Formal Reduction Of The Operator $\frac{d}{dx} + F$ To Canonical Form

Adam Graham-Squire
Fall 2009

Introduction

In this paper we analyze the differential operator $\frac{d}{dx} + F : \mathbb{C}((x))^n \to \mathbb{C}((x))^n$, where F is an $n \times n$ matrix with entries in $\mathbb{C}((x))$, so the entries are formal Laurent series. Our goal is to find a basis under which F corresponds to a canonical form. The structure of the paper is based in large part on [2], written in the more modern language of [1]. In particular we divide the main calculation into 4 cases, depending on the size of F and the order of pole of F. We will prove not only the existence and uniqueness of a canonical form, but also an algorithm to follow so as to bring F to canonical form.

Statement of Theorem

We will use the convention that for a matrix $M \in M_n\mathbb{C}((x))$ with a pole of order k, we can write

$$M = M_{-k}x^{-k} + M_{-k+1}x^{-k+1} + \cdots + M_{-1}x^{-1} + M_0 + M_1 x + M_2 x^2 + \cdots$$

where $M_h \in M_n\mathbb{C}$ for all h. With such convention in mind, our goal is to prove the following theorem.

Theorem. *For a given matrix $F \in M_n\mathbb{C}((x))$, the operator $\frac{d}{dx} + F : \mathbb{C}((x))^n \to \mathbb{C}((x))^n$ can be written in a canonical form $\frac{d}{dx} + \dot{F}$. Here the basis for $\dot{F}$ lies in $\mathbb{C}((x^{1/q}))^n$ and $\dot{F} \in M_n\mathbb{C}((x^{1/q}))$, for some positive integer $q \le n!$, and $\dot{F}_h = 0$ for all $h \ge 0$. Moreover, $\dot{F}$ is a block diagonal matrix where each block of size $m \times m$ is of the form $\sum_{i=1}^{u} b_i x^{(-1-\frac{i}{q})} I_m + \frac{R}{x}$. Here I_m is the $m \times m$ identity matrix, $b_i \in \mathbb{C}$, and all $\sum_{i=1}^{u} b_i x^{(-1-\frac{i}{q})}$ (which we often refer to as the "scalar terms") are pairwise distinct for different blocks. Also, $R \in M_m\mathbb{C}$ is in Jordan canonical form, and all eigenvalues $\rho = c + d\sqrt{-1}$ of R satisfy $c \in [0, 1/q)$. The canonical form is unique up to the order of the blocks.*

Method and Conventions

Our method will be to use a change of basis matrix $G \in GL_n\mathbb{C}((x^{1/q}))$ to reduce $\frac{d}{dx} + F$ to a form $\frac{d}{dx} + \dot{F}$. Specifically, we will apply a series of transitional change of basis matrices $G_{(i)} \in GL_n\mathbb{C}((x^{1/q}))$ to a series of transitional operators $\frac{d}{dx} + F_{(i)}$ such that $G_{(i)}^{-1}(\frac{d}{dx} + F_{(i)})G_{(i)} = \frac{d}{dx} + F_{(i+1)}$. Letting $F_{(1)} = F$ and $F_{(r)} = \dot{F}$, we have $G = G_{(1)}G_{(2)} \ldots G_{(r)}$ as our desired change of basis matrix. The method then simplifies to just finding the transitional $G_{(i)}$ matrices. Unless otherwise specified, for the sake of simplicity at each step we consider F to be our input matrix and C our output matrix after G has been applied.

We begin by analyzing what happens to the operator $\frac{d}{dx} + F$ when we conjugate it by G. Note that $G' = \frac{dG}{dx}$, and the product rule holds for matrices, i.e. $\frac{d}{dx}(GH) = G'H + GH'$.

Proposition. *Applying a change of basis G to $\frac{d}{dx}+F$ is equivalent to transforming the matrix F into the matrix $G^{-1}G'+G^{-1}FG$. Such a transformation $F \mapsto G^{-1}G'+G^{-1}FG$ is called a gauge transformation.*

Proof.

$$\begin{aligned}
[G^{-1}(\frac{d}{dx}+F)G](h) &= G^{-1}\frac{d}{dx}[G(h)]+G^{-1}FG(h)\\
&= G^{-1}(G\frac{d}{dx}(h)+G'(h)+G^{-1}FG(h)\\
&= \frac{d}{dx}(h)+G^{-1}G'(h)+G^{-1}FG(h)\\
&= [\frac{d}{dx}+G^{-1}G'+G^{-1}FG](h)
\end{aligned}$$

□

We let

$$G^{-1}G'+G^{-1}FG=C \tag{1}$$

and multiply both sides of (1) by G to get

$$G'+FG=GC \tag{2}$$

Equation (2) is the main equation we will work with.

Before we start to look at cases, let us state some conventions that will simplify our calculations. Frequently we can allow G to have no poles and let $G_0=I_n$. Thus we will often use

$$G=I_n+xG_1+x^2G_2+\cdots \tag{3}$$

and in this case

$$G'=G_1+2xG_2+3x^2G_3+\cdots$$

We will typically solve for G by equating coefficient matrices on both sides of (2). We first equate coefficients of the x^h term for the least value of h, then proceed to higher values of h. Thus we solve first for G_1, then G_2, and so on to verify that an appropriate G can be found. Lastly, for a matrix X with a block decomposition, X_{rs} will be the rs-block of X, $X_{h,rs}$ is the rs-block of X_h, and x_{ij} is the ij-entry of X unless otherwise specified.

We will now break the existence part of the proof of the Theorem into four cases and solve each case in turn. Once we have verified the existence of the canonical form, we will construct an algorithm and also prove uniqueness.

1. Case One: F has no poles

Claim 1.1. *If F has no poles then there exists a G, given by (3), such that $C=\dot{F}=0$. In other words, we can find a G such that (2) becomes*

$$G'+FG=0 \tag{4}$$

Proof. Here $F = F_0 + F_1 x + F_2 x^2 + \dots$, so equating coefficients in (4) gives

- Constant (x^0) term: $G_1 + F_0 I_n = 0$, therefore $G_1 = -F_0$.
- x^{h-1} term: Assuming that we have already determined G_i for all $i < h$, we have
$$G_h = -\frac{1}{h!}\left(\sum_{i=0}^{h-1} F_i G_{h-i-1}\right)$$
.

This gives a recursive method by which we can solve for an appropriate coefficient matrix G_h for all h, so we have found a G which satisfies (3). □

In case one, the Theorem is trivially satisfied since we let $C = \dot{F} = 0$.

2. Case Two: $n = 1$

Claim 2.1. *If $n = 1$ there exists a G, given by (3) and satisfying (2), such that $C_h = F_h$ for all $h < 0$ and $C_h = 0$ for all $h \geq 0$.*

Proof. Since $n = 1$, our matrices are abelian and (2) reduces to

$$G' + (F - C)G = 0 \tag{5}$$

where G is a Taylor series and F and C are Laurent series with a pole of order k. If we let $F_h = C_h$ for all $h < 0$ and $C_h = 0$ for all $h \geq 0$ then $F - C = \sum_{h=0}^{\infty} F_h$. This reduces (5) to case one and thus the Theorem will be satisfied, with the exception that the coefficient R for the x^{-1} term may not have the appropriate real part. This situation will be dealt with later in the paper. □

Remark: We note that if we make any gauge transformation with a constant matrix, $G = G_0$, where G_0 is not necessarily equal to I_n, then $G' = 0$ and (1) reduces to $G^{-1}FG = C$. Therefore we can assume without loss of generality that the leading coefficient matrix for F, call it F_{-k}, is in Jordan canonical form. Moreover, we can assume the Jordan blocks of F_{-k} are arranged so that all blocks with the same eigenvalue are next to one another, and we will refer to these (possibly larger) blocks as *spectral decomposition* blocks. We will assume the convention that the spectral decomposition block R_i will have the form

$$R_i = \begin{bmatrix} \rho_i & 0 & 0 & \dots & 0 \\ * & \rho_i & 0 & \dots & 0 \\ 0 & * & \rho_i & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & * & \rho_i \end{bmatrix}$$

where ρ_i is an eigenvalue of F_{-k} and the $*$ are either zero or one. From this point we will assume that the leading coefficient matrix of F is in spectral decomposition form, unless otherwise stated.

It will often be convenient to let our leading coefficient matrix for C equal the leading coefficient matrix for F. Thus if $C_{-k} = F_{-k}$, the calculation $F_{-k}G_i - G_iF_{-k}$ will frequently occur. Let us then consider the operator $T_{F_{-k}} : M_n\mathbb{C} \to M_n\mathbb{C}$, where $T_{F_{-k}}(X) = F_{-k}X - XF_{-k}$.

Lemma 1. *All eigenvalues of $T_{F_{-k}}$ are of the form $\rho_r - \rho_s$.*

Proof. Assuming F_{-k} has m distinct eigenvalues, we have

$$F_{-k} = \begin{bmatrix} R_1 & 0 & \dots & 0 \\ 0 & R_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & R_m \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1m} \\ X_{21} & X_{22} & \dots & X_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ X_{m1} & X_{m2} & \dots & X_{mm} \end{bmatrix} \tag{6}$$

where X is decomposed into blocks of the same size as F_{-k}. Now we calculate $F_{-k}X - XF_{-k} =$

$$\begin{bmatrix} R_1X_{11} - X_{11}R_1 & R_1X_{12} - X_{12}R_2 & \dots & R_1X_{1m} - X_{1m}R_m \\ R_2X_{21} - X_{21}R_1 & R_2X_{22} - X_{22}R_2 & \dots & R_2X_{2m} - X_{2m}R_m \\ \vdots & \vdots & \ddots & \vdots \\ R_mX_{m1} - X_{m1}R_1 & R_mX_{2m} - X_{2m}R_2 & \dots & R_mX_{mm} - X_{mm}R_m \end{bmatrix} \tag{7}$$

It suffices to prove the lemma for a single block of (7). Let us assume the rs-block of (7), $R_rX_{rs} - X_{rs}R_s$, has dimensions $n_r \times n_s$. Then the ij-entry of the block $R_rX_{rs} - X_{rs}R_s$ will have the form

$$(R_rX_{rs} - X_{rs}R_s)_{ij} = (\rho_r - \rho_s)x_{ij} + *x_{i+1,j} - *x_{i,j+1}$$

where $x_{0j} = x_{i,n_s+1} = 0$. Consider a new basis $\tilde{x}_{ij}$ for the rs-block of X, where

$$\tilde{x}_{ij} = (\rho_r - \rho_s)x_{ij} + *x_{i+1,j} - *x_{i,j+1}$$

for all $1 \le i \le n_r, 1 \le j \le n_s$. With respect to this new basis, the operator $T_{F_{-k}}$ can be viewed as a change of basis matrix from the basis $\{x_{ij}\}$ to the basis $\{\tilde{x}_{ij}\}$. To conclude the proof of the lemma, it suffices to show that such a matrix can be written as upper-triangular with $\rho_r - \rho_s$ for all of its diagonal entries. This is easy to see if we order the basis vectors correctly, which we shall demonstrate with an example. Consider a 3×2 rs-block, so $R_rX_{rs} - X_{rs}R_s$ would look like

$$\begin{bmatrix} (\rho_r - \rho_s)x_{11} - *x_{12} & (\rho_r - \rho_s)x_{12} \\ (\rho_r - \rho_s)x_{21} - *x_{22} + *x_{11} & (\rho_r - \rho_s)x_{22} + *x_{12} \\ (\rho_r - \rho_s)x_{31} - *x_{32} + *x_{21} & (\rho_r - \rho_s)x_{32} + *x_{22} \end{bmatrix}$$

We order our vectors by starting at the top of each column and moving down, beginning with the right-most column and moving left. Thus our first vector is $\tilde{x}_{12}$, the second is $\tilde{x}_{22}$, and the last is

$\tilde{x}_{31}$. With respect to this ordering, the change of basis matrix from x_{ij} to $\tilde{x}_{ij}$ will be

$$\begin{bmatrix} (\rho_r-\rho_s) & * & 0 & -* & 0 & 0 \\ 0 & (\rho_r-\rho_s) & * & 0 & -* & 0 \\ 0 & 0 & (\rho_r-\rho_s) & 0 & 0 & -* \\ 0 & 0 & 0 & (\rho_r-\rho_s) & * & 0 \\ 0 & 0 & 0 & 0 & (\rho_r-\rho_s) & * \\ 0 & 0 & 0 & 0 & 0 & (\rho_r-\rho_s) \end{bmatrix}$$

In general, since each $\tilde{x}_{ij}$ only involves entries above and to the right of it, if we order the $\tilde{x}_{ij}$ and x_{ij} appropriately (namely, starting at the upper-right and going down, then moving one column left and repeating the process), the change of basis matrix from the x_{ij}-basis to the $\tilde{x}_{ij}$-basis will be upper triangular with diagonal entries equal to $\rho_r - \rho_s$. □

Corollary. *$T_{F_{-k}}$ is an isomorphism from the subspace of off-block-diagonal matrices to itself.*

Proof. For an off-diagonal block we have $r \neq s$, thus $\rho_r - \rho_s \neq 0$ so all of our eigenvalues are nonzero. □

Lemma 2. *Assume that for all r, s $(\rho_r - \rho_s) \notin \mathbb{Z} - \{0\}$ and $c \in \mathbb{Z} - \{0\}$. Then the operator $cI_{n\times n} + T_{F_{-k}} : M_n\mathbb{C} \to M_n\mathbb{C}$ defined by $[cI_{n\times n} + T_{F_{-k}}](X) = cX + F_{-k}X - XF_{-k}$ is an isomorphism from $M_n\mathbb{C}$ to itself.*

Proof. Given the conclusions of Lemma 1, we have that the eigenvalues of $cI_{n\times n} + T_{F_{-k}}$ will have the form $\rho_r - \rho_s + c$. By the assumptions of Lemma 2 the $\rho_r - \rho_s + c$ will all be nonzero, even when $r = s$. It follows that $cI_{n\times n} + T_{F_{-k}}$ will be an isomorphism. □

Remark: Although the Corollary and Lemma 2 give us isomorphisms, we will only be using the property that the given maps are *surjective*.

3. Case Three: F has a pole of order one.

Claim 3.1. *Suppose F_{-1} (reduced to spectral decomposition form) has m distinct eigenvalues $\rho_i, 1 \leq i \leq m$, and none of the eigenvalues of F_{-1} differ by a non-zero integer ($(\rho_r - \rho_s) \notin \mathbb{Z} - \{0\}$). Then we can find a G, given by* (3) *and that satisfies (2), where $C_{-1} = F_{-1}$.*

Proof. Here we have

$$F_{-1} = \begin{bmatrix} R_1 & 0 & \dots & 0 \\ 0 & R_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & R_m \end{bmatrix}$$

So our calculations for (2) are as follows:

- x^{-1} term: Since G' has no poles, we have

$$F_{-1} - C_{-1} = 0, \text{ therefore } C_{-1} = F_{-1}$$

- x^j term: We can assume that all G_h, $0 \le h \le j$ have already been chosen to satisfy the previous calculations up through the x^{j-1} term, so our calculation at this level gives

$$(j+1)G_{j+1} + F_{-1}G_{j+1} + F_0G_j + F_1G_{j-1} + \cdots + F_jG_0 = G_{j+1}C_{-1}$$
$$\big((j+1)I_{n\times n} + T_{F_{-1}}\big)\,(G_{j+1}) = M$$

  Where M is equal to $-(F_0G_j + F_1G_{j-1} + \cdots + F_jG_0)$. The conditions of Lemma 2 are satisfied (with $c = j+1$), so $\big((j+1)I_{n\times n} + T_{F_{-1}}\big)$ is surjective and we conclude that we can choose entries for G_{j+1} to satisfy the calculation for the x^j term.

This gives a recursive method to choose values for all G_h so that G will satisfy (2) given the conditions for claim 3.1. □

Claim 3.2. *If F_{-1} has two blocks whose eigenvalues differ by a non-zero integer, we can apply a G such that the larger eigenvalue can be reduced until the eigenvalues are equal.*

Proof. We will illustrate the general method with the following specific case:

Subclaim 3.3. *Given a diagonal F_{-1}, we can apply a G (not given by* (3)*) which will reduce the upper-left eigenvalue by one.*

Proof. Here we have

$$F_{-1} = \begin{bmatrix} \rho_1 & 0 & \dots & 0 \\ 0 & \rho_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \rho_n \end{bmatrix}$$

To this matrix we apply the change of basis

$$G = \begin{bmatrix} x^{-1} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

with

$$G^{-1} = \begin{bmatrix} x & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \text{ and } G' = \begin{bmatrix} -x^{-2} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

For our calculation we need only look at the first two coefficient matrices of F, so with

$$F = \frac{1}{x}F_{-1} + F_0 + \text{higher order terms}$$

we have from (1)

$$G^{-1}G' + G^{-1}\frac{F_{-1}}{x}G + G^{-1}F_0G + G^{-1}(\text{h.o.t.})G = C$$

or

$$\begin{bmatrix} -x^{-1} & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} + \begin{bmatrix} \rho_1 x^{-1} & 0 & \dots & 0 \\ 0 & \rho_2 x^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \rho_n x^{-1} \end{bmatrix} + \begin{bmatrix} f_{11} & f_{12}x & \dots & f_{1n}x \\ f_{21}x^{-1} & f_{22} & \dots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1}x^{-1} & f_{n2} & \dots & f_{nn} \end{bmatrix} + \text{h.o.t.} = C$$

where the f_{ij} are the entries of F_0. Thus we conclude

$$C_{-1} = \begin{bmatrix} (\rho_1 - 1) & 0 & \dots & 0 \\ f_{21} & \rho_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & 0 & \dots & \rho_n \end{bmatrix}$$

We have reduced the eigenvalue ρ_1 by 1, and since C_{-1} is lower triangular, we can make another change of basis to return C_{-1} to spectral decomposition form and the eigenvalues will not change. □

In the general case, we will first arrange our blocks so that, given two blocks whose eigenvalues differ by an integer, the block with the larger eigenvalue (call it R_1) is in the upper left of F_{-1}. We then apply G given by

$$G = \begin{bmatrix} x^{-1}I_{n_1} & 0 \\ 0 & I_{n-n_1} \end{bmatrix}$$

where R_1 is of size $n_1 \times n_1$. This will reduce the eigenvalues of R_1 by one, changing all eigenvalues in the block at the same time. We may have f_{ij} entries below the diagonal that are pushed forward from the F_0 matrix, but we can return to spectral decomposition form (as described in the subclaim) without affecting the eigenvalues. This process of applying G and then returning to spectral decomposition form is repeated until the eigenvalues of the two blocks are equalized, at which point they will form one larger spectral decomposition block. □

The process described in claim 3.2 can be repeated for all blocks in F_{-1} with eigenvalues differing by integers. When all possible eigenvalues have been equalized, F_{-1} will have no eigenvalues differing by non-zero integers, and we can apply claim 3.1.

Remark: It is not difficult to see that if one interchanges the matrices for G and G^{-1}, the process above serves to *increase* a particular eigenvalue by one. Therefore any eigenvalue of our leading coefficient matrix has a real part that is unique only up to adding or subtracting an integer. It follows that we can bring all eigenvalues to a form where the real part of each eigenvalue lies in $[0, 1)$, if we so choose. Moreover, if we extend to allow $G \in GL_n\mathbb{C}((x^1/q))$ then a simple calculation shows that one can use the process above to bring the real part of each eigenvalue to be in $[0, 1/q)$. We will use this fact in the conclusion of the paper to explain part of the unicity of the canonical form.

This completes case 3, for C_{-1} is in Jordan canonical form, and by the remark above we can assure that for all $\rho = c + di$ we have $c \in [0, 1)$. Thus, with $q = 1$, our blocks will have the form R/x specified by the Theorem.

4. Case Four: F has a pole of order 2 or higher.

Goal. *If F has a pole of order k, $k \geq 2$, then we can find a G that will reduce F to a form which can be solved by one of the previous cases.*

The goal will be broken into two claims. First we reduce to looking at a single spectral decoposition block and then show that such a block can be reduced in a manner consistent with the goal.

Claim 4.1. *There exists a gauge transformation which reduces F to a diagonal block formation.*

Proof. We will find a change of basis matrix G such that when we apply G, F will be reduced to block diagonal form, where the size of the blocks are determined by the size of the spectral decomposition blocks of F_{-k}. In previous calculations C_{-1} would always equal F_{-1}, but in the case of higher poles this will not in general work. For the following calculations, we make the following assumptions:

- F_{-k} has m distinct eigenvalues of multiplicity n_r (so $\sum_{r=1}^{m} n_r = n$).
- G is given by (3).
- Diagonal blocks of G_h will equal zero for $h \geq 1$ ($G_{h,rr} = 0$ for $h \geq 1$ and all r).
- All off-diagonal blocks of C_h be zero ($C_{h,rs} = 0$ for $r \neq s$).

Thus only the off-diagonal blocks of G and the diagonal blocks of C are arbitrary and need to be determined. With these conventions, let us calculate (2).

- x^{-k} term: $F_{-k}G_0 = G_0C_{-k}$, so we have $F_{-k} = C_{-k}$
- x^{1-k} term: $F_{-k}G_1 + F_{1-k}G_0 = G_1C_{-k} + G_0C_{1-k}$, thus

$$F_{-k}G_1 - G_1F_{-k} + F_{1-k}G_0 - C_{1-k} = 0 \tag{8}$$

  Let us consider what (8) looks like on a diagonal block and on an off-diagonal block.
  Diagonal block: Each diagonal block of $F_{-k}G_1 - G_1F_{-k}$ will be zero since $G_{h,rr} = 0$, for $h \geq 1$. Then (8) will give $F_{1-k,rr} - C_{1-k,rr} = 0$ and since the diagonal block entries of C_h are arbitrary we can choose appropriate values for $C_{1-k,rr}$ to satisfy the expression.
  Off-diagonal block: Since the off-diagonal entries for C_{1-k} are zero and $G_0 = I_n$, the off-diagonal blocks of (8) will simply be $F_{-k}G_1 - G_1F_{-k} = -F_{1-k}$, or $T_{F_{-k}}(G_1) = -F_{1-k}$. It follows from the Corollary that we can find off-block-diagonal entries for G_1 to satisfy this expression.

- x^{t-k} term, $1 < t < k$: Assume we have already calculated values for G_i, $i < t$ and C_i, $i < t - k$. Then we still have no entries from G', so (2) gives

$$F_{-k}G_t + F_{1-k}G_{t-1} + \cdots + F_{t-k}G_0 = G_tC_{-k} + \cdots + G_0C_{t-k}, \text{ or}$$

$$F_{-k}G_t - G_tF_{-k} + M - C_{t-k} = 0 \tag{9}$$

  Where $M = F_{1-k}G_{t-1} + \cdots + F_{t-k}G_0 - G_{t-1}C_{1-k} - \cdots - G_1C_{t-k-1}$. We can now make the same diagonal and off-diagonal calculations for (9) as for (8), replacing M for F_{1-k}, G_t for G_1 and C_{t-k} for C_{1-k}. This satisfies (2) up through the x^{t-k} term for $1 < t < k$.

- x^{t-k} term, $t \geq k$: This calculation will be identical to the case where $1 < t < k$, except we will now have an entry coming from G'. This $(t-k+1)G_{t-k+1}$ matrix will already be determined by previous calculations, and we can add it into the matrix M and repeat the process above to satisfy equation (2) at this level.

This completes the calculation for claim 4.1, as we can recursively find matrices G_h and C_{h-k} to solve equation (2). We have thus found G such that C is now a diagonal-block matrix. $\square$

Once we have written F as a diagonal-block matrix, we have reduced our situation to the point where we can look at each block individually. Thus the only calculation that remains is to show that we can find a canonical form for F when we are given a single spectral decomposition block for F_{-k}. Our method for doing this will be to reduce a given block with pole of order 2 or higher to a previously proved case. Specifically, we want to either break the given spectral decomposition block into blocks of size 1×1, reduce the order of pole to one or zero, or some combination of the two. By induction, it suffices to show that we can either reduce the pole by an integer or break the block into smaller blocks, since a *finite* number of applications will then reduce us to a situation which can be solved by cases 1-3. It will not always be possible to make such a reduction on F itself, so at times we may need to replace F with an equivalent matrix that can be reduced. The details are explained in the following claim.

Claim 4.2. *Given a spectral decomposition block F with a single eigenvalue ρ, size $n \geq 2$, and a pole of order $k \geq 2$, we can bring F to a form C where one of two things will occur: C can be decomposed into blocks of smaller dimension, or we can replace C with a matrix $\tilde{C}$ such that the same G will bring both C and $\tilde{C}$ to canonical form but $\tilde{C}$ will either have a pole of integral order less than C or $\tilde{C}$ can be decomposed into smaller blocks. We will refer to $\tilde{C}$ as an equivalent-reduction matrix to C.*

Proof. We will first change F to a different form, which we will call rational canonical form. Here we use the fact that the transformation $\frac{d}{dx} + F$ acts on the $\mathbb{C}((x))[\frac{d}{dx} + F]$-module $\mathbb{C}((x))^n$. According to [1, Proposition 2.9 and Lemma 2.11], $\mathbb{C}((x))^n$ will have a cyclic vector e, and we can thus consider the basis $\{e, (\frac{d}{dx} + F)e, (\frac{d}{dx} + F)^2 e, \ldots, (\frac{d}{dx} + F)^{n-1} e\}$. After applying the appropriate G to bring us to such a basis, the matrix F will have the form

$$F = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_0(x) \\ 1 & 0 & \ldots & 0 & -a_1(x) \\ 0 & 1 & \ddots & \vdots & -a_2(x) \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \ldots & 0 & 1 & -a_{n-1}(x) \end{bmatrix} \tag{10}$$

where the 1's are on the first subdiagonal and the $a_i(x)$ are the coefficients for the characteristic polynomial of F:

$$y^n + a_{n-1}(x)y^{n-1} + \cdots + a_1(x)y + a_0(x)$$

Remark: Bringing F to rational canonical form may change the order of the pole of F, however the induction process of reducing the order of pole or the size of the matrix will still apply.

If all of the $a_i(x)$ have either no pole or a pole of order 1, we can apply case 1 or case 3 and reduce F to canonical form. Thus we assume that at least one $a_i(x)$ has a pole of order two or higher.

We will now apply the diagonal transformation

$$G = ||\delta_{ij} x^{(i-1)t}|| = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & x^t & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & x^{(n-1)t} \end{bmatrix}$$

to $\frac{d}{dx} + F$. For such a G, the transformation $F \mapsto G^{-1}G' + G^{-1}FG$ is called the *shearing transformation.* we calculate

$$G^{-1}G' = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & tx^{-1} & 0 & \dots & \vdots \\ 0 & 0 & 2tx^{-1} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & t(n-1)x^{-1} \end{bmatrix} \tag{11}$$

and

$$G^{-1}FG = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0(x)\cdot x^{(n-1)t} \\ x^{-t} & 0 & \dots & 0 & -a_1(x)\cdot x^{(n-2)t} \\ 0 & x^{-t} & \ddots & \vdots & -a_2(x)\cdot x^{(n-3)t} \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & x^{-t} & -a_{n-1}(x) \end{bmatrix} \tag{12}$$

Note that (11) has a pole of only order one. Thus it will not affect any of our calculations on terms of higher order poles, so we leave it out of the calculation to find $\dot{t}$. As usual, $G^{-1}G' + G^{-1}FG = C$.

Our first goal is to find the smallest value of t such that (12) will have a pole of order t. We write $\operatorname{ord}(a_i)$ for the order of pole of the Laurent series $a_i(x)$, with the convention that $\operatorname{ord}(a_i) = -\infty$ if $a_i(x)$ has no poles. The pole of $a_i(x)\cdot x^{(n-1-i)t}$ will be of order $\operatorname{ord}(a_i) - (n-1-i)t$. The condition that (12) have the smallest possible pole of order t is therefore equivalent to the system of n inequalities

$$\operatorname{ord}(a_i) - (n-1-i)t \le t, \ \text{ for } i = 0, 1, \dots, n-1 \tag{13}$$

and we choose $\dot{t}$ to be the desired t-value, so

$$\dot{t} = \max\left\{\frac{\operatorname{ord}(a_i)}{n-i}, 1\right\}_{i=0,1,\dots,n-1} \tag{14}$$

Note that $\dot{t}$ is defined, since by assumption at least one $a_i(x)$ has a pole of order two or higher. If $\dot{t} = 1$ we can apply case 3 to $G^{-1}G' + G{-}1FG$ and we will be finished, so we assume that $\dot{t} > 1$.

Letting $t = \dot{t}$, we have

$$C = \begin{bmatrix} 0 & 0 & \dots & 0 & -a_0(x)\cdot x^{(n-1)\dot{t}} \\ x^{-\dot{t}} & 0 & \dots & 0 & -a_1(x)\cdot x^{(n-2)\dot{t}} \\ 0 & x^{-\dot{t}} & \ddots & \vdots & -a_2(x)\cdot x^{(n-3)\dot{t}} \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & x^{-\dot{t}} & -a_{n-1}(x) \end{bmatrix} + G^{-1}G' \tag{15}$$

Where, by construction, at least one of the entries in the right-hand column has a pole of order $\dot{t} > 1$, and all others have a pole of less than or equal to $\dot{t}$.

Our next goal is to show that we can reduce (15) by either decomposing it into submatrices or replacing it by an equivalent-reduction matrix with a lower order of pole. First, consider the leading coefficient matrix of (15)

$$C_{-\dot{t}} = \begin{bmatrix} 0 & 0 & \dots & 0 & c_0 \\ 1 & 0 & \dots & 0 & c_1 \\ 0 & 1 & \ddots & \vdots & c_2 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ 0 & \dots & 0 & 1 & c_{n-1} \end{bmatrix} \tag{16}$$

where, by construction, at least one of the c_i is non-zero. Note that the $G^{-1}G'$ has no impact on (16) and $C_{-\dot{t}}$ is in rational canonical form over $\mathbb{C}$. Here we have two possibilities: (i) The matrix (16) has more than one distinct eigenvalues, or (ii) The matrix (16) has exactly one eigenvalue. Before analyzing these possibilities, let us first consider the relationship between the number of distinct eigenvalues of $C_{-\dot{t}}$ the value of $\dot{t}$, and the entries of $C_{-\dot{t}}$.

Lemma 3. *If (16) has only one eigenvalue, ρ, then all entries c_i are nonzero, $\rho \neq 0$, and $\dot{t}$ is an integer.*

Proof. Suppose (16) has only one eigenvalue, ρ. Then the characteristic equation for (16) will be

$$(y-\rho)^n = y^n - n\rho y^{n-1} + \cdots + (-1)^n \rho^n \tag{17}$$

where $c_i = \binom{n}{i}\rho^{n-i}$, for $0 \le i \le n-1$. Since each c_i has ρ to some power greater than zero in it, we have $\rho = 0$ if and only if $c_i = 0$ for all i. Also $\rho \neq 0$ if and only if $c_i \neq 0$ for all i, and by assumption, at least one $c_i \neq 0$. Thus $\rho \neq 0$ and $c_i \neq 0$ for all i. In particular, this means that all the inequalities in (13) are in fact strict equalities for $t = \dot{t}$. Choosing $i = n-1$ and $t = \dot{t}$, (14) gives us $\mathrm{ord}(a_{n-1}) = \dot{t}$. Since $\mathrm{ord}(a_{n-1})$ is an integer, this proves the first part of Lemma 3. ☐

In situation (i), we convert (16) back to spectral decomposition form and apply claim 4.1. This will break (16) into submatrices associated with the different eigenvalues as desired. However, this means that $\dot{t}$ may not be an integer. We know by construction that for some i, $\dot{t} = \frac{\mathrm{ord}(a_i)}{n-i} \in \mathbb{Q}$, so fractional exponents may thus be introduced, by necessity. Since all our previous calculations involve integer exponents, some comments are in order regarding how our reduction will be done if fractional exponents appear.

In the case where a fractional exponent, say $x^{p/q}$, has been introduced we first make the change of variable

$$\tau = x^{\frac{1}{q}}$$

Then we have

$$\frac{d\tau}{dx} = \frac{1}{q} x^{(1-q)/q}$$

and

$$\frac{d}{dx} = \frac{d}{d\tau} \cdot \frac{1}{q}\tau^{1-q}$$

In terms of the operator $\frac{d}{dx} + F(x)$ (here we write the (x) to emphasize which variable we are using), the change of variable will have the following effect:

$$\begin{aligned} \frac{d}{dx} + F(x) &= \frac{d}{d\tau} \cdot \frac{1}{q}\tau^{1-q} + F(\tau^q) \\ &= \frac{1}{q}\tau^{1-q}\left[\frac{d}{d\tau} + \hat{F}(\tau)\right] \end{aligned} \tag{18}$$

Here the entries in $\hat{F}(\tau)$ will be Laurent polynomials.

Lemma 4. *Suppose that G transforms $(\frac{d}{dx}+F)$ to $(\frac{d}{dx}+C)$. Then for any scalar matrix α, G also transforms $\alpha(\frac{d}{dx}+F)$ to $\alpha(\frac{d}{dx}+C)$.*

Proof. Since α is scalar, it will commute with G^{-1}, thus

$$G^{-1}\alpha(\frac{d}{dx}+F)G = \alpha G^{-1}(\frac{d}{dx}+F)G = \alpha(\frac{d}{dx}+C)$$

□

We can now apply gauge transformations to (18) which will reduce $\frac{d}{d\tau}+\hat{F}(\tau)$ to canonical form. Once we have reduced $\hat{F}(\tau)$ to canonical form $\dot{F}(\tau)$ we can return to the variable x by means of the calculation

$$\begin{aligned} \frac{1}{q}\tau^{1-q}\left[\frac{d}{d\tau}+\dot{F}(\tau)\right] &= \frac{d}{d\tau}\cdot\frac{1}{q}\tau^{1-q}+\frac{1}{q}\tau^{1-q}\dot{F}(\tau^q) \\ &= \frac{d}{dx}+\frac{1}{q}x^{(1-q)/q}\dot{F}(x^{1/q}) \end{aligned} \tag{19}$$

According to Lemma 4, (18), and (19), the same G which transforms $\frac{d}{d\tau}+\hat{F}(\tau)$ to $\frac{d}{d\tau}+\dot{F}(\tau)$ will also transform $\frac{d}{dx}+F(x)$ to $\frac{d}{dx}+\frac{1}{q}x^{(1-q)/q}\dot{F}(x^{1/q})$.

Lemma 5. *If $\dot{F}(\tau)$ is in canonical form, then $G_0^{-1}(\frac{d}{dx}+\frac{1}{q}x^{(1-q)/q}\dot{F}(x^{1/q}))G_0$ will also be in canonical form, where G_0 is a gauge transformation which will assure that the coefficient matrix for x^{-1} is in Jordan canonical form.*

Proof.

$$\dot{F}(\tau) = \sum_{i=1}^{u}\tilde{b}_i\tau^{(-1-\frac{1}{\tilde{q}})}I_n+\frac{\tilde{R}}{\tau}$$

so

$$\begin{aligned} \frac{1}{q}x^{(1-q)/q}\dot{F}(x^{1/q}) &= \frac{1}{q}x^{(1-q)/q}\left(\sum_{i=1}^{u}\tilde{b}_i(x^{1/q})^{(-1-\frac{1}{\tilde{q}})}I_n+\frac{\tilde{R}}{x^{1/q}}\right) \\ &= \sum_{i=1}^{u}\frac{\tilde{b}_i}{q}x^{(-1-\frac{1}{q\tilde{q}})}I_n+\frac{(1/q)\tilde{R}}{x} \end{aligned}$$

Since $\tilde{R}$ is in Jordan canonical form, $\frac{1}{q}\tilde{R}$ may not be in Jordan canonical form. We fix this by applying a gauge transformation G_0 which will bring $\frac{1}{q}\tilde{R}$ to R, where R is in Jordan canonical form. $\frac{1}{q}\tilde{R}$ is lower triangular, so R and $\frac{1}{q}\tilde{R}$ will have the same eigenvalues. Lastly, let $b_i=\frac{\tilde{b}_i}{q}$. The $b_i x^{(-1-\frac{1}{q\tilde{q}})}I_n$ terms will not be affected by the application of G_0 since they are scalars and thus will commute with G_0. Multiplication by $(1/q)$ will also preserve the pairwise distinctness of $\sum\tilde{b}_i(x^{1/q})$, so $\sum b_i x^{[-1-(1/q\tilde{q})]}+R/x$ will be in canonical form, proving Lemma 5. □

Now we consider situation (ii), when (16) has only one eigenvalue. If $C_{-\dot{t}}$ given by (16) has only one eigenvalue, ρ, we will replace C with an equivalent-reduction matrix $\widetilde{C}$ that will either have a pole of order $\dot{t}-j$, (where j is a positive integer) or will decompose into smaller blocks. To finish the proof of Claim 4.2 we need the following two subclaims:

Subclaim 4.3. *Suppose $\beta \in \mathbb{C}$, $A, \tilde{B} \in M_n\mathbb{C}((x))$, $\tilde{A} = A - \beta I_n$, and for some G we have $G^{-1}(\frac{d}{dx} + \tilde{A})G = \frac{d}{dx} + \tilde{B}$. Then $G^{-1}(\frac{d}{dx} + A)G = \frac{d}{dx} + (\tilde{B} + \beta I_n)$.*

In our situation, if we let $\beta = \rho x^{-\dot{t}}$ then subclaim 4.3 implies that any G that will bring $C - \rho x^{-\dot{t}} I_n$ to canonical form will also bring C to canonical form. Thus subclaim 4.3 allows us to replace C with the equivalent-reduction matrix $\tilde{C} = C - \rho x^{-\dot{t}} I_n$.

Remark: The difference between the canonical forms of C and $\tilde{C}$ will be the scalar matrix $\rho x^{-\dot{t}} I_n$. These ρ form some (if not all) of the b_i that show up in the Theorem. The relation between the ρ and b_i will be explained in more detail in the conclusion, however we point out here why the $\sum b_i x^{[-1-(i/q)]}$ will be pairwise distinct for different blocks. Distinct blocks only occur when distinct eigenvalues show up, thus for two different blocks at least one b_i will be distinct.

Subclaim 4.4. *The matrix $\tilde{C} = C - \rho x^{-\dot{t}} I_n$, after being brought to rational canonical form and then applying the shearing transformation, will have a pole of lower order than C.*

Once we have proved subclaim 4.4, there are two possibilities for the leading coefficient matrix of $\tilde{C}$ after it has been put into rational canonical form and then applying the shearing transformation: (I) It has more than one eigenvalue, in which case we break it into submatrices following the method of situation (i), or (II) It has only one eigenvalue, in which case the lower pole is an integer strictly less than $\dot{t}$ and thus we have reduced the order of pole by an integer. In either case claim 4.2 is proved. $\square$

All that remains is to prove the two subclaims.

Proof of Subclaim 4.3. Since βI_n is a scalar matrix it will commute with all matrices. Thus

$$\begin{aligned} G^{-1}\left(\frac{d}{dx} + A\right)G &= G^{-1}\left(\frac{d}{dx} + (A - \beta I_n) + \beta I_n\right)G \\ &= G^{-1}\left(\frac{d}{dx} + \tilde{A}\right)G + \beta I_n \\ &= \frac{d}{dx} + \tilde{B} + \beta I_n \end{aligned}$$

$\square$

Proof of Subclaim 4.4. To prove this claim, it suffices to prove the following lemma.

Lemma 6. *$\tilde{C} = C - \rho x^{-\dot{t}} I_n$ has a rational canonical form*

$$\begin{bmatrix} 0 & 0 & \dots & 0 & -\tilde{a}_0(x) \\ 1 & 0 & \dots & 0 & -\tilde{a}_1(x) \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & -\tilde{a}_{n-2}(x) \\ 0 & \dots & 0 & 1 & -\tilde{a}_{n-1}(x) \end{bmatrix} \tag{20}$$

where the order of pole of $\tilde{a}_i(x)$ is less than $\dot{t}(n-i)$, for all i.

Lemma 6 gives $\operatorname{ord}(\tilde{a}_i) < \dot{t}(n-i)$, so for all i we have

$$\frac{\operatorname{ord}(\tilde{a}_i)}{n-i} < \dot{t}$$

and we can apply the shearing transformation to (20) and choose a value $\tilde{t} = \max\left\{\frac{\text{ord}(\tilde{a}_i)}{n-i}, 1\right\}$ that will be strictly less than $\dot{t}$. Since $\dot{t}$ is the order of pole of C, Lemma 6 proves subclaim 4.4. $\square$

Proof of Lemma 6. By assumption, C has a leading matrix with a single eigenvalue ρ, and the entries in the right hand column of C come from the coefficients of $(y-\rho)^n = y^n - n\rho y^{n-1} + \cdots \pm \rho^n$. Thus, with $\dot{t} > 1$ we have

$$\tilde{C} = C - \rho x^{-\dot{t}} I_n = \begin{bmatrix} -\rho & 0 & \dots & 0 & (-\rho)^n \\ 1 & -\rho & \ddots & \vdots & \vdots \\ 0 & 1 & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & -\rho & \vdots \\ 0 & \dots & 0 & 1 & (n-1)\rho \end{bmatrix} x^{-\dot{t}} + \text{lower order terms}$$

We will now construct the basis under which $\tilde{C}$ will be in rational canonical form. Let e be the $n \times 1$ vector

$$e = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Then

$$\left(\frac{d}{dx} + \tilde{C}\right) e = \begin{pmatrix} -\rho \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} x^{-\dot{t}} + \text{lower order terms}$$

and in general

$$\left(\frac{d}{dx} + \tilde{C}\right)^i e = \begin{pmatrix} * \\ \vdots \\ * \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} x^{-i\dot{t}} + \text{lower order terms}, \ 0 \le i < n \tag{21}$$

where the 1 is in the i^{th} position (considering the top entry as the zero entry), all entries below the i^{th} position are zero, and the $*$'s represent the coefficients of the expansion $(\rho - 1)^i$.

We also have

$$\left(\frac{d}{dx} + \tilde{C}\right)^n e = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} x^{-n\dot{t}} + \text{lower order terms} \tag{22}$$

The vectors in (21) form a basis, though (since together their leading coefficients would form an upper-triangular matrix), and in this new basis we can write

$$\left(\frac{d}{dx} + \tilde{C}\right)^n e = \tilde{a}_0(x) e + \tilde{a}_1(x)\left(\frac{d}{dx} + \tilde{C}\right) e + \cdots + \tilde{a}_{n-1}(x)\left(\frac{d}{dx} + \tilde{C}\right)^{n-1} e \tag{23}$$

for some $\tilde{a}_i(x) \in \mathbb{C}((x))$.

Remark: In particular, combining (22) and (23) illustrates that for each i, one of two things must occur: Either $\tilde{a}_i(x)\left(\frac{d}{dx}+\tilde{C}\right)^i e$ has order of pole strictly less than $n\dot{t}$ or, if a pole of higher order occurs, that pole is canceled by an equivalent pole from some other $\tilde{a}_w(x)\left(\frac{d}{dx}+\tilde{C}\right)^w e$, $w \neq i$, (or sum thereof).

Proposition. *If* $m = \max\{ord(\tilde{a}_i) - \dot{t}(n-i)\}$, *then* $m < 0$.

Proof. Suppose, for the sake of contradiction, that $m \geq 0$. For at least one i we have $\text{ord}(\tilde{a}_i) - \dot{t}(n-i) = m$. Let j be the largest value of i that satisfies $\text{ord}(\tilde{a}_i) - \dot{t}(n-i) = m$, and consider $\tilde{a}_j\left(\frac{d}{dx}+\tilde{C}\right)^j e$. We have that $\text{ord}(\tilde{a}_j) = m + \dot{t}n - \dot{t}j$ and the j^{th} position of $\left(\frac{d}{dx}+\tilde{C}\right)^j e$ will be

$$x^{-j\dot{t}} + \text{ lower order terms}$$

So the j^{th} entry of $\tilde{a}_j(x)\left(\frac{d}{dx}+\tilde{C}\right)^j e$ will have order of pole

$$j\dot{t} + m + \dot{t}n - j\dot{t} = m + \dot{t}n$$

Therefore $\text{ord}(\tilde{a}_j(\frac{d}{dx}+\tilde{C})^j e) \geq \dot{t}n$, so by the remark above, the pole in the j^{th} entry must be canceled by an equivalent pole from some $\tilde{a}_w\left(\frac{d}{dx}+\tilde{C}\right)^w e$, $w \neq j$. However, for $w < j$ there is a zero in the j^{th} position of $\left(\frac{d}{dx}+\tilde{C}\right)^w e$, therefore no pole coming from any $w < j$ could cancel the pole at j. By assumption, we chose the largest value for j, thus there are no $w > j$ that could cancel the pole at j. We have arrived at a contradiction, so we conclude that $m < 0$. ☐

Since $m < 0$ and m is a maximum, we have $\text{ord}(\tilde{a}_i) < \dot{t}(n-i)$ for all i. This concludes the proof of Lemma 6. ☐

5. Structure of Algorithm

The algorithm below could be used to reduce F to canonical form. We use the indices r and s to help reconstruct the blocks once we have reduced our calculations to applying cases 1, 2, or 3. The algorithm ends when case 1, 2 or 3 is applied.

(1) If F has size $n = 1$, apply case 2.
(2) If F has order of pole $k \leq 1$, apply either case 1 or case 3.
(3) Let $r = 1$ and $s = 1$.
(4) If F has size $n \geq 2$ and pole of order $k \geq 2$, then apply claim 4.1 to reduce F to block diagonal form. This step reduces the algorithm to individual spectral decomposition blocks, so the remaining steps only apply to individual blocks.
(5) Repeat steps 1 and 2 on the individual block. We can assume for all further steps that the single block F has size $n \geq 2$ and pole of order $k \geq 2$.
(6) Change F to rational canonical form and repeat step 2.
(7) Apply the shearing transformation to F and call the new matrix C. Repeat step 2 for C.
(8) If the leading coefficient matrix for C, $C_{-\dot{t}}$, has more than one eigenvalue, introduce the change of variable $\tau_r = (\tau_{r-1})^{1/q_r}$. Here we use the conventions that $\tau_0 = x$ and $q_0 = 1$. Replace C with $F(\tau_r)$, increase the index r by one and repeat the process beginning with step 4.

(9) If $C_{-\tilde{t}}$ has only one eigenvalue, ρ_s, replace C with $\tilde{C} = C - \rho_s \tau_{r_s}^{-\tilde{t}_s}$. Increase the index of s by one. Repeat the process beginning with step six, using $F = \tilde{C}$.

After following the algorithm up to this point, one can construct most of the structure of the canonical form given by the Theorem. The final step to complete the canonical form is to apply G such that all eigenvalues ρ of the various R have the property that $\mathrm{Re}(\rho) \in [0, 1/q)$. This is explained below. First we describe the construction of an individual block. Suppose when a given block reaches the point where we can apply case 1, 2 or 3 that it has indices $r = w$ and $s = v$. When we apply case 1, 2, or 3 and return to the variable x we will get a canonical form $\tilde{F}(x^{1/q_1 q_2 \cdots q_w})$. To this we must add the scalar terms that were removed for this block at each step (9) of the algorithm. These will have the form $\sum_{s=1}^{v} \rho_s \tau_{r_s}^{-\tilde{t}_s} I_n$, and when written in terms of the variable x they will look like

$$\sum_{s=1}^{v} \rho_s \left(x^{1/q_1 \cdots q_{r_s}}\right)^{-\tilde{t}_s} I_n$$

which simplifies to

$$\sum_{s=1}^{v} \rho_s x^{\left(-1 - \frac{1+\tilde{t}_s}{q_1 \cdots q_{r_s}}\right)} I_n \tag{24}$$

These terms will be combined with those from the canonical form $\tilde{F}(x^{1/q_1 q_2 \cdots q_w})$ to give us the canonical form *for the block*.

The combination of (24) with $\tilde{F}(x^{1/q_1 q_2 \cdots q_w})$ will occur in different ways depending on whether the last step was an application of case 1, 2 or 3.

- If case 1 was applied then $\tilde{F}(x^{1/q_1 q_2 \cdots q_w}) = 0$ and only the terms from (24) will make up the canonical form.
- If case 2 was applied then $R = \tilde{F}_{-1}$ and any other terms remaining from $\tilde{F}(x^{1/q_1 q_2 \cdots q_w})$ will be added in to the scalar terms from (24).
- If case 3 was applied then $\tilde{F}(x^{1/q_1 q_2 \cdots q_w}) = \tilde{F}_{-1} = R$ and the scalar terms from (24) will be the scalar terms given in the Theorem.

The next to last step to get the canonical form given in the Theorem is that one must find the lowest common denominator for *all* blocks to find the appropriate value of $q \leq n!$, and then rewrite all fractional exponents to correspond to q. Once this is done, one can apply the method of claim 3.2 to reduce or increase eigenvalues for the R matrices. We call this the last step in the algorithm:

Last Step of Algorithm: Applying an appropriate $G \in GL_n\mathbb{C}((x^{1/q}))$ allows us to reduce or increase eigenvalues of any R matrix by multiples of $1/q$. Apply such G until all eigenvalues have a real part that lies in $[0, 1/q)$.

After this final step, the canonical form will be completely constructed from the algorithm. All that remains is to confirm the uniqueness of the canonical form, up to permutations of the order of the blocks.

6. Uniqueness of the Canonical Form

Before we begin the proof of the uniqueness, we remark that the uniqueness is far from obvious. Although the algorithm gives specific steps to follow, the construction of a cyclic vector in [1, Lemma 2.11] indicates that there are many possible choices for the construction of rational canonical form. Thus it is believable that one could have two different canonical forms, say $\tilde{F}$ and $\dot{F}$, that both come from the same original F. We will show that this does not occur.

Claim 6.1. *Suppose that for a given F we have $\dot{F}$ and $\tilde{F}$ both in canonical form, such that $G^{-1}(\frac{d}{dx}+\tilde{F})G = \frac{d}{dx} + \dot{F}$. Then $\dot{F}$ and $\tilde{F}$ have identical blocks.*

Proof. By assumption and (2), we have that

$$G' + \tilde{F}G = G\dot{F} \tag{25}$$

Let G have pole of order v and let the maximum pole of $\tilde{F}$ and $\dot{F}$ have order k, and let us assume that $k > 1$ (the $k = 1$ case will be dealt with in the course of our calculations). We assume without loss of generality that we have integer exponents (the case of fractional exponents is dealt with in a remark at the end of the section), and we use the notation that $\tilde{F}$ has blocks $\tilde{B}_g$, where $\tilde{B}_g$ has the form $\tilde{B}_g = \sum \tilde{b}_{r,g} x^{-1-r} I_{\tilde{n}_g} + \tilde{R}_g/x$ (and similarly for $\dot{F}$). Thus $\tilde{F}_{-k}$ and $\dot{F}_{-k}$ have block decompositions

$$\tilde{F}_{-k} = \begin{bmatrix} \tilde{B}_{-k,1} & 0 & \dots & 0 \\ 0 & \tilde{B}_{-k,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \tilde{B}_{-k,\tilde{m}} \end{bmatrix} \text{ and } \dot{F}_{-k} = \begin{bmatrix} \dot{B}_{-k,1} & 0 & \dots & 0 \\ 0 & \dot{B}_{-k,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \dot{B}_{-k,\dot{m}} \end{bmatrix} \tag{26}$$

where at least one of the $\tilde{B}_{-k,i} = \tilde{b}_{k-1,i} I_{\tilde{n}_i}$ or $\dot{B}_{-k,j} = \dot{b}_{k-1,j} I_{\dot{n}_j}$ is nonzero. Note that the block decomposition is the one given by the canonical form (as opposed to spectral decomposition form which we have used before), so we cannot assume that $\tilde{b}_{k-1,i} \neq \tilde{b}_{k-1,p}$ for $i \neq p$ (and the same holds for $\dot{b}$).

Subclaim 6.2. *$\tilde{b}_{k-1,1} = \dot{b}_{k-1,1}$. Also, $G_{-v,i1} = 0$ for all i such that $\tilde{b}_{k-1,i} \neq \tilde{b}_{k-1,1}$ and $G_{-v,1j} = 0$ for all j such that $\dot{b}_{k-1,j} \neq \dot{b}_{k-1,1}$.*

Proof. Equating coefficients in (25), we see that since there are initially no entries from the G' term we have

x^{-k-v} term:

$$\tilde{F}_{-k} G_{-v} = G_{-v} \dot{F}_{-k} \tag{27}$$

Let $G_{-v} = |G_{-v,ij}|$ be an appropriate block decomposition of G_{-v}. Then (27) reduces to equations

$$\tilde{B}_{-k,i} G_{-v,ij} = G_{-v,ij} \dot{B}_{-k,j} \tag{28}$$

on blocks. By assumption, at least one $G_{-v,ij}$ is nonzero. Since the order of the blocks is irrelevant, we assume that $|g_{su}| = G_{-v,11} \neq 0$, which from (28) gives equations $\tilde{b}_{k-1,1} g_{-v,su} = g_{-v,su} \dot{b}_{k-1,1}$ where at least one $g_{-v,su} \neq 0$. Thus $\tilde{b}_{k-1,1} = \dot{b}_{k-1,1}$, which gives the first part of the subclaim. The second part of the subclaim follows from making the same analysis as above by looking at (28) on the entries where either $i = 1$ or $j = 1$ but $i \neq j$. □

Note that we cannot conclude at this point that the sizes of the blocks $\tilde{B}_{-k,1}$ and $\dot{B}_{-k,1}$ are the same. We can say that $G_{-v,11}$ is an $\tilde{n}_1 \times \dot{n}_1$ block, however, and later we will show that $\tilde{n}_1 = \dot{n}_1$

Subclaim 6.3. *With the same notation as in subclaim 6.2, $\tilde{b}_{h,1} = \dot{b}_{h,1}$ for all $h < k-1$. Also, $G_{h-v,i1} = 0$ for all i such that $\tilde{b}_{k-1,i} \neq \tilde{b}_{k-1,1}$ and $G_{h-v,1j} = 0$ for all j such that $\dot{b}_{k-1,j} \neq \dot{b}_{k-1,1}$.*

Proof. We assume by induction that for all $p > h$ we have $\tilde{b}_{p,1} = \dot{b}_{p,1}$, $G_{p-v,i1} = 0$ for all i such that $\tilde{b}_{k-1,i} \neq \tilde{b}_{k-1,1}$, and $G_{p-v,1j} = 0$ for all j such that $\dot{b}_{k-1,j} \neq \dot{b}_{k-1,1}$. These h values correspond to x^{h-k-v} terms in which the G' does not contribute. Equating coefficient matrices for the x^{h-k-v} term gives

$$\tilde{F}_{-k}G_{h-v} - G_{h-v}\dot{F}_{-k} + \tilde{F}_{1-k}G_{h-v-1} - G_{h-v-1}\dot{F}_{1-k} + \cdots + \tilde{F}_{h-k}G_{-v} - G_{-v}\dot{F}_{h-k} = 0 \tag{29}$$

In general, on a given block (29) gives

$$\tilde{B}_{-k,i}G_{h-v,ij} - G_{h-v,ij}\dot{B}_{-k,j} + \cdots + \tilde{B}_{h-k,i}G_{-v,ij} - G_{-v,ij}\dot{B}_{h-k,j} = 0 \tag{30}$$

In terms of the upper-left block, (30) gives

$$\tilde{B}_{-k,1}G_{h-v,11} - G_{h-v,11}\dot{B}_{-k,1} + \cdots + \tilde{B}_{h-k,1}G_{-v,11} - G_{-v,11}\dot{B}_{h-k,1} = 0 \tag{31}$$

and the assumption that $\tilde{b}_{p,1} = \dot{b}_{p,1}$ for all $p > h$ reduces (31) to just $\tilde{B}_{h-k,1}G_{-v,11} - G_{-v,11}\dot{B}_{h-k,1}$. We can now follow the process of subclaim 6.2 to get the first part of the result. For the second part of the subclaim, we note that the induction assumption (on blocks where $i = 1$ but $j \neq 1$ and such that $\dot{b}_{k-1,j} \neq \dot{b}_{k-1,1}$) reduces (30) to $\tilde{B}_{-k,1}G_{h-v,1j} - G_{h-v,1j}\dot{B}_{-k,j} = 0$. From this we easily conclude the result that $G_{h-v,1j} = 0$ (and the result that $G_{h-v,i1} = 0$ is found in the same manner). ◻

Subclaim 6.4. $G_{-v,1j} = G_{-v,i1} = 0$ *for all* i, j.

Proof. We give the proof for $G_{-v,1j}$, the other proof is similar. Suppose that for a given j, there exists a p such that $\dot{b}_{p,1} = \dot{b}_{p,j}$ for all $h < p \leq k-1$ and $\dot{b}_{h,1} \neq \dot{b}_{h,j}$. Then for the appropriate block, (30) reduces to $\tilde{B}_{h-k,1}G_{-v,1j} - G_{-v,1j}\dot{B}_{h-k,j} = 0$ from which we conclude that (since $\dot{b}_{h,1} \neq \dot{b}_{h,j}$) we must have $G_{-v,1j} = 0$. By the pairwise distinctness condition for the canonical form, such an h must exist for each j, which gives the result. ◻

Subclaim 6.5. $G_{1j} = G_{i1} = 0$ *for all* i, j, *thus* G *has the form*

$$G = \begin{bmatrix} G_{11} & 0 & \dots & 0 \\ 0 & ? & \dots & ? \\ \vdots & \vdots & \ddots & \vdots \\ 0 & ? & \dots & ? \end{bmatrix}$$

Proof. We give the proof for G_{1j}, the other half of the proof is similar. Specifically, we will show that coefficient matrices $G_{p,1j}$ equal zero for all p. Given the results of subclaim 6.4, (30) reduces to

$$\tilde{B}_{-k,i}G_{h-v,ij} - G_{h-v,ij}\dot{B}_{-k,j} + \cdots + \tilde{B}_{h-k-1,i}G_{1-v,ij} - G_{1-v,ij}\dot{B}_{h-k-1,j} = 0 \tag{32}$$

We can now repeat the calculations of subclaims 6.3 and 6.4 to conclude that $G_{1-v,1j} = 0$. By induction, we conclude that $G_{p,1j} = 0$ for all $p < k-1-v$ given the conditions for subclaim 6.3. For all other values of p we must slightly change our calculations to allow for the terms coming from G'. It is easy to see, however, that the terms coming from G' will be canceled out (for example, since $G_{-v,1j} = 0$ for all j, the expression $-vG_{-v,1j}$ will have no effect on later calculations) and thus we have the same result, that $G_{p,1j} = 0$ for all $p \geq k - 1 - v$. We conclude that $G_{1j} = 0$ for all j as desired. □

By induction and subclaim 6.5, we can infer that G has a block form where only the diagonal blocks are nonzero. Since G is invertible, we conclude further that each diagonal block must be invertible, thus the blocks G_{ii} are square. It follows that the size of blocks for $\tilde{F}$ and $\dot{F}$ are the same.

Subclaim 6.6. *The upper left eigenvalues $\tilde{\rho}_1$ and $\dot{\rho}_1$ of $\tilde{R}_1$ and $\dot{R}_1$ are equal. Also, $v = 0$.*

Proof. For the upper left block of the x^{-v-1} term, we will have equation (31) (with $h = k - 2$) plus the terms coming from when $h = k - 1$. Here we write $\tilde{B}_{-1,11} = \tilde{R}_1$ (and similarly for $\dot{R}$) to coincide with our notation for the canonical form. We will then have

$$\tilde{B}_{-k,1}G_{k-1-v,11} - G_{k-1-v,11}\dot{B}_{-k,1} + \cdots + \tilde{R}_1 G_{-v,11} - G_{-v,11}\dot{R}_1 - vG_{-v,11} = 0 \tag{33}$$

By the first part of subclaim 6.3 this reduces to just

$$(\tilde{R}_1 - v)G_{-v,11} - G_{-v,11}\dot{R}_1 = 0 \tag{34}$$

We recall that the R_1 matrices are in Jordan canonical form, with the real part of all eigenvalues in $[0, 1)$. Without loss of generality, assume that the nonzero term of $G_{-v,11}$ occurs in the upper left block of $G_{-v,11}$, asumming that $G_{-v,11}$ is given a block decomposition to match the decomposition of the Jordan forms of the R_1 matrices. We can now make a calculation along the lines of those in Lemmas 1 and 2 to find that $\tilde{\rho}_1$ and $\dot{\rho}_1$ must satisfy the equation

$$\tilde{\rho}_1 - \dot{\rho}_1 - v = 0$$

Now since $v \in \mathbb{Z}$ and $\tilde{\rho}_1, \dot{\rho}_1 \in [0, 1)$ we conclude that $\tilde{\rho}_1 = \dot{\rho}_1$ and $v = 0$, thus G has no poles. □

Subclaim 6.7. *$\tilde{R}_1$ and $\dot{R}_1$ are equal.*

Proof. $G^{-1}(\frac{d}{dx} + \tilde{F})G = \frac{d}{dx} + \dot{F}$ if and only if $\frac{d}{dx} + \tilde{F} = G(\frac{d}{dx} + \dot{F})G^{-1}$. Thus there is a symmetry in the role of G and G^{-1}, and by subclaim 6.6 we can conclude that G^{-1} has no poles as well. It then follows that the leading coefficient matrices of G and G^{-1} must be invertible. Thus $G_{0,11}$ must be invertible because of the block diagonal decomposition of G_0, and we return to the calculation (34). Now that we know that $G_{0,11}$ is invertible, we can conclude that $\tilde{R}_1$ and $\dot{R}_1$ are conjugate. Since both are in Jordan canonical form, they must be equal. □

Together, subclaims 6.3 and 6.7 imply that the upper left blocks of $\tilde{F}$ and $\dot{F}$ are equal, and we can use induction to conclude equality for all blocks. In the case of a fractional exponent with denominator q, it is easy to see that the calculations on the scalar terms of the canonical form will be no different. For our conclusions about the R matrices and the order of pole of G, we use the fact that $Re(\rho) \in [0, 1/q)$ for all eigenvalues to achieve the same results as above. This concludes the proof of claim 6.1 in its full generality. □

Conclusion

This concludes our verification of the Theorem. We have shown that given an operator $\frac{d}{dx} + F$, we can write F in a canonical form. Moreover, we have constructed an algorithm for how to bring F to canonical form, and shown that the choices made in the calculation of the algorithm are unimportant in terms of the final product, as the canonical form is unique.

References

[1] M. van der Put and M.F. Singer. *Galois Theory Of Linear Differential Equations*. Springer-Verlag, Berlin, 2003.
[2] H.L. Turritin. Convergent Solutions of Ordinary Linear Homogeneous Differential Equations In The Neighborhood Of An Irregular Singular Point, *Acta Math.*, 93:27-66, 1955.