. Extensions of Differential Structures

Here K is a differential field with derivation $k \mapsto k'$, V is a K-space (i.e., a vector space over K) *of dimension n, and $D : V \to V$ is a differential structure. Primes $'$ will also be used to indicate the derivation on any differential extension field $L \supset K$.*

Recall[7] that when $L \supset K$ is an extension field of K (not necessarily differential) the tensor product $L \otimes_K V$ over K admits an L-space structure characterized by $\ell \cdot (m \otimes_K v) = (\ell m) \otimes_K v$. This structure will always be assumed. By means of the K-embedding

$$v \in V \mapsto 1 \otimes v \in L \otimes_K V \tag{8.1}$$

one views V as a K-subspace of $L \otimes_K V$ when the latter is considered as a K-space. In particular, any (ordered) basis $\mathbf{e}$ of V can be regarded as a subset of $L \otimes_K V$.

Proposition 8.2 ("Extension of the Base") : *Assuming the notation of the previous paragraph any basis of the K-space V is also a basis of the the L-space $L \otimes_K V$. In particular,*

$$\dim_K V = \dim_L (L \otimes_K V)\,. \tag{i}$$

Proof : See, e.g., [Lang, Chapter XVI, §4, Proposition 4.1, p. 623]. **q.e.d.**

Proposition 8.3 : *Suppose W, $\hat{V}$ and $\hat{W}$ are finite-dimensional*[8] *K-spaces and $T : V \to W$ and $\hat{T} : \hat{V} \to \hat{W}$ are K-linear mappings. Then there is a K-linear mapping $T \otimes_K \hat{T} : V \otimes_K \hat{V} \to W \otimes_K \hat{W}$ characterized by*

$$(T \otimes_K \hat{T})(v \otimes_K \hat{v}) = Tv \otimes_K \hat{T}\hat{v}, \qquad v \otimes_K \hat{v} \in V \otimes_K \hat{V}. \tag{i}$$

Proof[9] **:** There is a standard characterization of the tensor product $V \otimes_K \hat{V}$ in terms of K-bilinear mappings of $V \times \hat{V}$ into K-spaces Y, e.g., see [Lang, Chapter XVI, §1, p. 602]. The proposition results from considering the K-bilinear mapping $(v, \hat{v}) \in V \times \hat{V} \mapsto Tv \otimes_K \hat{T}\hat{v} \in W \otimes_K \hat{W}$. **q.e.d.**

---

[7] As a general reference for the remarks in this paragraph see, e.g., [Lang, Chapter XVI, §4, pp. 623-4, particularly Example 2]. Except for references to bases, most of what we say does not require V to be finite-dimensional.

[8] The finite-dimensional hypothesis is not needed; it is assumed only because this is a standing hypothesis for V.

[9] We offer only a quick sketch. The result is more important for out purposes than a formal proof, and filling in all the details would lead us too far afield.

Proposition 8.4 : *To any differential field extension* $L \supset K$ *there corresponds a unique differential structure* $D_L : L \otimes_K V \to L \otimes_K V$ *extending* $D : V \to V$, *and this structure is characterized by the property*

$$\text{(i)} \qquad D_L(\ell \otimes_K V) = \ell' \otimes_K v + \ell \otimes_K Dv, \qquad \ell \otimes_K v \in L \otimes_K V.$$

Recall from (8.1) that we are viewing V as a K-subspace of $L \otimes_K V$ by identifying V with its image under the embedding $v \mapsto 1 \otimes_K v$. Assuming (i) we have $D_L(1 \otimes_K v) = 1 \otimes_K Dv \simeq Dv$ for any $v \in V$, and this is the meaning of D_L "extending" D.

In the proof we denote the derivation $\ell \mapsto \ell'$ by $\delta : L \to L$, and we also write the restriction $\delta|_K$ as δ.

One is tempted to prove the proposition by invoking Proposition 8.3 so as to define mappings $\delta \otimes_K \mathrm{id}_V : L \otimes_K V \to L \otimes_K V$ and $\mathrm{id}_L \otimes_K D : L \otimes_K V \to L \otimes_K V$ and to then set $D_L := \delta \otimes_K \mathrm{id}_V + \mathrm{id}_L \otimes_K D$. Unfortunately, Proposition 8.3 does not apply since D is not K-linear.

Proof[10] **:** The way around the problem is to first use the fact that D is K_C-linear; one can then conclude from Proposition 8.3 (with K replaced by K_C) that a K_C-linear mapping $\hat{D} : L \otimes_{K_C} V \to L \otimes_{K_C} V$ is defined by

$$\text{(ii)} \qquad \hat{D} := \delta \otimes_{K_C} \mathrm{id}_V + \mathrm{id}_{K_C} \otimes_{K_C} D\,.$$

The next step is to define $Y \subset L \otimes_{K_C} V$ to be the K_C-subspace generated by all vectors of the form $\ell k \otimes_{K_C} v - \ell \otimes_{K_C} kv$, where $\ell \in L$, $k \in K$ and $v \in V$. Then from the calculation

$$\begin{aligned} \hat{D}(\ell k \otimes_{K_C} v - \ell \otimes_{K_C} kv) &= \delta(\ell k) \otimes_{K_C} v + \ell k \otimes_{K_C} Dv \\ &\qquad - \delta(\ell) \otimes_{K_C} kv - \ell \otimes_{K_C} D(kv) \\ &= \ell k' \otimes_{K_C} v + k\ell' \otimes_{K_C} v + \ell k \otimes_{K_C} Dv \\ &\qquad - \ell' \otimes_{K_C} kv - \ell \otimes_{K_C} (k'v + kDv) \\ &= \ell k' \otimes_{K_C} v - \ell \otimes_{K_C} k'v \\ &\qquad + \ell' k \otimes_{K_C} v - \ell \otimes_{K_C} k'v \\ &\qquad + \ell k \otimes_{K_C} Dv - \ell \otimes_{K_C} kDv \end{aligned}$$

[10]Footnote 9 applies here also.

we see that Y is $\hat{D}$-invariant, and $\hat{D}$ therefore induces a K_C-linear mapping $\tilde{D} : (L \otimes_{K_C} V)/Y \to (L \otimes_{K_C} V)/Y$ which by (ii) satisfies

$$\text{(iii)} \qquad \tilde{D}([\ell \otimes_{K_C} v]) = [\ell' \otimes_{K_C} v] + [\ell \otimes_{K_C} Dv],$$

where the bracket [ ] denotes the equivalence class (i.e., coset) of the accompanying element.

Now observe that when $L \otimes_{K_C} V$ is viewed as an L-space (*resp.* K-space), Y becomes an L-subspace (*resp.* a K-subspace), and it follows from (iii) that $\tilde{D}$ is a differential structure when the L-space (*resp.* K-space) structure is assumed.

In view of the K-space structure on $(L \otimes_{K_C} V)/Y$ the K-bilinear mapping[11] $(\ell, v) \mapsto [\ell \otimes_{K_C} v]$ induces a K-linear mapping $T : L \otimes_K V \to (L \otimes_{K_C} V)/Y$ which one verifies to be K-isomorphism. It then follows from (iii) and (iv) that the mapping $D_L := T^{-1} \circ \tilde{D} \circ T : L \otimes_K V \to L \otimes_K V$ satisfies (i), and it follows that D_L is a differential structure on the L-space $L \otimes_K V$.

As for uniqueness, suppose $\check{D} : L \otimes_K V \to L \otimes_K V$ is any differential structure extending D, i.e., having the property

$$\check{D}(1 \otimes_K v) = 1 \otimes_K Dv, \qquad v \in V .$$

Then for any $\ell \otimes_K v \in L \otimes_K V$ one has

$$\begin{aligned} \check{D}(\ell \otimes_K v) &= \check{D}(\ell \cdot (1 \otimes_K v)) \\ &= \ell' \cdot (1 \otimes_K v) + \ell \cdot \check{D}(1 \otimes_K v) \\ &= \ell' \otimes_K v + \ell \cdot (1 \otimes_K Dv) \\ &= \ell' \otimes_K v + \ell \otimes_K Dv \\ &= D_L(\ell \otimes_K v), \end{aligned}$$

hence $\check{D} = D_L$. **q.e.d.**

When considered over the differential field $\mathbb{C}(z) = (\mathbb{C}(z), \frac{d}{dz})$ the linear differential equation $y'' = y$, has only the trivial solution, but if we "allow solutions" from the differential field extension $\mathbb{C}(z)(\exp(z)) = (\mathbb{C}(z)\exp(z), \frac{d}{dz})$ this is no longer the case. At the conceptual level, "allowing solutions from a differential field extension" simply means considering extensions of the given differential structure. However, at the computational level these extensions play no role, as one sees from the following result.

[11] Which we note is not L-bilinear, since for $v \in V$ the product ℓv is only defined when $\ell \in K$.

Proposition 8.5 : *Suppose* $\mathbf{e}$ *is a basis of* V *and*

(i) $$x' + Bx = 0$$

is the defining $\mathbf{e}$*-equation for* D*. Let* $L \supset K$ *be a differential field extension and consider* $\mathbf{e}$ *as a basis for the* L*-space* $L \otimes_K V$*. Then the defining* $\mathbf{e}$*-equation for the extended differential structure* $D_L : L \otimes_K V \to L \otimes_K V$ *is also* (i).

Proof : Since D_L extends D the $\mathbf{e}$ matrices of these two differential structures are the same. **q.e.d.**

A differential field extension $L \supset K$ has *no new constants* if $L_C = K_C$. (Note that $L_C \supset K_C$ is automatic.)

Proposition 8.6 : *Suppose* $\mathbf{e}$ *and* $\hat{\mathbf{e}}$ *are two bases of* V *and*

(i) $$x' + Bx = 0$$

and

(ii) $$x' + Ax = 0$$

are the defining $\mathbf{e}$ *and* $\hat{\mathbf{e}}$*-equations of* D *respectively. Let* $L \supset K$ *be a no new constant differential field extension in which each equation admits a fundamental matrix solution. Then the field extensions of* K *generated by the entries of these fundamental matrix solutions are the same.*

Proof : In view of the discussion surrounding (2.14) we can assume A and B are related by

$$A = PBP^{-1} - P'P^{-1},$$

where $P \in \mathrm{GL}(n, K)$.

Let $M\, N \in \mathfrak{gl}(n, L)$, be fundamental matrix solutions of (i) and (ii) respectively and set $\hat{M} := PM$. Then from

$$\begin{aligned} \hat{M}' &= (PM)' \\ &= PM' + P'M \\ &= P(-BN) + P'M \\ &= -PBP^{-1}\hat{M} + P'P^{-1}\hat{M} \end{aligned}$$

we see that

$$0 = \hat{M}' + (PBP^{-1} - P'P^{-1})\hat{M} = \hat{M}' + A\hat{M},$$

and we conclude that $\hat{M} \in \mathrm{GL}(n, L)$ is also a fundamental matrix solution of (ii). By Proposition 5.5 we have $PM = \hat{M} = NC$ for some $C \in \mathfrak{gl}(n, L_C)$, and we can therefore write

$$\text{(iii)} \qquad M = P^{-1}NC.$$

The entries of P^{-1} are in K, and by the no new constant hypothesis the same holds for the entries of C. The result follows. **q.e.d.**

9. The Differential Galois Group

Here K is a differential field and (V, D) is a differential K-module. We assume $\dim_K V = n < \infty$.

A *Picard-Vessiot extension* for (V, D) is a differential field extension $L \supset K$ satisfying the following conditions:

(a) the extension has no new constants;

(b) the L-space $L \otimes_K V$ admits a basis consisting of horizontal vectors of D_L; and

(c) when $M \supset K$ is any other differential extension satisfying (a) and (b) there is a differential field embedding $\varphi : L \to M$ over K, i.e., a field embedding over K satisfying

$$\varphi \circ \delta_L = \delta_M \circ \varphi \,. \tag{9.1}$$

The intuitive idea behind (c) is that $L \supset M$ is "minimal" among differential extensions $M \supset K$ satisfying properties (a) and (b): any other such extension is "at least as big" in the sense that it must contain an isomorphic copy of L.

The *differential Galois group* of (V, D) corresponding to an associated Picard-Vessiot extension $L \supset K$ is the group G_L of automorphisms of L over K which commute with the derivation δ_L on L. This group obviously depends on L, but, as we will see, only up to isomorphism.

We need a few preliminaries. Define a *fundamental matrix solution of* (V, D) *in* $\mathrm{GL}(n, L)$ to be any fundamental matrix solution of any defining equation of D_L.

Proposition 9.2 : *Condition* (b) *is equivalent to the existence of a fundamental matrix solution for* $(L \otimes_K V, D_L)$ *in* $\mathrm{GL}(n, L)$.

Proof : By Corollary 5.4.

I thank J. Kovacic for the next observation.

Proposition 9.3 : *When* $L \supset K$ *is a Picard-Vessiot extension for* (V, D) *the following equivalent conditions hold.*

(c$_1$) *If* $L \supset M \supset K$ *is an intermediate differential field, and if the differential field extension* $M \supset K$ *also satisfies* (a) *and* (b), *then there is a differential field embedding* $\phi : L \to M$ *over* K.

(c$_2$) *For any fundamental matrix solution* $Z \in \mathrm{GL}(n, L)$ *for* (V, D) *one has* $L = K(Z)$, *i.e.,* L *is generated by the entries of* Z.

(c$_3$) *If* $L \supset M \supset K$ *is an intermediate differential field, and if the differential field extension* $M \supset K$ *also satisfies* (a) *and* (b), *then* $M = L$.

Proof : Condition (c$_1$) is immediate from (c).

(c$_1$) $\Rightarrow$ (c$_2$) : Set $M := K(Z) \subset L$. We must prove that $L \subset M$.

Since the extension $L \supset K$ has no new constants the same is true of $M \supset K$, and by construction M contains the fundamental solution matrix Z of (V, D). It follows from Proposition 9.2 that conditions (a) and (b) in the definition of a Picard-Vessiot extension are satisfied. By (c) there is a differential embedding $\phi : L \to K(Z)$, and from (9.1) one sees that $\phi(Z)$ is also a fundamental matrix solution for (V, D). From Proposition 5.5 we conclude that $ZC = \phi(Z)$ for some $C \in \mathrm{GL}(n, L_C) = \mathrm{GL}(n, K_C)$, the last equality by (a).

Choose any $\ell \in L$. Then $\phi(\ell) \in M = K(Z)$, and we can therefore write

$$\phi(\ell) = \frac{p(Z)}{q(Z)}, \qquad \text{where} \qquad p, q \in K[X_{ij}] \qquad \text{and} \qquad q(Z) \neq 0.$$

Consider the elements $p(ZC^{-1})$ and $q(ZC^{-1})$ of M. Since ϕ is an embedding over K and $C \in \mathrm{GL}(n, K_C) \subset \mathrm{GL}(n, K)$ we have

$$\phi(p(ZC^{-1})) = p(\phi(ZC^{-1})) = p(\phi(Z)\phi(C^{-1})) = p(\phi(Z)C^{-1}) = p(Z)$$

and, similarly,

$$\phi(q(ZC^{-1})) = q(Z).$$

Note from this last identity and $q(Z) \neq 0$ that $q(ZC^{-1}) \neq 0$. Since ϕ is injective we see from

$$\phi\left(\ell - \frac{p(ZC^{-1})}{q(ZC^{-1})}\right) = \phi(\ell) - \frac{p(Z)}{q(Z)} = 0$$

that $\ell = \frac{p(ZC^{-1})}{q(ZC^{-1})} \in M$, and $L \subset M$ follows.

$(c_2) \Rightarrow (c_3)$: If $W \in \mathrm{GL}(n, M)$ is a fundamental matrix solution for (V, D) then $M = K(W)$ by (c_2). However, W may also be considered as a fundamental matrix solution of (V, D) in $\mathrm{GL}(n, L)$, and, as above, we therefore have $W = ZC$ for some $C \in \mathrm{GL}(n, K_C)$. The equalities $M = K(W) = K(ZC) = K(Z) = L$ follow.

$(c_3) \Rightarrow (c_1)$: Take $\phi := \mathrm{id}_L$.

q.e.d.

Corollary 9.4 : *When $L \supset K$ is a Picard-Vessiot extension for (V, D) any differential field embedding $\phi : L \to L$ over K is an automorphism.*

Proof : $\phi(L) \subset L$ satisfies (a) and (b), hence $\phi(L) = L$ by (c_3). **q.e.d.**

Theorem 9.5 : *Suppose $L \supset K$ and $M \supset K$ are Picard-Vessiot extensions for (V, D). Then there is a differential field isomorphism $\phi : L \to M$ over K, and the associated differential Galois groups are isomorphic.*

The mapping ϕ is not unique, e.g., for any $g \in G_M$ the composition $g \circ \phi : L \to M$ is another differential field isomorphism over K. Nevertheless, the result will be used to justify reference to "the" Picard-Vessiot extension of a differential K-module (V, D). A similar convention is used with differential Galois groups: one refers to "the" differential Galois group of (V, D).

Proof : By definition we have differential embeddings $L \rightleftarrows M$ over K which by Corollary 9.4 must be isomorphisms. If we let $\phi : L \to M$ denote the upper arrow we obtain an isomorphism $\sigma : G_L \to G_M$ between the differential Galois groups by assigning $g \in G_L$ to $\phi \circ g \circ \phi^{-1} \in G_M$; the inverse is given by $h \in G_M \mapsto \phi^{-1} \circ h \circ \phi \in G_L$. **q.e.d.**

Picard-Vessiot extensions and differential Galois groups have traditionally been associated with homogeneous linear ordinary differential equations, i.e., with what we are viewing as basis representations of differential modules. Readers familiar with that approach will see immediately from Proposition 9.3(c_2) that the extensions and groups we have defined for a differential module agree with the traditional definitions for any defining equation of that structure.

One question we have not addressed is existence, i.e., does a Picard-Vessiot extension exist for any differential structure? In general the answer is no: the standard proof one needs characteristic zero and algebraically closed hypotheses on K_C.

Bibliography

[C-K] R.C. Churchill and J. Kovacic, Cyclic Vectors, in *Differential Algebra and Related Topics*, (Li Guo, P. Cassidy, W. Keigher and W. Sit, eds.), World Scientific, Singapore, 2002.

[C-L] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, 1955.

[Lang] S. Lang, *Algebra*, Revised Third Edition, Springer-Verlag, New York, 2002.

[Poole] E.G.C. Poole, *Introduction to the Theory of Linear Differential Equations*, Dover Publications, New York, 1960,

[vdP-S] M. van der Put and M.F. Singer, *Galois Theory of Linear Differential Equations*, Springer-Verlag, Berlin, 2003.

R.C. Churchill
Department of Mathematics
Hunter College and the Graduate Center, CUNY,
and the University of Calgary
e-mail rchurchi@math.hunter.cuny.edu