

An Introduction to the Algebraic Theory of Differential Equations

Taylor Dupuy

September 23, 2022

Contents

FrontMatter	v
0.1 Why do these notes exist?	v
0.2 Where can I get a digital copy of these notes?	vii
0.3 How do I cite these notes?	vii
0.4 Notation	ix
 1 Differential Algebra Basics	 1
1.1 The Basic Objects	1
1.1.1 Δ -Rings and ∂ -Rings	1
1.1.2 Morphisms of Δ -Rings and ∂ -Rings	2
 2 Monodromy and Hilbert's 21st Problem	 5
2.1 The Monodromy Representation	6
2.1.1 Wronskians	6
2.1.2 Stalks and Germs of Holomorphic and Meromorphic Functions	9
2.1.3 Reduction to First Order Systems	10

2.1.4	Linear Systems	11
2.1.5	Holomorphic Linear Systems	13
2.1.6	Monodromy of Holomorphic Linear Systems	14
2.2	Fuchsian Condition	18
2.3	Hilbert's 21st Problem	22
2.3.1	Representations of $\pi_1(\mathbb{P}^1 \setminus T)$	23
2.3.2	Gauge Transformations	24
2.4	Classification of Fuchsian Equations	26
2.4.1	Some Reminders About Riemann Surfaces	26
2.4.2	The Residue Theorem For Meromorphic Differential Forms	27
2.4.3	Classification of Fuchsian Differential Equations on $\mathbb{P}^1 \setminus S$	29
2.5	Hypergeometric Equations	32
2.5.1	Weyl Algebras and D -Modules	34
2.5.2	Linear Systems and D -Modules	35
2.5.3	Cyclic Vectors and Katz's Theorem	36
2.5.4	Local Exponents	38
2.5.5	Theta Operator and Indicial Equations	39
2.5.6	Fuchs' Relation	43
2.5.7	Local Solutions of Exponent ρ	44
2.5.8	Exponents Determine Equations	46

<i>CONTENTS</i>	iii
3 Riemann-Hilbert	51
3.1 Vector Bundles and Connections	52
3.1.1 Vector Bundles (for the uninitiated)	54
3.1.2 Babymost case: Covariant Derivatives and Connections Associated to ODEs	58
3.1.3 Connections	59
3.1.4 Christoffel Symbols	60
3.1.5 Covariant Derivatives	60
3.1.6 Connection 1-forms	62
3.1.7 Curvature	65
3.2 Plemelj's Construction	71
3.3 Holomorphic Riemann-Hilbert Correspondence	71
4 Poincaré-Fuchs	73
4.1 Differential Equations from the Foliation-Theoretic Perspective	73
4.1.1 Exterior Differential Systems	73
4.1.2 Comparison to Δ -Schemes	73
5 Differential Galois Theory	75
6 Poincaré Problem	77
7 Dimension Theory	79

A	Analytic Elements	81
A.1	Solutions of Homogeneous Differential Equations	81
A.2	Cauchy-Kovalevskya (Cauchy-Kowalevski)	81

FrontMatter

0.1 Why do these notes exist?

These notes are from a course taught in Fall 2022 at UVM entitled *The Algebraic Theory of Differential Equations*. I decided to write these notes because there are a lot different sources for graduate students starting out in Differential Algebra but none of them cover exactly what I would like to cover.

So what is the Algebraic Theory of Differential Equations? For me the starting place are Ritt's books on Differential Algebra. I love Ritt's books [Rit32] and [Rit50]. The issue is that they are a bit out of date and have a lot of dependencies within chapters that are not clearly marked so you almost have to read the books linearly. They are also missing a lot of the classical theory from the 1800's which motivated the subject. The successor to Ritt's books is Kolchin's books [Kol73] which, while mathematically very useful, invokes notation and terminology that gives me nightmares. Also, the algebraic geometry there largely ignores the development of scheme theory between from 1950 to 1970 by the French school. An alternative to these two is Kaplansky's book [Kap76] which I love but is perhaps too brief. Following the spirit of these Differential Algebra Books are Buium's books [Bui86] (influenced by Matsuda's book [Mat80]), [Bui92], and [Bui94] which are probably the most influential on my perspective. Again, to really understand those books (for example the Poincaré-Fuchs theorem on equations of the form $P(t, y, y') = 0$ for a polynomial P whose solutions have no movable singularities) one needs to go back to some of the older material which is best covered elsewhere. They are about differential field theory, differential algebraic groups, and

applications of differential algebra to diophantine geometry.

Also, in the spirit of these Differential Algebra books are the important books on Picard-Vessiot Theory by Singer and van der Put [vdPS03] and Magid [Mag94]. These books are about the Galois theory for linear differential equations.

To cover this perspective one would like to talk about hypergeometric functions, the Painlevé equations, and monodromy more generally. There is the classic book by Ince [Inc44] and a standard text [IKSY91] which is nice but focuses a lot of computations and de-emphasizes global geometry. There are great discussion of Hilbert’s 21st problem in [BGK⁺87] and a more modern algebro geometric version in [Del70]. Marrying this material with the field theoretic methods of [Bui86] is something that I want to do.

Once one gets into the Painlevé equations more algebraic geometry surfaces. The Japanese following Okamoto [Oka87a, Oka87b, Oka86, Oka87c] (and many many papers which I’m not going to list following this thread) showed that there exist rational surfaces of “spaces of initial conditions” for the Painlevé which capture a lot of geometry.

Also, there are so-called Lax Pairs for these Painlevé equations which leads to a theory of “algebraic complete integrability”. The notion of algebraic complete integrability is discussed in, say, [Bea90][AvMV04]. From here one can see that equations like the KP equation admit Lax Pairs and this theory again makes connections to algebraic geometry (this time abelian varieties) through so-called Jacobian Flows and Krichever modules [Mul94].

On top of all this there is a general differential Galois Theory beyond linear equations developed by Umemura [Ume11] and a general theory of Riemann-Hilbert Problems and D-Modules following Malgrange and Kashiwara [BGK⁺87].

I haven’t even mentioned differential algebraic geometry (it’s associated tussles with dimension theory) and the geometry of foliations. To make things worse, much of this material generalizes beyond differential equations, to difference equations, p -derivations, and other operations.

Understandably, I can’t cover this all. I’m not even going to pretend to try. My goal is to survey material. Because of this, I’m going to need to assume

some mathematics at times — there already exists excellent references for much of the material we need to source. This will at times include basic Differential [Inc44] and Partial Differential Equations [Eva10], Commutative Algebra [AM16], Galois Theory [Cox12], Complex Analysis [Ull08], Algebraic Topology [Hat02], Manifolds [Lee13], and Algebraic Geometry [Vak17]. At the same time, I’m not crazy. I don’t want to be writing to nobody. Things that I feel are part of a good introduction for well-prepared graduate students I will review.

In addition to helping graduate students, I want to help myself. I have a number of things I would like to understand better. What is a τ -function? What is a space of initial conditions? What is a Jacobian flow? What proofs work for differential equations but not for difference equations? What do we *really* mean when we say X equation is a limiting case of Y equation? What are the most fundamental examples to keep in mind and teach students when talking about this material? How are classical asymptotic methods “enriched by D -modules and sheaves”?

The subject is vast and I hope we have a fun time exploring it. It may be that I don’t get anywhere on any of this material and we spend 3 months defining what a differential ideal is. We’ll see.

0.2 Where can I get a digital copy of these notes?

A link to the .tex can be found here:

<https://tdupu.github.io/diff-alg-public/diff-alg.pdf>.

0.3 How do I cite these notes?

```
@Unpublished{Dupuy2022,
author = {Dupuy, Taylor},
```

```
title    = {An Introduction to the Algebraic Theory of Differential Equations},
year     = {2022},
comment  = {Course notes from lectures given at the University of Vermont in Fa
journal  = {preprint},
url      = {https://tdupu.github.io/diff-alg-public/diff-alg.pdf},
}
```

0.4 Notation

- R^Δ (or R^∂) the constants of the derivations (or derivation).
- $R\{x\}$ the ring of differential polynomials over R .
- $K(S)_\partial = K(\{S\})$ the field extension of K ∂ -generated by S .
- $\mathbb{C}\langle t - t_0 \rangle$ convergent power series at t_0
- $\mathbb{C}\langle\langle t - t_0 \rangle\rangle$ Laurent series of meromorphic functions (so finite poles).
- $R[\Delta]$ the Weyl Algebra associated to a partial differential ring.

Chapter 1

Differential Algebra Basics

I would skip this for now and only come back to this chapter when we need it.

1.1 The Basic Objects

1.1.1 Δ -Rings and ∂ -Rings

In this book, unless stated otherwise, all rings are going to be commutative with a multiplicative unit. Let R be a commutative ring. By a *derivation* on R we mean a map of sets $\partial : R \rightarrow R$ that satisfied

$$\partial(a + b) = \partial(a) + \partial(b), \quad \forall a, b \in R,$$

$$\partial(ab) = \partial(a)b + a\partial(b), \quad \forall a, b \in R,$$

$$\partial(1) = \partial(0) = 0.$$

Derivations are completely formal here. We don't care about limits.

Exercise 1.1.1.1. Check that all the usual rules hold. For example if $\partial : R \rightarrow R$ is a differential ring then

1. For $n \in \mathbb{Z}_{\geq 0}$ and $a \in R$ we have $\partial(a^n) = na^{n-1}\partial(a)$.
2. For $a \in R$ and $b \in R^\times$ we have $\partial(a/b) = (\partial(a)b - a\partial(b))/b^2$. Here R^\times denotes the elements which have a multiplicative inverse.
3. For $f \in R[x]$ and $a \in R$ we have $\partial(f(a)) = f^\partial(a) + f'(a)\partial(a)$. If $f(x) = \sum_{i=0}^d b_i x^i$ then $f^\partial(x) = \sum_{i=0}^d \partial(b_i)x^i$.

Note that the one exception for derivative rules holding is the chain rule. For an abstract ring R there is not a defined composition of elements $a \circ b$ (although you can compose with polynomials as above).

Definition 1.1.1.2. A *differential ring* or (Δ -ring) is a tuple (R, Δ) where R is a commutative rings with unity and $\Delta = \{\partial_1, \dots, \partial_m\}$ is a collection of commuting derivations $\partial_i : R \rightarrow R$.

When $\Delta = \{\partial\}$ then we call (R, Δ) a ∂ -ring and will use the notation (R, ∂) . We also call such a ring an ordinary differential ring.

- Example 1.1.1.3.**
1. The ring of polynomials in on variable $(\mathbb{C}[t], \frac{d}{dt})$
 2. The ring of rational functions $(\mathbb{C}(t), \frac{d}{dt})$, this is an example of a differential field. In general a *differential field* is a differential ring (K, Δ) where the underlying ring K is a field.
 3. The ring of holomorphic functions $\text{Hol}(U)$ for some $U \subset \mathbb{C}^m$ is an example of a Δ -ring, $(\text{Hol}(U), \{\frac{\partial}{\partial t_1}, \dots, \frac{\partial}{\partial t_m}\})$. Here we are using (t_1, \dots, t_m) for the complex variables $t_j = \sigma_j + i\tau_j$ where $\sigma_j, \tau_j \in \mathbb{R}$.
 4. We can do the same thing with meromorphic functions $\text{Mer}(U)$. These will give a differential field.

1.1.2 Morphisms of Δ -Rings and ∂ -Rings

Let (A, Δ) and (B, Δ) be differential rings where we use $\Delta = \{\partial_1, \dots, \partial_m\}$ for the derivatives on both A and B .

Definition 1.1.2.1. A *morphism* of differential rings is a ring homomorphism $f : A \rightarrow B$ such that for each $\partial_i \in \Delta$ we have $f(\partial_i(a)) = \partial_i(f(a))$ for each $a \in A$.

Chapter 2

Monodromy and Hilbert's 21st Problem

In the theory of linear differential equations with one independent variable z , I wish to indicate an important problem one which very likely Riemann himself may have had in mind. This problem is as follows: To show that there always exists a linear differential equation of the Fuchsian class, with given singular points and monodromic group. The problem requires the production of n functions of the variable z , regular¹ throughout the complex z -plane except at the given singular points; at these points the functions may become infinite of only finite order, and when z describes circuits about these points the functions shall undergo the prescribed linear substitutions. The existence of such differential equations has been shown to be probable by counting the constants, but the rigorous proof has been obtained up to this time only in the particular case where the fundamental equations of the given substitutions have roots all of absolute magnitude unity. L. Schlesinger (1895) has given this proof, based upon Poincaré's theory of the Fuchsian zeta-functions. The theory of linear differential equations would evidently have a more finished appearance if the problem here sketched could be disposed of by some perfectly general method.

Hilbert's 21st Problem

In this chapter we are going to move towards Hilbert's 21st problem and some of the classical theory of monodromy of solutions of differential equations.

¹Hilbert means holomorphic.

2.1 The Monodromy Representation

In this section we develop some basic tools we need to construct a monodromy representation associated to a linear differential equation on \mathbb{P}^1 . ♠♠♠ [Taylor: \[Finish\]](#)

2.1.1 Wronskians

Let (R, ∂) be a differential ring. Let $f_1, \dots, f_n \in R$. The *Wronskian* of f_1, \dots, f_n is

$$W(f_1, \dots, f_n) = \det \begin{pmatrix} f_1 & f_2 & \cdots & f_n \\ f_1' & f_2' & \cdots & f_n' \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)} & f_2^{(n-1)} & \cdots & f_n^{(n-1)} \end{pmatrix}.$$

The Wronskian gives us a test for linear dependence over the constants of a differential field.

Theorem 2.1.1.1. *Let (K, ∂) be a differential field. Let $f_1, \dots, f_n \in K$. Let $C = K^\partial$ be the constants. We have that f_1, \dots, f_n are linearly dependent over C if and only if $W(f_1, \dots, f_n) = 0$.*

Proof. Suppose that f_1, \dots, f_n are linearly dependent over C . Then there exists $c_1, \dots, c_n \in C$ not all zero such that

$$c_1 f_1 + \cdots + c_n f_n = 0.$$

Taking derivatives gives

$$\begin{pmatrix} f_1 & f_2 & \cdots & f_n \\ f_1' & f_2' & \cdots & f_n' \\ \vdots & \vdots & \ddots & \vdots \\ f_1^{(n-1)} & f_2^{(n-1)} & \cdots & f_n^{(n-1)} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} = 0. \quad (2.1.1)$$

Since the $c_i \in K$, this means the matrix B such that $\det(B) = W$ is singular and hence $W = \det(B) = 0$.

Conversely, suppose that $W = 0$. Then there exists some $c_1, \dots, c_n \in K$ not all zero such that (2.1.1) holds. To prove our result, we need to show that $c_1, \dots, c_n \in C$. If some proper subset of $\{f_1, \dots, f_n\}$ have a non-trivial dependence relation we can replace our set with that subset and hence we can assume without loss of generality that $\{f_2, f_3, \dots, f_n\}$ are linearly independent over K . We can also suppose that $c_1 \neq 0$. Furthermore we can scale the vector (c_1, \dots, c_n) by $1/c_1$. Hence we can further assume that $c_1 = 1$.

Now for $1 \leq j \leq n-2$ (note the $n-2$ here) we can take a derivative of

$$c_1 f_1^{(j)} + \dots + c_n f_n^{(j)} = 0$$

To get

$$\begin{aligned} 0 &= c_1 f_1^{(j+1)} + \dots + c_n f_n^{(j+1)} + c'_1 f_1^{(j)} + \dots + c'_n f_n^{(j)} \\ &= c'_2 f_2^{(j)} + \dots + c'_n f_n^{(j)}. \end{aligned}$$

But f_2, \dots, f_n are linearly independent. This implies that $c'_2 = \dots = c'_n = 0$ which implies that $c_1, \dots, c_n \in C$ which proves our result. \square

We want to show that the Wronskian satisfies a linear differential equations. To do this we need a couple things.

In what follows one needs to recall the definition of an adjugate matrix and how cofactor expansion works. Recall that if A is an invertible $n \times n$ matrix then the *adjugate* is defined by

$$\text{adj}(A) = \det(A)A^{-1}.$$

This is the best way to remember the formula. The adjugate is just what would be the inverse would be had we not inverted the determinant. Unlike inverse, it turns out that every $n \times n$ matrix and we can obtain its formula from cofactor expansion. We have

$$\text{adj}(A)_{ji} = (-1)^{i+j} \det(\tilde{A}_{ij})$$

where \tilde{A}_{ij} is the matrix obtains by deleting the i th row and j th column. This all comes from the formula for the inverse of a matrix using cofactor expansion (sometimes also called ‘‘Laplace’s Formula’’).

Finally, we need to know what the partial derivative of the determinant is with respect to each of its entries. In what follows we are going to consider $X = (x_{ij})$ as an abstract $n \times n$ matrix with entries being variables. This means that $\det(X)$ will be viewed as a polynomial in $\mathbb{Z}[x_{ij} : 1 \leq i, j \leq n]$.

Lemma 2.1.1.2. *Let $X = (x_{ij})$ be a symbolic matrix.*

$$\frac{\partial \det(X)}{\partial x_{ij}} = \text{adj}(X)_{ji}.$$

Proof. The proof is direct. By cofactor expansion we have $\det(X) = \sum_{j=1}^n x_{ij} \text{adj}(X)_{ji}$ hence

$$\begin{aligned} \frac{\partial \det(X)}{\partial x_{ij}} &= \frac{\partial}{\partial x_{ij}} \left[\sum_{\ell=1}^n x_{i\ell} \text{adj}(X)_{\ell i} \right] \\ &= \sum_{\ell=1}^n \frac{\partial x_{i\ell}}{\partial x_{ij}} \text{adj}(X)_{\ell i} + x_{i\ell} \frac{\partial}{\partial x_{ij}} \text{adj}(X)_{\ell i} \\ &= \sum_{\ell=1}^n \delta_{\ell j} \text{adj}(X)_{\ell i} = \text{adj}(X)_{ji}. \end{aligned}$$

Note that on the second to last equality we used that $\frac{\partial}{\partial x_{ij}} \text{adj}(X)_{\ell i} = 0$ since $\text{adj}(X)_{\ell i}$ has no terms with i in the first entry and ℓ in the second entry (this is the cofactor expansion formula). \square

To apply this we need the formula for the dot product of matrices. Sometimes this is called the “Killing form”.² If you have never done this exercise in your life you should do it.

Exercise 2.1.1.3. Let $A, B \in M_n(R)$ for a commutative ring R . One has

$$\text{Tr}(A^T B) = \sum_{1 \leq i, j \leq n} A_{ij} B_{ij}.$$

We can now prove our result.

²Named after Wilhelm Killing 1847–1923

Theorem 2.1.1.4. *Let $A = (a_{ij}) \in M_n(R)$ with (R, ∂) a differential ring. We have*

$$\partial(\det(A)) = \text{Tr}(\text{adj}(A)\partial(A))$$

where $\partial(A)$ denotes the matrix $\partial(A) = (\partial(a_{ij}))$. Furthermore if $A \in \text{GL}_n(R)$ then

$$\partial(\det(A)) = \text{Tr}(A^{-1}\partial(A))\det(A).$$

Proof. Let $X = (x_{ij})$. We are going to use the chain rule

$$\begin{aligned} \partial(\det(X)) &= \sum_{1 \leq i, j \leq n} \frac{\partial \det(X)}{\partial x_{ij}} \partial(x_{ij}) \\ &= \sum_{1 \leq i, j \leq n} \text{adj}(X)_{ji} \partial(x_{ij}) \\ &= \text{Tr}(\text{adj}(X)\partial(X)). \end{aligned}$$

To get the last formula, if X is invertible we use the previous formula $\text{adj}(X) = \det(X)X^{-1}$. \square

2.1.2 Stalks and Germs of Holomorphic and Meromorphic Functions

Recall that for $U' \subset U$ open subset of \mathbb{C}^m we have injectures $\text{Hol}(U) \rightarrow \text{Hol}(U')$ and $\text{Mer}(U) \rightarrow \text{Mer}(U')$ given by restricting the domain of some $f(z)$ to U' . Both of these ring homomorphisms are injective by the analytic continuation principle (which holds in several variables as well as one variable).³ The *stalk* at some $t_0 \in \mathbb{C}$ is

$$\text{Hol}_{t_0} = \varinjlim_{U \ni t_0} \text{Hol}(U), \quad \text{Mer}_{t_0} = \varinjlim_{U \ni t_0} \text{Mer}(U)$$

where the direct limit is taken over open set U containing t_0 . Any element of a stalk is called a *germ*.

³If you have never shown that analytic continuation works in two variables this is a good exercise.

It is important to know that there is always a ring homomorphism $\text{Hol}(U) \rightarrow \text{Hol}_{t_0}$ and that any element of $\text{Hol}(U)$ is determined by its stalk. Same goes for meromorphic functions.

Remark 2.1.2.1. For the uninitiated, we recall that if I is a partially ordered set then a directed system is a collection $((R_i)_{i \in I}, (f_{i,j})_{i < j})$ consisting of rings R_i and morphisms $f_{i,j} : R_i \rightarrow R_j$ whenever $i < j$.

The direct limit of the directed system then is the ring

$$\varinjlim R_i = (\coprod_{i \in I} R_i) / \sim$$

where $r_i \in R_i$ and $r_j \in R_j$ are declared equivalent when for some $k > i, j$ we have $f_{i,k}(r_i) = f_{j,k}(r_j)$.

In the one variable case we for $a \in \mathbb{C}$ we are going to use the notation

$$\mathbb{C}\langle t - a \rangle := \text{Hol}_a, \quad \mathbb{C}\langle\langle t - a \rangle\rangle = \text{Mer}_a$$

And in the several variable case for $(a_1, \dots, a_m) \in \mathbb{C}^m$ we will use the notation

$$\mathbb{C}\langle t_1 - a_1, \dots, t_m - a_m \rangle = \text{Hol}_{(a_1, \dots, a_m)}, \quad \mathbb{C}\langle\langle t_1 - a_1, \dots, t_m - a_m \rangle\rangle = \text{Mer}_{(a_1, \dots, a_m)}.$$

In other books they use $\mathbb{C}\{t\}$ for convergent power series but we are going to reserve this symbol for the ring of differential polynomials.

2.1.3 Reduction to First Order Systems

Any system of PDEs is equivalent to a first order system of PDEs. The idea is that we can always introduce more variables every times we need to take a new derivative so that all of our expressions only involve single derivatives of variables. Later, for linear differential equations we will see that we can actually go backwards.

We illustrate this in the case of linear first order differential equations in one differential indeterminate. Here we consider the equation

$$y^{(r)} + a_{r-1}y^{(r-1)} + \dots + a_0y = 0.$$

By introducing “velocity variables” $v_j = y^{(j)}$ for $j = 0, 1, \dots, r-1$ we get a new system

$$\begin{cases} v'_0 = v_1, \\ v'_1 = v_2, \\ \vdots \\ v'_{r-1} = -a_{r-1}v_{r-1} - a_{r-2}v_{r-2} - \cdots - a_0v_0. \end{cases}$$

which then can be written in matrix form

$$V' = AV$$

where

$$V = \begin{pmatrix} v_0 \\ v_1 \\ \vdots \\ v_{r-1} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_{r-1} & -a_{r-2} & -a_{r-3} & \cdots & -a_0 \end{pmatrix}.$$

Here A is just the transpose of a companion matrix. We will often convert between higher order equations and first order equations in this way.

2.1.4 Linear Systems

Let $A \in M_n(R)$ where R is a differential ring. The system

$$Y' = AY \tag{2.1.2}$$

is called a *linear system at over R* in the indeterminates $Y = (y_1, \dots, y_n)$ (I am going to allow myself to abusively conflate row and column vectors). The letter n is sometimes called the *rank* of the linear system.

Exercise 2.1.4.1. The solutions of a linear system form an R^∂ -module.

A matrix $\Phi \in \text{GL}_n(R)$ is called a fundamental set of solutions or *fundamental solution* if

$$\Phi' = A\Phi,$$

where the derivatives of Φ in the expression Φ' are taken component-wise. The idea is that the columns of the matrix Φ form a basis of solutions over the constants R^∂ .

Fundamental matrices are unique. Any solution of the linear system takes the form ΦZ for some vector $Z \in (R^\partial)^{\oplus n}$. This means that if $\tilde{\Phi}$ is another fundamental matrix there exists some $M \in \text{GL}_n(R^\partial)$ such that

$$\tilde{\Phi} = \Phi M. \quad (2.1.3)$$

In the theory of monodromy, these will become the monodromy matrices and in the Picard-Vessiot theory of linear differential algebraic extensions of differential fields these matrices are going to become the Galois group elements. This is so important we are going to put it in a theorem environment.

Theorem 2.1.4.2 (Existence of “Monodromy” Matrices). *If Φ and $\tilde{\Phi}$ are two fundamental matrices of a rank n linear system over a differential ring (R, ∂) then there exists some $M \in \text{GL}_n(R^\partial)$ such that $\tilde{\Phi} = \Phi M$.*

To prove a fundamental set of solutions we are going to use existence and uniqueness together with the following lemma.

Lemma 2.1.4.3. *Let K be a ∂ -field. If $Y_1, \dots, Y_n \in K^n$ are linearly independent over K then they are linearly independent over $C = K^\partial$.*

Proof. We prove this by proving they are linearly dependent over K if and only if they are linearly dependent over C . If they are linearly dependent over C then clearly they are linearly dependent over K . Conversely, suppose that they are linearly dependent over K . We will prove this by induction so we can suppose that no proper subset is linearly dependent over K otherwise we could apply the inductive hypothesis. The base case is immediate.

Now we do the inductive step. By clearing denominators we have $Y_1 = \sum_{j=2}^n c_j Y_j$ for some $c_j \in K$. We have

$$\begin{aligned} 0 &= Y_1' - AY_1 \\ &= \sum_{j=2}^n c_j' Y_j + \sum_{j=2}^n c_j Y_j' - \sum_{j=2}^n c_j AY_j \\ &= \sum_{j=2}^n c_j' Y_j \end{aligned}$$

But since Y_2, \dots, Y_n were assumed to be linearly independent we must have $c'_2 = \dots = c'_n = 0$ which proves the c_j 's are constants. \square

2.1.5 Holomorphic Linear Systems

A *holomorphic linear system* at $t_0 \in \mathbb{C}$ is a linear system over $R = \mathbb{C}\langle t - t_0 \rangle$. That is, it is a system of linear differential equations

$$Y' = AY$$

where the matrix A is holomorphic at $t_0 \in \mathbb{C}$.

We now prove the existence and uniqueness theorem for holomorphic linear systems.

Theorem 2.1.5.1 (Existence and Uniqueness). *Let $t_0 \in \mathbb{C}$. Let $A \in M_n(\mathbb{C}\langle t - t_0 \rangle)$. Let $Y_0 \in \mathbb{C}^n$. There exists a unique $Y \in \mathbb{C}\langle t - t_0 \rangle^{\oplus n}$ such that*

$$\begin{cases} Y' = AY, \\ Y(t_0) = Y_0. \end{cases}$$

There are three ways of doing this. I might add some more details later.

1. Use power series expansions, then prove a convergence result.
2. Big Hammer: Use Cauchy-Kowalevski⁴ This theorem is morally the same as above just with more complicated PDEs. One shows that there is a power series solution then proves convergence.
3. Bigger Hammer: Use the existence of differentially closed fields \widehat{K} is the ∂ -closure of the field $K \subset \mathbb{C}\langle t - t_0 \rangle$ given by $K = \mathbb{Q}(a_{ij} : 1 \leq i, j \leq n)_{\partial}$. This is the differential field generated by the coefficients of the matrix A . The Sierdenberg embedding theorem then tells us that $\widehat{K} \subset \mathbb{C}\langle t - t_0 \rangle$, and this gives us a holomorphic solution of $Y' = AY$. By the property of differential closures once we find a solution we can keep adjoining solutions using Blum's axiom. ♠♠♠ Taylor: [\[explain this further\]](#).

⁴This is the same as Cauchy-Kovalevskaya. Some people spell the Russian name differently.

We now prove the existence of a fundamental matrix.

Lemma 2.1.5.2. *Every holomorphic linear system which is holomorphic at $t_0 \in \mathbb{C}$ admits a fundamental matrix $\Phi(t)$ which is holomorphic at t_0 .*

Proof. By existence and uniqueness we can always find a solution $Y_i \in \mathbb{C}\langle t - t_0 \rangle^{\oplus n}$ satisfying

$$Y_i' = AY_i, \quad Y_i(t_0) = e_i$$

where e_i is an elementary column vector (it has zeros everywhere except for the i th position). The solutions Y_1, \dots, Y_n are linearly independent over $K = \mathbb{C}\langle t - t_0 \rangle^{\oplus n}$ because e_1, \dots, e_n are linearly independent over K . Hence by Lemma 2.1.4.3 we get that the solutions are linearly independent over \mathbb{C} . The matrix

$$\Phi = [Y_1 | Y_2 | \dots | Y_n]$$

is our fundamental system. □

2.1.6 Monodromy of Holomorphic Linear Systems

Consider a holomorphic linear system

$$Y' = AY, \quad A = A(t) \in M_n(\text{Hol}(U)),$$

where $U \subset \mathbb{C}$ a connected open set. By the previous section for each $t_0 \in U$ there exists a fundamental matrix Φ which is holomorphic in a neighborhood of t_0 . We are going to want to analytically continue Φ along every path γ starting at t_0 and obtain Φ^γ which will eventually allow us to cook-up a group homomorphism from the fundamental group of paths starting at t_0 to $\text{GL}_n(\mathbb{C})$ which measures how much Φ changed once we take it around the loop.

The group homomorphism

$$\rho : \pi_1(U, t_0) \rightarrow \text{GL}_n(\mathbb{C}), \quad \rho(\gamma) = M_\gamma$$

is called the *monodromy representation*. We will now explain what $M_\gamma \in \text{GL}_n(\mathbb{C})$ is supposing Φ_γ exists: since Φ and Φ_γ are both fundamental matrices at t_0 then as in (2.1.3) there exists some M_γ such that

$$\Phi_\gamma = \Phi M_\gamma.$$

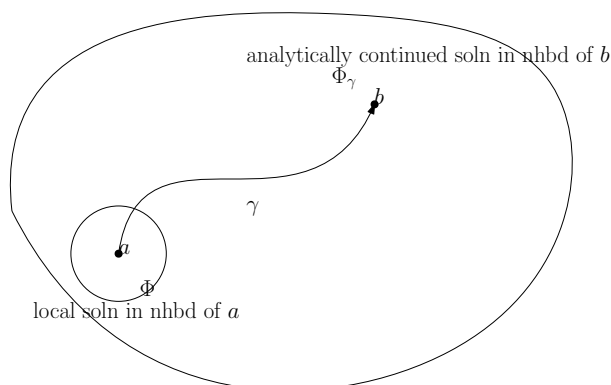


Figure 2.1: A picture of analytically continuing a local fundamental matrix along a path.

That is all.

We need to set our convention for concatenation of paths. If γ_1 and γ_2 are two paths in U where the endpoint of γ_2 is the starting point of γ_1 then we will let $\gamma_2\gamma_1$ denote the path which first performs γ_1 then performs γ_2 .

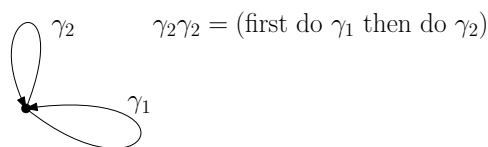


Figure 2.2: The convention we use for composition of paths. Other people use other conventions and it will mess up your formulas.

Remark 2.1.6.1 (WARNING). Conventions on concatenation of paths changes from text to text and this will mess with your formulas.

With our convention for concatenation of paths we have.

$$\Phi_{\gamma_2\gamma_1} = (\Phi_{\gamma_1})_{\gamma_2}$$

On one hand we have $\Phi_{\gamma_2\gamma_1} = \Phi M_{\gamma_2\gamma_1}$. On the other hand we have $(\Phi_{\gamma_1})_{\gamma_2} = (\Phi M_{\gamma_1})_{\gamma_2} = \Phi_{\gamma_2}(M_{\gamma_1})_{\gamma_2} = \Phi M_{\gamma_2} M_{\gamma_1}$. This proves that

$$M_{\gamma_2\gamma_1} = M_{\gamma_2} M_{\gamma_1}.$$

Finally, suppose that Ψ is another fundamental matrix at t_0 . Then $\Psi = \Phi M$ for some t_0 and let $\Psi_\gamma = \Psi N_\gamma$. Then we have

$$\Phi M N_\gamma = \Psi N_\gamma = \Psi_\gamma = \Phi_\gamma M = \Phi M_\gamma M,$$

which implies

$$N_\gamma = M^{-1} M_\gamma M.$$

This proves the representation is independent of the choice of fundamental matrix up to conjugation.

Example 2.1.6.2 (Babymost Example). In the rank one case we have a differential equations

$$y'(t) = a(t)y(t), \quad a(t) \in \text{Hol}(U).$$

This has a solution $\phi(t) = \exp(\int_{t_0}^t a(s)ds)$ which is also the fundamental matrix. This formula makes sense in a small disc around t_0 . For things to be interesting we need $\int_\gamma a(s)ds$ to have monodromy.

If $a(t) = 1/t$ this would be the simplest case. This is a little too simple as $\int_{t_0}^t \frac{ds}{s}$ would give a branch of $\log(s)$ which would only change the exponent by $2\pi i$.

If $a(t) = c/t$ for some constant c , then things get a little interesting. One then has $y(t) = t^c := \exp(c \log(t))$ as a solution. In this case if we let γ_0 be a loop around the origin and $M_0 = M_{\gamma_0}$ we find that

$$M_0 = \exp 2\pi i c.$$

□

For any path γ in U starting at $a \in U$ and ending at $b \in U$ and any fundamental matrix Φ in a neighborhood of a we are going to show that we can analytically continue Φ along γ to get a new fundamental matrix Φ^γ which is the analytic continuation of Φ along gamma. There are some issues that we need to address.

1. How do we know that the fundamental matrix doesn't have a natural stopping point where it can't be continued further?

2. How do we know that the continuation Φ^γ doesn't degenerate after leading the initial ball B where the power series defining it converged? How do we know solutions don't become linearly dependent?

Let's address the first issue. Suppose that Φ is analytic in some ball B around a and that there is some a_1 on the boundary of B where Φ doesn't extend. Well since $a \in U$ we know that there exist some Φ_1 a fundamental matrix which is valid in some neighborhood B_1 of a_1 . Then on $B \cap B_1$ we there exists some matrix $M_1 \in \text{GL}_n(\mathbb{C})$ such that

$$\Phi_1 = \Phi M_1.$$

By analytic continuation we could actually extend Φ to $B \cap B_1$ and hence by the sheaf property there exists some unique Φ_2 such defined on $B \cup B_1$ which restricts to Φ and Φ_1 on there respective domains.

Let's now address the second issue. Let $\det(\Phi) = W$. We need to show that $W(t)$ is never zero on these continuations. We know that

$$W'(t) = \text{Tr}(\Phi^{-1}\Phi')W(t).$$

But since $\Phi' = A\Phi$ we have that $\Phi^{-1}\Phi' = \Phi^{-1}(t)A(t)\Phi(t)$ and since trace is invariant under conjugation our scalar equation becomes

$$W'(t) = \text{Tr}(A(t))W(t),$$

and we see that

$$W'(t) = \exp\left(\int_{t_0}^t \text{Tr}(A(s))ds\right),$$

where the integral is understood to be a path integral. This is never zero which implies that $\Phi_\gamma(t)$ always remains a fundamental system of solutions.

Finally, we just want to make the remark that Φ_γ only depends on the homotopy class $[\gamma]$ of γ . This is because path integrals are well-defined on homotopy classes.

Theorem 2.1.6.3. *For every $U \subset \mathbb{C}$ and every $A(t) \in \text{Hol}(U)$ and every $t_0 \in U$, monodromy of a fundamental set of solutions is well-defined and hence induces a well-defined monodromy representation $\pi_1(U, t_0) \rightarrow \text{GL}_n(\mathbb{C})$, given by $[\gamma] \mapsto M_\gamma$ where M_γ is the matrix $\Phi_\gamma = \Phi M_\gamma$ for a fundamental matrix Φ .*

Example 2.1.6.4 (Euler Systems). In a punctured neighborhood around $0 \in \mathbb{C}$, consider the system

$$Y' = \frac{A}{t}Y.$$

Consider the function $t^A = \exp(A \log(t))$ for some branch $\log(t)$ and \exp denoting the matrix exponential. We have

$$\frac{d}{dt} [t^A] = \exp(A \log(t)) A \frac{1}{t} = \frac{A}{t} t^A,$$

so the matrix $\Phi(t) = t^A$ is a local matrix solution of this equation. Since $\det(e^B) = e^{\text{Tr}(B)}$ for any matrix B we have $\det(\Phi) = t^{\text{Tr}(A)}$ which is never zero and hence $\Phi(t)$ is a fundamental matrix.

Now let γ be a loop in U that encloses the origin. We can compute

$$\Phi_\gamma(t) = \exp(A(\log(t) + 2\pi i)) = \Phi(t) \exp(2\pi i A)$$

and hence $M_\gamma = \exp(2\pi i A)$.

Exercise 2.1.6.5. Every matrix $M \in \text{GL}_n(\mathbb{C})$ can appear as the monodromy matrix of some system. (Hint: use the Euler system and show that for every $M \in \text{GL}_n(\mathbb{C})$ there exist some $A \in M_n(\mathbb{C})$ such that $\exp(2\pi i A) = M$. This needs some ideas like a matrix logarithm or using a Jordan canonical form.)

2.2 Classification of Fuchsian Equations

The Fuchsian condition is a condition on meromorphic differential equations that we impose that make it so that solutions aren't divergent. Maybe this isn't obvious but if one applies the power series technique to innocent looking differential equations they can have formal power series solutions which are completely divergent. The next example shows this.

Exercise 2.2.0.1. Consider the equation

$$t^3 y''(t) + (t^2 + t)y'(t) - y(t) = 0.$$

If we expand in a power series we find that for $y(t) = \sum_{n=0}^{\infty} a_n t^n$ to be a solution one had the initial value difference equation

$$\begin{cases} a_0 = 0, \\ a_1 = a, \\ a_n = -(n-1)a_{n-1}. \end{cases}$$

where $a \in \mathbb{C}$ is arbitrary. One finds that

$$y(t) = a \sum_{n=1}^{\infty} (-1)^{n+1} (n-1)! t^n \in \mathbb{C}[[t]] \setminus \mathbb{C}\langle t \rangle$$

is a divergent power series solution! Note that $|a_{n+1}|/|a_n| = n \rightarrow \infty$ as $n \rightarrow \infty$.

So what is the issue? The issue is that when we convert this equation into a first order system of differential equations it has a pole of order bigger than one. One can check that if we let $y'(t) = v(t)$ in the above example we see that $v'(t) = -\frac{t^2+t}{t^3}v(t) + \frac{1}{t^3}y(t)$ and letting $Y(t) = (y(t), v(t))$ we get the first order system

$$Y' = A(t)Y$$

where

$$A(t) = \begin{pmatrix} 0 & 1 \\ -\frac{t+1}{t^2} & \frac{1}{t^3} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{t^3} + \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \frac{1}{t^2} + \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} \frac{1}{t} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

The matrix expansion of $A(t)$ has a pole of order bigger than one at $t = 0$.

In what follows we are going to let \mathbb{P}^1 denote the projective line (equivalently the Riemann sphere). In order to avoid all of the divergent behavior we introduce the notion of a Fuchsian differential equation. We will later prove that these differential equations have “regular singular points”.

Definition 2.2.0.2 (Fuchsian Differential Equations). Consider a rank n first order system of differential equations

$$Y' = A(t)Y. \tag{2.2.1}$$

with $A(t) \in M_n(\text{Hol}(\mathbb{P}^1 \setminus T))$ for $T \subset \mathbb{P}^1$ a finite collection of points. We say the system is *Fuchsian* at $t_0 \in T$ if $A(t)$ has the form

$$A(t) = \frac{B(t)}{t - t_0},$$

where $B(t)$ is holomorphic at t_0 . We say the system is *Fuchsian* at if it is Fuchsian at every point in T .

We extend this concept to higher order differential equations in one variable by saying that they are Fuchsian and Fuchsian at a point if their associated first order system is.

Exercise 2.2.0.3 (Fuch's Criterion For ODEs In One Variable). Consider a univariate holomorphic system on $\mathbb{P}^1 \setminus S$. A first order system

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_0(t) = 0$$

is Fuchsian at $t = t_0$ if and only if the poles of the coefficients are restricted by $\text{mult}_{t=t_0}(a_j(t)) \geq n - j$. This means that the equation takes the form

$$y^{(n)} + \frac{b_{n-1}(t)}{t - t_0}y^{(n-1)} + \cdots + \frac{b_0(t)}{(t - t_0)^n} = 0,$$

where the $b_j(t)$ are holomorphic at $t = t_0$

Example 2.2.0.4 (Airy Equation). Consider the Airy equation

$$y'' = ty.$$

One can see that this system is regular singular at $t \in \mathbb{P}^1 \setminus \infty$. At $t = \infty \in \mathbb{P}^1$ we need to change variables $t = 1/s$ and we find that $dt = \frac{-1}{s^2}ds$ which means $\frac{d}{dt} = -s^2 \frac{d}{ds}$ and

$$\frac{d^2}{dt^2} = s^2 \frac{d}{ds} s^2 \frac{d}{ds} = s^2 \left(s^2 \frac{d}{ds} + 2s \right) \frac{d}{ds} = s^4 \frac{d^2}{ds^2} + 2s^3 \frac{d}{ds},$$

which gives

$$\frac{d^2 y}{ds^2} + \frac{2}{s} \frac{dy}{ds} - \frac{1}{s^5} y = 0.$$

From this we see that there is an irregular singular point at $s = 0$.

Remark 2.2.0.5. There are two ways to compute what $\frac{d^2}{dt^2}$ in the chart at infinity. The first way is to act on an unknown function f by the operator $-s^2 \frac{d}{ds}$ twice and then pretend line you never used the symbol $f = f(s)$ for a computation. The second way is to consider the non-commutative ring $\mathbb{C}[s, \partial]$ subject to the relations $\partial s = s\partial + 1$. This is the a ring of linear differential operators on $\mathbb{C}[s]$ called the *Weyl algebra*. The second way is really equivalent to the first way.

As stated before, we care about Fuchsian differential equations because they tell us that the solutions are nice. By “nice” we mean that the singularities are not out of control. By “out of control” we mean, regular singular. This means that in every sector $S_{t_0}(\alpha, \beta)$, if we approach the points $t = t_0$ with bounded angle of variation then the solution must have at worst a pole. Here is a picture of such a sector:

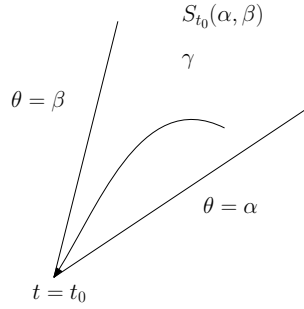


Figure 2.3: A sector used in the definition of regular singular.

In what follows we will let $S_{t_0}(\alpha, \beta) = \{t \in \mathbb{C} : \alpha < \arg(t - t_0) < \beta \text{ where } t_0 \in \mathbb{C}, \alpha, \beta \in [0, 2\pi] \text{ with } \alpha > \beta \text{ and } \arg \text{ the branch of the argument taking valued in } [0, 2\pi)\}$. We will also let $B_R(t_0)$ denote the open disc of radius R centered at t_0 . A set of the form $S_{t_0}(\alpha, \beta) \cap B_R(t_0)$ will be called a bounded sector emanating from $t = t_0$, and a bounded sector contained in another bounded sector as an open set will be called a bounded subsector.

Definition 2.2.0.6. We say that $t = t_0$ is a *regular singular point* if and only if for every local sector at $t = t_0$ there exists a holomorphic basis of solutions $Y_1(t), \dots, Y_n(t)$ with $Y_j(t) = (y_{j1}(t), \dots, y_{jn}(t))^T$ and $\lambda \in \mathbb{C}$ such that

$$\lim_{t \rightarrow t_0} (t - t_0)^\lambda y_{ji}(t) = 0.$$

That seems like a lot but all this is saying is that as you approach your point in question you don't blow up like an essential singularity.

Theorem 2.2.0.7 (Fuch's Criterion). *Solutions of Fuchsian systems only have at worst regular singular points locally.*

Proof. The trick in this proof is to use the isomorphic $\mathbb{C}^n \cong \mathbb{R}^{2n}$ and make estimates as if the functions were real valued. Here we will view $A(t)$ as a function to $M_{2n}(\mathbb{R})$ which is real analytic and a solution $Y(t)$ as a function to \mathbb{R}^{2n} which is real analytic. Also we can observe that a solution $Y(t)$ from this real analytic perspective has $|Y(t)|^2 = Y(t) \cdot Y(t)$. Also, we let $Y'(t)$ denote it's usual real analytic derivative which coincides with its complex analytic one (after again changing the complex analytic one again to real analytic function). We have

$$\frac{d}{dt} [\ln |Y(t)|^2] = \frac{2Y' \cdot Y}{|Y|^2} = 2 \frac{(AY) \cdot Y}{|Y|^2}.$$

Taking norms and using $|V \cdot W| \leq |V| \cdot |W|$ with $|AV| \leq \|A\| \cdot |V|$ we get

$$\left| \frac{d}{dt} [\ln |Y(t)|] \right| \leq \|A(t)\| \leq \frac{C_0}{|t|}$$

where in the last line we used the Fuchsian hypothesis. This then gives along a given contour $\gamma(r) = e^{i\theta_0}(r_0 - r)$ starting at $\gamma(r_0) = t_0 = e^{i\theta_0}r_0$ and ending at $t = e^{i\theta_0}(r_0 - r)$ that

$$\ln |Y(t)| \leq C_1 + \int_{\gamma} \frac{C_0}{|s|} d|s| = C_1 - \int_0^r \frac{C_0}{r - r_0} |e^{i\theta_0} dr| = C_1 - C_0 \ln |r - r_0|$$

which implies that $|Y(t)| \leq e^{C_1} |t|^{-C_0}$. ♠♠♠ Taylor: [\[FIXME, add Gronwall-like statement to appendix\]](#) □

2.3 Hilbert's 21st Problem

We are now in a position to state Hilbert's 21st problem. The Monodromy map associated to every Fuchsian system on \mathbb{P}^1 with poles contained in T a

representation of its fundamental group.

$$\frac{\{\text{Fuchsian systems, rank } n \text{ on } \mathbb{P}^1 \text{ with poles on } T\}}{(\text{global holomorphic gauge trans})} \rightarrow \frac{\{\text{repns } \rho : \pi_1(\mathbb{P}^1 \setminus T, t_0) \rightarrow \text{GL}_n(\mathbb{C})\}}{(\text{matrix conjugation})}$$

Hilbert's 21st problem asks if this map of sets is surjective.

Problem 2.3.0.1 (Hilbert's 21st Problem). Is it the case that every representation $\rho : \pi_1(\mathbb{P}^1 \setminus T) \rightarrow \text{GL}_n(\mathbb{C})$ comes from a Fuchsian differential equation with poles supported on T ?

This problem has a rather crazy history. The problem was first posed by Hilbert in 1900 during the International Congress of Mathematicians (ICM). This is the event where the give out Field's Medals and occurs once every four years. In 1907 Plemelj⁵ published a positive answer to the question. In 1983, Treibich-Koch published a gap in the proof; it turns out that previous work from Dekkers in 1979 implies that the map is indeed surjective in the rank two case. Finally, in 1990, Bolibruch showed that the map is not surjective in rank higher than two disproving the conjecture.

♠♠♠ Taylor: [\[EXPAND\]](#)

2.3.1 Representations of $\pi_1(\mathbb{P}^1 \setminus T)$

The representations $\pi_1(\mathbb{P}^1 \setminus T) \rightarrow \text{GL}_n(\mathbb{C})$ are rather easy to describe. The key observation is that \mathbb{P}^1 minus some points is homotopy equivalent to a bouquet of circles:

$$\mathbb{P}^1 \setminus \{t_1, \dots, t_n\} \approx \underbrace{S^1 \vee \dots \vee S^1}_{(n-1)\text{-times}}$$

where \approx denotes homotopy equivalence and \vee denotes the wedge product of topological spaces. A picture of this homotopy equivalence for $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ is given in Figure 2.3.1. The convenient description allows us to see that the fundamental group is just the free group on $(n-1)$ generators

$$\pi_1(\mathbb{P}^1 \setminus \{t_1, \dots, t_n\}) \cong F_{n-1} \cong \langle \gamma_1, \gamma_2, \dots, \gamma_{n-1} \rangle,$$

⁵Nalini Joshi pronounces this “Plum-ell-i”, I’m not sure how to pronounce this name

the generators then can be taken to be homotopy classes of loops around each of the points t_1, \dots, t_{n-1} . The last loop γ_n around t_n satisfies the relation

$$\gamma_n \cdots \gamma_2 \gamma_1 = 1.$$

You can actually see this loop is trivial if you think about it a little bit.

Anyway, with this description the representations $\pi_1(\mathbb{P}^1 \setminus T) \rightarrow \mathrm{GL}_n(\mathbb{C})$ are determined by tuples $(M_1, M_2, \dots, M_{n-1}) \in \mathrm{GL}_n(\mathbb{C})^{n-1}$ modulo simultaneous conjugation by an element in $\mathrm{GL}_n(\mathbb{C})$. Here $M_j = \rho(\gamma_j)$.

2.3.2 Gauge Transformations

It remain to describe the equivalence relation of differential equations that we are using in the Hilbert's 21st problem. Let (R, ∂) be a ∂ -ring and consider the equation

$$Y' = AY \tag{2.3.1}$$

where $Y = (y_1, \dots, y_n)$ and $Y' = (y'_1, \dots, y'_n)$ and $A \in M_n(R)$. One can change coordinate in this differential equation and suppose that

$$Y = \Phi \tilde{Y} \tag{2.3.2}$$

for some $\Phi \in \mathrm{GL}_n(R)$. In this situation we get a new equation

$$\tilde{Y}' = \tilde{A} \tilde{Y} \tag{2.3.3}$$

which is said to be *gauge equivalent* to the previous equation. We will now compute what \tilde{A} is by plugging $Y = \Phi \tilde{Y}$ into $Y' = AY$. We obtain $Y' = (\Phi \tilde{Y})' = \Phi' \tilde{Y} + \Phi \tilde{Y}'$. We also obtain $AY = A\Phi \tilde{Y}$. Putting these together gives $\Phi \tilde{Y}' = A\Phi \tilde{Y} - \Phi' \tilde{Y}$ or

$$\tilde{Y}' = \tilde{A} \tilde{Y}, \quad \tilde{A} = \Phi^{-1} A \Phi - \Phi^{-1} \Phi'.$$

Both $Y \mapsto \Phi^{-1} Y$ and $A \mapsto A^\Phi := \Phi^{-1} A \Phi + \Phi^{-1} \Phi'$ are called *gauge transformations* and define right group actions of $\mathrm{GL}_n(R)$ on $R^{\oplus n}$ and $M_n(R)$. The equations (2.3.1) and (2.3.3) are called *gauge equivalent*.

For holomorphic and meromorphic linear systems we can consider holomorphic and meromorphic gauge transformations. These gauge transformations

can be local or global. What is interesting is that sometimes we can take a meromorphic linear systems and then convert it into a holomorphic linear systems by some meromorphic gauge transformation. In the case that we can do this the singularities of the original linear system are called *apparent singularities*.

Example 2.3.2.1. Consider the linear system

$$Y' = \begin{pmatrix} 1 & \frac{1}{t^2} - \frac{2}{t} \\ t^2 & 0 \end{pmatrix} Y \quad (2.3.4)$$

which is holomorphic on $\mathbb{C} \setminus \{0\}$. The singularity at $t = 0$ is actually just apparent as it is gauge equivalent to the system

$$\tilde{Y}' = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \tilde{Y}.$$

To see this one uses a meromorphic gauge transformation. The point here is that the singularities of (2.3.4) are just apparent and that they can be removed by using

$$Y = \begin{pmatrix} t^2 & 0 \\ 0 & 1 \end{pmatrix} \tilde{Y}.$$

As an exercise one needs to compute

$$\tilde{A} = \begin{pmatrix} \frac{1}{t^2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} t^2 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \frac{1}{t^2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2t & 0 \\ 0 & 0 \end{pmatrix}.$$

which comes from the formula for gauge transformations.

The most important example is the case where Φ is a fundamental matrix.

Example 2.3.2.2. Consider a linear differential system $Y' = AY$ over a differential ring (R, ∂) . Suppose that the system admits a fundamental matrix $\Phi \in \text{GL}_n(R)$. Then setting $Y = \Phi \tilde{Y}$ we find that

$$A^\Phi = \Phi^{-1}A\Phi - \Phi^{-1}\Phi' = \Phi^{-1}(A\Phi - \Phi') = 0$$

and so that system is gauge equivalent to the trivial system

$$\tilde{Y}' = 0.$$

It is important to observe that there usually aren't *global* fundamental matrices. This is what prevents us from trivializing all differential equations.

2.4 Classification of Fuchsian Differential Equations on \mathbb{P}^1

We are going to show that every Fuchsian differential system on \mathbb{P}^1 with polar locus $S = \{a_1, a_2, \dots, a_m\} \subset \mathbb{P}^1 \setminus \{\infty\}$ takes the form

$$Y' = A(t)Y, \quad A(t) = \frac{A_1}{t - a_1} + \dots + \frac{A_m}{t - a_m},$$

where $A_1, A_2, \dots, A_m \in M_n(\mathbb{C})$ are constant matrices. To do this we first need to review some facts about residues and Riemann surfaces.

2.4.1 Some Reminders About Riemann Surfaces

Riemann Surfaces are just topological spaces equipped with a system of holomorphic charts that make them locally isomorphic to open subsets of \mathbb{C} . A description of these charts for \mathbb{P}^1 is given in Figure 2.4.1. This allows us to make sense of what a holomorphic map is and make sense of what computations “at infinity” are.

One can upgrade this \mathbb{C} to \mathbb{C}^n and get a category of complex manifolds. It turns out that the category of compact Riemann surfaces is equivalent to the category of smooth projective algebraic curves over \mathbb{C} . These are curves which are cut out by homogeneous polynomial equations in some complex projective space \mathbb{P}^n . The same is not true for higher dimensional compact complex manifolds, there exist compact complex manifolds which aren't projective varieties (see for example [Sha13, pg 161]). It is true however, a theorem called Chow's theorem, that every compact complex manifold embedded into complex projective space is a projective variety (i.e. it is cut out by homogeneous equations).

The case of complex projective curves (or equivalently Riemann Surfaces) is especially nice because this category is equivalent to the category of fields $K/\mathbb{C}(t)$ which are algebraic. In the case of connected Riemann surfaces X the naturally associated field is the field of meromorphic functions on X which we denote by $\text{Mer}(X)$. In the case of projective curves C the naturally

associated field is the function field $\kappa(C)$. Miraculously they are isomorphic even though they have drastically different descriptions away from the case of $X = C = \mathbb{P}^1$. This case is rather easy to describe.

We will prove that $\text{Mer}(\mathbb{P}^1) = \mathbb{C}(t)$ which is easily seen to be the fraction field $\mathbb{C}[t]$ of the polynomial functions on one of its open sets. We write $\mathbb{P}^1 = U_0 \cup U_\infty$ and note that some $f \in \text{Mer}(\mathbb{P}^1)$ has finitely many poles on U_0 . This means there exists a polynomial $g(t)$ such that $f(t)g(t)$ is entire. Since $f(t)$ is meromorphic and $g(t)$ is meromorphic the order of vanishing at infinity is finite. Here $\text{ord}_{t=\infty}(f(t)g(t)) = \text{ord}_{s=0}(f(1/s)g(1/s))$. This means that $f(t)g(t) = g(t) \in \mathbb{C}[t]$. Hence $f(t) = h(t)/g(t)$ which proves that every Meromorphic function is rational.

In general $\text{Mer}(X)$ is a finite algebraic extension of $\mathbb{C}(t)$. To give an idea of how different-looking $\text{Mer}(X)$ and $\kappa(X)$ can be consider the case of an elliptic curve E . As a Riemann surface we like to describe this as \mathbb{C}/Λ for some lattice $\Lambda \subset \mathbb{C}$. In this situation, we have $\text{Mer}(E) = \mathbb{C}(\wp_\Lambda(t), \wp'_\Lambda(t))$ where $\wp_\Lambda(t)$ is the Weierstrass \wp -function associated to the lattice Λ . In the case where we want to present E algebraically, then away from ∞ (some curves may have more than one “point at infinity” just not the traditional presentations of \mathbb{P}^1 and E) we have $E \subset \mathbb{C}^2$ given by the equation $y^2 = x^3 + ax + b$ for some $a, b \in \mathbb{C}$. Here we are using (x, y) for complex coordinates. The crazy part is that there is a map $\mathbb{C}/\Lambda \rightarrow E$ given by $x = \wp(t)$ and $y = \wp'(t)$ which gives the isomorphism. Here the point $0 \in \mathbb{C}/\Lambda$ maps to ∞ in the projective model of the elliptic curve.

2.4.2 The Residue Theorem For Meromorphic Differential Forms

We will need the following theorem about the sum of residues being zero later as we try to classify Fuchsian equations. Here we briefly recall that for any meromorphic differential ω on a compact Riemann surface X we can find a local parameter $t = t_b$ at $b \in X$ and then write ω as $f(t)dt$.⁶ We can then

⁶In coordinates on say \mathbb{C} the local parameter for $b \in \mathbb{C}$ is $t_b = t - b$ where t is the usual complex variable.

develop $f(t)$ in a Laurent series to get

$$\omega = \left(\frac{a_{-n}}{t^n} + \cdots + \frac{a_{-1}}{t} + a_0 + a_1 t + \cdots \right) dt$$

and define the residue at b by the usual formula

$$\operatorname{res}_{t=b}(\omega) = a_{-1}.$$

We will extend this to vector valued differential forms $A(t)dt$ by doing this component by component and taking the residues there.

Theorem 2.4.2.1 (Residue Theorem). *Let ω be a meromorphic differential on a compact Riemann surface X . Then $\sum_{a \in X} \operatorname{res}_{t=a}(\omega) = 0$.*

Proof. We give a proof in the case that $X = \mathbb{P}^1$. A complete proof can be found at [Sch14, Proposition 6.6] and those notes can be found online as of 2022 by a simple Google search.

The basic idea as depicted in figure 2.4.2 is to take a simple closed contour γ_1 and its opposite contour γ_2 and realize that on one hand

$$\int_{\gamma_1} \omega + \int_{\gamma_2} \omega = 0$$

while on the other hand we have the classic residue theorem from complex analysis for each of these integrals

$$\int_{\gamma_1} \omega + \int_{\gamma_2} \omega = 2\pi i \sum_{b \in \mathbb{P}^1} \operatorname{res}_{t=b}(\omega).$$

□

For $A(t) = (a_{ij}(t)) \in M_n(\mathbb{C}((t)))$ we will do Laurent series developments entry-by-entry and write

$$A(t) = \sum_{j=-\infty}^{\infty} A_j(t - t_0)^j, \quad A_j \in M_n(\mathbb{C}).$$

For entries which are truly meromorphic, then for closed curves γ we will have

$$\int_{\gamma} A(t)dt = \left(\int_{\gamma} a_{ij}(t)dt \right).$$

As above if $A(t)dt$ is a matrix of meromorphic differential forms with poles at t_1, \dots, t_m on a Riemann surface X with residues R_j for $1 \leq j \leq m$ then we get

$$\sum_{j=1}^m R_j = 0.$$

Corollary 2.4.2.2. *The sum of the residues of meromorphic matrix valued differential forms is zero.*

2.4.3 Classification of Fuchsian Differential Equations on $\mathbb{P}^1 \setminus S$

Fuchsian differential equations on \mathbb{P}^1 have a very simple form. For a polar locus $S = \{s_1, \dots, s_m\}$ we will often assume that S takes the form $S = \{0, 1, \infty, s_4, \dots, s_m\}$ which we can do by using a Möbius transformation. We can also if we want assume that the polar locus not contain ∞ .

Theorem 2.4.3.1. *Consider a Fuchsian differential equation on \mathbb{P}^1 ,*

$$Y' = A(t)Y, \quad A(t) \in M_n(\mathbb{C}(t)).$$

If $A(t)$ has polar locus $S = \{s_1, \dots, s_m\} \subset \mathbb{P}^1$ not containing infinity then

$$A(t) = \frac{A_1}{t - s_1} + \dots + \frac{A_m}{t - s_m}$$

where $A_j \in M_n(\mathbb{C})$ and $A_1 + \dots + A_m = 0$.

Proof. We can suppose without loss of generality that there are no poles at infinity. To prove this one needs to recall that by the Mittag-Leffler theorem [?, Proposition 2.19] if $f(z)$ is meromorphic on \mathbb{C} with poles at s_1, \dots, s_m then there exists $p_j(z) \in \mathbb{C}[z]$ polynomials such that

$$f(z) - \sum_{j=1}^m p_j\left(\frac{1}{z - s_j}\right)$$

is entire and the degree of p_j is the order of the pole of f at $z = s_j$. In our application we have that $A(t)$ has at most a pole at each s_j . Hence there exists some matrices $A_1, \dots, A_m \in M_n(\mathbb{C})$ such that the components of

$$B(t) = A(t) - \frac{A_1}{t - s_1} - \dots - \frac{A_m}{t - s_m}$$

are holomorphic on \mathbb{C} . Also note that $A_j/(t - s_j) = A_j s/(1 - s s_j)$ is also holomorphic at $s = 0$ or $t = \infty$. This means that $B(t)$ is entire and bounded and hence constant. But we know that $\lim_{t \rightarrow s_j} B(t) = 0$ by construction which means that it must be the constant function zero.

The second part about the sum of the residues being zero follows from the Residue theorem (§2.4.2) but doing it component by component in the matrix $A(t)dt$. \square

The residue matrices are so important we give them a name. They are called the *local exponents* of the linear differential equation. We will see that if ρ is a local exponent of $A(t)/t$ where $A(t) \in M_n(\mathbb{C}[[t - t_0]])$ at $t = t_0$ with the property that $\rho + r$ is not an eigenvalue for any integer $r > 1$ then the system admits a solution of the form $Y(t) = (t - t_0)^\rho Z(t)$ where $Z(t) \in \mathbb{C}[[t]]^n$. If $A(t)$ is holomorphic then $Z(t)$ will be holomorphic.

We conclude this subsection with a classification of equations with one, two, three, and four singular points. The case of three singular points will end up leading to the theory of hypergeometric differential equations. The case of four singular points ends up leading to the theory of isomonodromic deformations and P_{VI} the 6th Painlevé equation.

In the following examples the singular locus $S = \{s_1, \dots, s_m\} \subset \mathbb{P}^1$ can be taken without loss of generality to be $S = \{0, 1, \infty, s_4, \dots, s_m\}$ since any three points can map to any other three points by a Möbius transformation.

Example 2.4.3.2 (one singular point). Consider a Fuchsian differential equations with $S = \{0\}$. Then we have

$$\frac{dY}{dt} = \frac{A_0}{t} Y$$

for some constant matrix $A_0 \in M_n(\mathbb{C})$. We then can use the chart at infinity $\partial_t = -s^2 \partial_s$ to conclude that the equation becomes $-s^2 \frac{dY}{ds} = s A_0 Y$ which

gives

$$\frac{dY}{ds} = -\frac{A_0}{s}Y,$$

which is not holomorphic at ∞ unless $A_0 = 0$. Hence every such system is equivalent to

$$Y' = 0.$$

Example 2.4.3.3 (two singular points). Consider a Fuchsian differential equation with polar locus $S = \{0, \infty\}$. From the previous example we see that it has the form

$$\frac{dY}{dt} = \frac{A_0}{t}$$

and that $A_0 = -A_\infty$.

The case of three singular points is sometimes called the Gauss case of the hypergeometric case because of its connections to the hypergeometric differential equations.

Example 2.4.3.4 (three singular points). Consider a Fuchsian differential equation with polar locus $S = \{0, 1, \infty\}$. Such an equation has the form

$$\frac{dY}{dt} = \left(\frac{A_0}{t} + \frac{A_1}{t-1} \right) Y$$

where $A_0 + A_1 + A_\infty = 0$. Explicitly after changing coordinates to the chart at infinity (letting $t = 1/s$) we find

$$-s^2 \frac{dY}{ds} = \left(\frac{A_0}{1/s} + \frac{A_1}{1/s-1} \right) Y$$

which implies

$$\frac{dY}{ds} = - \left(\frac{A_0 + A_1}{s} + \frac{A_1}{1-s} \right) Y,$$

and we can see $A_\infty = -A_0 - A_1$ explicitly.

In section §2.5 we will show every rank two Fuchsian equation with polar locus $S = \{a, b, c\} \subset \mathbb{P}^1$ can be reduced to the Gauss hypergeometric equation in a single dependent variable

$$t(t-1)y'' + (c - (a+b+1)t)y' - aby = 0.$$

The case of four singular points is sometimes called the Painlevé case because of its connections to the Painlevé equations.

Example 2.4.3.5 (four singular points). Every Fuchsian differential equation with polar locus containing four points now cannot be normalized to a standard set of points. We can bring the first three points of S to $0, 1, \infty$ but a third point $\lambda \in \mathbb{C}$ remains. We will have $S = \{0, 1, \infty, \lambda\}$ and the differential equation will take the form

$$\frac{dY}{dt} = \left(\frac{A_0}{t} + \frac{A_1}{t-1} + \frac{A_\lambda}{t-\lambda} \right) Y$$

where $A_0, A_1, A_\lambda \in M_n(\mathbb{C})$ and we define A_∞ by the sum of the residues being zero $A_0 + A_1 + A_\lambda + A_\infty = 0$.

A fun game to play here will be to determine the conditions under which we may vary A_λ as a function of λ and preserve the monodromy representation. Note that $\pi_1(\mathbb{P}^1 \setminus \{0, 1, \infty, \lambda_1\}) \cong \pi_1(\mathbb{P}^1 \setminus \{0, 1, \infty, \lambda_2\})$ for every pair of λ_1 and λ_2 so it makes sense to ask for monodromy representations to change. Such deformations are called *isomonodromic*. The criterion for deformations to be isomonodromic are given by Schlesinger's equations for matrices which give rise to the Painlevé equations.

2.5 Hypergeometric Differential Equations: Fuchsian Differential Equations of rank two on $\mathbb{P}^1 \setminus \{0, 1, \infty\}$

The Gauss hypergeometric equation is the following homogeneous ordinary differential equation

$$y'' + \frac{(a+b-1)t-c}{t(1-t)}y' + \frac{ab}{t(1-t)}y = 0. \quad (2.5.1)$$

Its solutions are so-called hypergeometric functions and has the remarkable property that any rank two Fuchsian differential equations on \mathbb{P}^1 can be reduced to this equation for some collection of parameters (a, b, c) . The

(a, b, c) really encode eigenvalues of the residue matrices (=local exponents) and the form of the equation is really a consequence of the restriction of these local exponents in part due to Fuch's theorem which says that the sum of the local exponents in this rank 2 case with three singularities must be equal to one.

Exercise 2.5.0.1 (Hypergeometric Exponents). Show that the local exponents of the hypergeometric differential equation with parameters (a, b, c) fall into the following table:

0	1	∞
0	0	a
$1 - c$	$c - a - b$	b

This is sort of a hard computation now but gets easier once more tools are developed in the rest of the section. I would try it now for 30 minutes, then try it again once you have the indicial equation, then revisit it once more once you have Fuchs' relation.

The first mystery is showing that $n \times n$ first order linear systems can actually be reduced to an order n equation in one dependent variable. This is sort of the opposite of taking an equation of high order and reducing it to an equation a low order but in more variables. The key to this is the theory of D -modules. We will show that linear systems gives rise to a D -modules and conversely any D -module with a choice of basis gives a linear differential equation. Changing the basis will changes the differential equation by a gauge transformation. Now knowing that D -modules encode linear differential equations we apply Katz's theorem and show that D -modules will admit so called cyclic vectors. In the case of rank two Fuchsian differential systems with three poles this reduces our equation to an order two equation linear equation in one dependent variable.

Finally, once we are in the order two case we need to show that all of our equations are determined by the local exponents and that we can manipulate these exponents by a series of gauge transformations and automorphisms of \mathbb{P}^1 to bring our general equations into the Gauss hypergeometric case. This involves a basic lemma about how local exponents change under Gauge transformations of the form $Y(t) = t^\rho \tilde{Y}(t)$.

2.5.1 Weyl Algebras and D -Modules

Let (R, Δ) be a Δ -algebra.

Definition 2.5.1.1. A *Weyl algebra* associated to (R, Δ) is the ring $R[\Delta]$ of linear operators on associated to (R, Δ) . It is the non-commutative ring $R[\partial: \partial \in \Delta]$ where one has

$$\partial a = a\partial + \partial(a), \quad a \in R, \partial \in \Delta.$$

One also has $\partial_1\partial_2 = \partial_2\partial_1$ for $\partial_1, \partial_2 \in \Delta$.

The idea behind the formula $\partial a = a\partial + \partial(a)$ comes from looking at $a \in R$ when viewed as an element $a \in R[\Delta]$ as the linear operator “multiplication by a ”. In this situation we have

$$(\partial a) \cdot f = \partial(af) = \partial(a)f + a\partial(f) = [\partial(a) + a\partial] \cdot f,$$

which justifies the rule.

Definition 2.5.1.2. A D -module is a $R[\Delta]$ -module. We will simply call these $R[\Delta]$ -modules.

Authors like to get cutesy with the above definition and it is worth pointing some things out. First, many authors define $\mathcal{D} = R[\Delta]$ and then talk about \mathcal{D} -modules. See for example Singer and van der Put. Some authors only define Weyl-algebras for polynomial rings and refer to this particular Weyl algebra as *the* Weyl algebra. In this case they take $R = \mathbb{C}[x_1, \dots, x_m]$ with $\Delta = \{\partial_{x_1}, \dots, \partial_{x_m}\}$ and then only talk about Weyl algebras (as we have defined above) as the only Weyl algebras. This is useful when searching the literature for propositions about Weyl algebras that you need. Finally, many authors restrict to the case $\Delta = \{\partial\}$ which will be the case we are interested in mostly and call these ∂ -modules. In this case some authors (like Nick Katz) like to define D as the derivation operator on the module V which satisfies $D(av) = \partial(a)v + aD(v)$ for $v \in V$ and $a \in R$. I reserve the right to use a mixture of these perspectives (and you should too).

2.5.2 Linear Systems and D -Modules

There is a procedure for converting between linear differential equations and D -modules which will be useful that we will now explain. In this subsection we will restrict to the case of a single derivative.

Given a rank n linear system over (R, ∂) given by

$$Y' = AY, \quad A = (a_{ij}) \in M_n(R),$$

we can define a D -module structure on $V = R^{\oplus n}$. Let e_1, \dots, e_n be a standard basis for V . Then we define

$$D(e_j) = \sum_{i=1}^n a_{ij} e_i.$$

Let $V_0 = (R^\partial)^{\oplus n}$. We now have a R^∂ -linear operator on V_0 and we extend this to all of V by specifying

$$D(bv_0) = \partial(b)v_0 + bD(v_0), \quad v_0 \in V_0, b \in R.$$

Exercise 2.5.2.1. Check that this is well defined. This means that if $bv_0 = cw_0$ for some other $c \in R$ and $w_0 \in V_0$ then $D(bv_0) = D(cw_0)$. [This is a silly easy problem.]

Conversely, given a D -module structure on $V = R^{\oplus n}$ one then takes a basis v_1, \dots, v_n and finds that

$$D(v_j) = \sum_{i=1}^n a_{ij} v_i$$

for some $a_{ij} \in R$. This allows us to set up a linear differential equation

$$Y' = AY, \quad A = (a_{ij}) \in M_n(R).$$

One then finds that the linear differential equation associated to the D -module is again the D -module with v_1, \dots, v_n identifying with the standard basis vectors.

If instead we had chosen a different basis one can check that one will obtain a new differential equation

$$\tilde{Y}' = \tilde{A}\tilde{Y}$$

which is gauge equivalent to the first equation. This gives us a procedure for assigning a linear system of rank n over R (up to gauge equivalent) to every $R[\partial]$ -module V of finite rank n . The point here is that change of basis of the D -module gives rise to a gauge transformation of the associated linear system.

Exercise 2.5.2.2. Show that indeed a change of coordinates on the R -module induces a gauge transformation of the linear differential equation.

2.5.3 Cyclic Vectors and Katz's Theorem

In order to convert first order linear systems of rank n into linear differential equations of order n in a single variable we need the notion of a cyclic vector.

Definition 2.5.3.1. An $R[\Delta]$ -module V is *cyclic* if and only if there exists some $v \in V$ such that $V = R[\Delta] \cdot v$. Such a vector $v \in V$ where $V = R[\Delta] \cdot v$ is called a *cyclic vector*.

In the case that $\Delta = \{\partial\}$ and $V \cong R^n$ a $R[\partial]$ -module a vector $v \in V$ is cyclic if and only if

$$v, \partial(v), \dots, \partial^{n-1}(v)$$

form a basis for V . This is probably the most important case. Before proving such cyclic vectors exist, let's take a moment to realize our goal reducing a first order linear system of rank n to an order n linear differential equation in a single dependent variable.

Following our procedure we set $v_0 = v$ and $v_i = \partial^i(v)$ which gives us $\partial(v_i) = v_{i+1}$ for $0 \leq i \leq n-2$ and then $\partial(v_{n-1}) = b_0 v_0 + b_1 v_1 + \dots + b_{n-1} v_{n-1}$ for $b_i \in R$ and we get the linear systems

$$Y' = BY, \quad B = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ b_0 & b_1 & b_2 & \cdots & b_{n-1} \end{pmatrix}$$

which gives the linear differential equation

$$y^{(n)} = b_0 y + b_1 y' + b_2 y'' + \cdots + b_{n-1} y^{(n-1)}.$$

The following Theorem can be found in [Kat87] (which is just five pages including citations).

Theorem 2.5.3.2 (Katz's Theorem). *Let (R, ∂) be a ∂ -ring with $t \in R$ satisfying $\partial(t) = 1$. Let V be a free R -module of finite rank n which has the structure of a $R[\partial]$ -module. Then if R is local, and $(n-1)!$ is invertible in R then V admits a cyclic vector of the form*

$$v = \sum_{j=0}^{n-1} \frac{(t-a)^j}{j!} \sum_{i=0}^j \binom{j}{i} D^i(e_{j-i})$$

where $a \in R^\partial$ and e_1, \dots, e_n is the standard elementary basis for R^n .

We will prove this for $R = \mathbb{C}[[t]]$.

Lemma 2.5.3.3. *Let $V = \mathbb{C}[[t]]^{\oplus n}$ be a D -module.*

1. *Let $h_0, \dots, h_{n-1} \in V$ be horizontal (i.e. suppose $\partial(h_j) = 0$). Consider*

$$v = \sum_{j=0}^n \frac{t^j}{j!} h_j. \text{ The vector } v \text{ is cyclic.}$$

2. *For each $v_0 \in V$ consider the system*

$$\begin{cases} v \equiv v_0 \pmod{tV} \\ \partial(v) = 0 \end{cases} \quad (2.5.2)$$

The element $v := e^{-t\partial} v_0 = \sum_{j \geq 0} (-1)^j \frac{t^j}{j!} \partial^j(v_0)$ is t -adically convergent and is the unique element in V satisfying (2.5.2).

Proof. Taking derivatives we have

$$\begin{aligned} v &= h_0 + th_1 + \frac{t^2}{2!} h_2 + \cdots + \frac{t^{n-1}}{(n-1)!} h_{n-1} \\ \partial(v) &= h_1 + th_2 + \frac{t^2}{2!} h_3 + \cdots + \frac{t^{n-2}}{(n-2)!} h_{n-2} \\ &\vdots \\ \partial^{n-1}(v) &= h_{n-1} \end{aligned}$$

starting from the bottom of the list and going up one can see linear independence as they each introduce a new h_j .

To prove the second part we just compute the derivative of v and expand using the product rule term by term. For uniqueness, suppose that w is another solution. One then has $w = v + tu$ for some $u \in V$. We then get $u + t\partial(u) = 0$, by $0 = \partial(w) = \partial(v) + \partial(tu) = \partial(tu)$. We can expand u in a power series to get $u(t) = \sum_j a_j t^j$ and we find that $\partial(u(t)) = u^\partial(t) + u'(t)$ which gives $a_0 = 0$ and then $a_j^\partial + (j+1)a_{j+1} + a_{j+1} = 0$. This allows us to conclude all of the $a_j = 0$ inductively. \square

The proof of the following theorem will use Nakayama's Lemma which can be found in Atiyah-MacDonald [AM16, pg 21].

Proof of Katz's Theorem for Formal Power Series. Let $V = R^n$. Let e_0, \dots, e_{n-1} be a basis, then it is a basis modulo tV . Hence by Nakayama, $\tilde{e}_j := e^{-t\partial}e_j$ is also a basis for V since it is a basis modulo tV . Furthermore, by the Lemma $\partial(\tilde{e}_j) = 0$. We now apply part one of Lemma 2.5.3.3 to get

$$\begin{aligned} \sum_{j=0}^{n-1} \frac{t^j}{j!} \tilde{e}_j &= \sum_{j=0}^{n-1} \frac{t^j}{j!} \sum_{i \geq 0} \frac{t^i}{i!} \partial^i(e_j) \\ &= \sum_{j=0}^{n-1} \sum_{i \geq 0} \frac{t^{i+j}}{i!j!} \partial^i(e_j). \end{aligned}$$

We can trim this down (using Nakayama again). If v is cyclic then $v + t^n c$ is also cyclic. The "large" power t^n ensures that it remains a basis after n derivatives. This allows us to kill off terms with $j + k \geq n$. Hence

$$\sum_{j=0}^{n-1} \sum_{i=0}^{n-1-j} (-1)^i \frac{t^i}{i!} \partial^i(e_j)$$

gives a cyclic vector. \square

2.5.4 Local Exponents

Consider a first order meromorphic system of rank n on \mathbb{P}^1 given by

$$Y' = A(t)Y, \quad A(t) \in M_n(\mathbb{C}(t)).$$

The eigenvalues of residue matrices play such an important role in the local behavior of solutions of differential equations we give them a name.

Definition 2.5.4.1. Let $R = \text{Res}_{t=t_0}(A(t)) \in M_n(\mathbb{C})$ be a residue at $t = t_0$. An eigenvalue of R is called a *local exponent* of the system at $t = t_0$.

If $S = \{s_1, s_2, \dots, s_m\}$ is the polar locus for a differential equation of rank n with eigenvalues $\rho_1(s_j), \rho_2(s_j), \dots, \rho_n(s_j)$ at the points j we will often write down a so-called *Riemann table* in the form

s_1	s_2	\cdots	s_m
$\rho_1(s_1)$	$\rho_1(s_2)$	\cdots	$\rho_1(s_m)$
$\rho_2(s_1)$	$\rho_2(s_2)$	\cdots	$\rho_2(s_m)$
\vdots	\vdots	\ddots	\vdots
$\rho_n(s_1)$	$\rho_n(s_2)$	\cdots	$\rho_n(s_m)$

We will now go on to show that solutions of $Y(t)$ locally have the form $t^\rho Z(t)$ for ρ “non-resonant” local exponents. We say that an eigenvalue ρ of R is non-resonant provided there doesn’t exist another eigenvalue μ of R such that $\rho - \mu \in \mathbb{Z}$.

2.5.5 Theta Operator and Indicial Equations

Consider a linear differential equation in one variable

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_0(t) = 0$$

which is formally Fuchsian at $t = 0$ so that

$$b_j(t) := t^{n-1}a_j(t) \in \mathbb{C}[[t]].$$

We wish to derive an equation for the local exponents of this equation at $t = 0$. To do this it will be convenient to write our operator (which we view as an element of the Weyl algebra)

$$L = \partial^n + a_{n-1}(t)\partial^{n-1} + \cdots + a_0(t) \in \mathbb{C}[[t]][\partial]$$

in terms of the *theta operator*

$$\theta = t\partial_t \in \mathbb{C}[[t]][\partial].$$

We remark that this operator is called the *Euler operator* in [IKSY91] and is denoted by δ .

The following basic identities will be useful.

Exercise 2.5.5.1. In this problem $\partial = \partial_t$. Show that

1. $t^n \partial^n = \theta(\theta - 1) \cdots (\theta - n + 1)$
2. $\theta(t^\rho f) = t^\rho(\theta + \rho)f$.
3. $(\theta + \rho)t^j = (j + \rho)t^j$ for all $j \geq 0$

If we let $M = t^n L$ then we see that

$$M = \sum_{j=0}^n a_j t^{n-j} t^j \partial^j = \sum_{j=0}^n b_j \theta(\theta - 1) \cdots (\theta - j + 1).$$

For concreteness we write out the order two case.

Example 2.5.5.2. We have $L = a_0 + a_1 \partial + \partial^2$ and $M = b_0 + b_1 \theta + \theta(\theta - 1) = b_0 + (b_1 - 1)\theta + \theta^2$. We can be even more explicit with $b_0 = t^2 a_0$ and $b_1 = t a_1$ so that

$$M = t^2 a_0 + (t a_1 - 1)\theta + \theta^2.$$

One also has a cute form of the hypergeometric differential equation.

Exercise 2.5.5.3. Check that the hypergeometric equation has the form

$$\theta(\theta + 1 - c)y - t(\theta + a)(\theta + b)y = 0.$$

Now in order to derive the indicial equation for the local exponents of a linear differential operator we will seek solutions of $My = 0$ in the form

$$f(t) = t^\rho \sum_{j=0}^{\infty} c_j t^j$$

and conclude a necessary identity about the exponent $\rho \in \mathbb{C}$. To proceed we write each $b_i(t)$ for $1 \leq i \leq n$ as $b_i(t) = \sum_{j=0}^{\infty} b_{ij}t^j$. We then just proceed with a computation

$$\begin{aligned}
 Mf &= \left(\sum_{i=0}^n \sum_{j=0}^{\infty} b_{ij}t^j \theta^i \right) \left(t^\rho \sum_{k=0}^{\infty} c_k t^k \right) \\
 &= t^\rho \sum_{i=0}^n \sum_{j=0}^{\infty} b_{ij}t^j (\theta + \rho)^i \sum_{k=0}^{\infty} c_k t^k \\
 &= t^\rho \sum_{i=0}^n \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} b_{ij}c_k t^j (k + \rho)^i t^{k+j} \\
 &= t^\rho \sum_{m=0}^{\infty} \left(\sum_{i=0}^n \sum_{j=0}^m b_{ij}c_{m-j} (\rho + m - j)^i \right) t^m.
 \end{aligned}$$

If $Mf = 0$ as an element of $\mathbb{C}[[t]][t^\rho]$ then by the linear independence of t^m this gives a system of equations for each m given by

$$\sum_{i=0}^n \sum_{j=0}^m b_{ij}c_{m-j} (\rho + m - j)^i = 0.$$

In the case $m = 0$ we can pull out c_0 and use that $b_{i0} = b_i(0)$ and the indicial equation.

Theorem 2.5.5.4 (Indicial Equation). *If $t^\rho f(t)$ is a formal solution of L then ρ is a solution of*

$$b_0(0) + b_1(0)\rho + \cdots + b_{n-1}(0)\rho^{n-1} + \rho^n = 0. \quad (2.5.3)$$

Equation 2.5.3 is called the indicial equation for the differential equation at $t = 0$. It is an exercise to compute derive the indicial equation at other points.

The basic idea is to use $\theta = (t - t_0) \frac{d}{d(t - t_0)}$ rather than $t \frac{d}{dt}$. Similarly, for an equation at infinity one needs to change coordinates to s where $t = 1/s$ and $-s^2 \frac{d}{ds} = \frac{d}{dt}$.

For a later application to Fuch's relation it will be useful to compute $b_{n-1}(0)$ explicitly the second to top coefficient is always the sum of the roots:

$$b_{n-1}(0) = -\rho_1 - \rho_2 - \cdots - \rho_n.$$

This formula is related to a residue and we will later apply the Residue theorem.

Corollary 2.5.5.5. $b_{n-1}(0) = \text{res}_{t=0}(a_{n-1}(t)dt) - \binom{n}{2}$.

Proof. One sees that

$$\begin{aligned} M &= \sum_{j=0}^n a_j t^{n-j} \theta(\theta-1) \cdots (\theta-j+1) \\ &= \theta^n + (-1-2-\cdots-(n-1))\theta^{n-1} + ta_{n-1}(t)\theta^{n-1} + \cdots \\ &= \theta^n + \left(a_{n-1}(t)t - \binom{n}{2} \right) \theta^{n-1} + \cdots \end{aligned}$$

and hence the statement follows. \square

Exercise 2.5.5.6. For this problem consider a Fuchsian differential equation

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_0(t)y = 0,$$

where $a_j(t) \in \mathbb{C}(t)$ for $0 \leq j \leq n-1$. Let $S \subset \mathbb{P}^1$ be the polar locus of this equation.

1. Show that the general indicial equation at $t = t_0 \neq \infty$ takes the form

$$\sum_{j=0}^n c_j \rho(\rho-1) \cdots (\rho-j+1) = 0$$

where $c_j = \lim_{t \rightarrow t_0} (t - t_0)^{n-j} a_j(t)$.

2. Show that if $t_0 = \infty \in S$ then the indicial equation becomes

$$\sum_{j=0}^n (-1)^j c_j \rho(\rho-1) \cdots (\rho-j+1) = 0$$

where $c_j = \lim_{t \rightarrow \infty} t^{n-j} a_j(t)$.

Using the formulas above one now has a more systematic approach to computing the local exponents for the hypergeometric functions.

Exercise 2.5.5.7. Compute the local exponents of the hypergeometric equation

$$y'' + \frac{(a+b-1)t-c}{t(1-t)}y' + \frac{ab}{t(1-t)}y = 0,$$

for each $t_0 \in S = \{0, 1, \infty\}$ using the indicial formulas.

2.5.6 Fuchs' Relation

We now state Fuchs relation which tells us more about the local exponents of the hypergeometric function (for example).

Theorem 2.5.6.1 (Fuchs' Relation). *Consider a Fuchsian linear differential equation on \mathbb{P}^1 of the form*

$$y^{(n)} + a_{n-1}(t)y^{(n-1)} + \cdots + a_0(t)y = 0.$$

Let $S \subset \mathbb{P}^1$ denote the polar locus of the differential equation and assume that $\infty \in S$. If $\rho_1(a), \dots, \rho_n(a)$ denote the local exponents at $a \in S$ then

$$\sum_{a \in S} (\rho_1(a) + \cdots + \rho_n(a)) = (\#S - 2) \binom{n}{2}.$$

Fuchs' relation imposes an extra constraint on the possible eigenvalues of matrices. Note that in the case $n = 2$ and $\#S = 3$ (the hypergeometric case) we have

$$\sum_{a \in S} (\rho_1(a) + \rho_2(a)) = 1.$$

Exercise 2.5.6.2. Check that the exponents in the Riemann table in Exercise 2.5.0.1 satisfy Fuchs' relation.

We now give the proof of Fuchs' relation.

Proof. We know that for $a \in S \setminus \infty$ we have

$$\rho_1(a) + \rho_2(a) + \cdots + \rho_n(a) = \binom{n}{2} - \operatorname{res}_{t=a}(a_{n-1}(t)dt),$$

similarly for $a = \infty$ we have

$$\rho_1(\infty) + \rho_2(\infty) + \cdots + \rho_n(\infty) = -\binom{n}{2} - \operatorname{res}_{t=\infty}(a_{n-1}(t)dt).$$

Using the residue formula (Theorem 2.4.2.1) we have

$$\begin{aligned}
0 &= - \sum_{a \in S} \operatorname{res}_{t=a} (a_{n-1}(t) dt) \\
&= \sum_{a \in S \setminus \infty} \left(\rho_1(a) + \rho_2(a) + \cdots + \rho_n(a) - \binom{n}{2} \right) \\
&\quad + \rho_1(\infty) + \rho_2(\infty) + \cdots + \rho_n(\infty) + \binom{n}{2} \\
&= -(\#S - 2) \binom{n}{2} + \sum_{a \in S} (\rho_1(a) + \rho_2(a) + \cdots + \rho_n(a)),
\end{aligned}$$

which proves the result. \square

2.5.7 Local Solutions of Exponent ρ

Consider a first order system of rank n which is formally Fuchsian at $t = 0$. We will write

$$Y' = \frac{A(t)}{t} Y, \quad A(t) \in M_n(\mathbb{C}[[t]]).$$

Note that this system is equivalent to $\theta(Y) = A(t)Y$ where θ operators component-by-component. We will let expand $A(t)$ in a power series

$$A(t) = A_0 + A_1 t + \cdots,$$

and then consider power series solutions of the form $Y(t) = t^\rho Z(t)$ and develop $Z(t)$ as a power series

$$Z(t) = Z_0 + Z_1 t + \cdots.$$

We then find that

$$\theta(Y) = \theta(t^\rho Z) = t^\rho(\theta + \rho)Z, \quad AY = t^\rho AZ,$$

which leads us to

$$\begin{aligned}
(\theta + \rho)Z(t) &= \rho Z_0 + (\rho Z_1 + Z_1)t + (\rho Z_2 + 2Z_2)t^2 + \cdots, \\
A(t)Z(t) &= A_0 Z_0 + (A_1 Z_0 + A_0 Z_1)t + (A_2 Z_0 + A_1 Z_1 + A_0 Z_2)t^2 + \cdots
\end{aligned}$$

which when we equate coefficients tells us that Z_0 is an eigenvector of A_0 with eigenvalue ρ and that for $n \geq 1$ we have the equation

$$nZ_n + \rho Z_n = A_0 Z_n + A_1 Z_{n-1} + \cdots + A_n Z_0.$$

This equation allows us to solve inductively as long as $(\rho + n)$ is not an eigenvalue of A_0 for $n \geq 1$ since we have the expression

$$(\rho + n + A_0)Z_n = A_1 Z_{n-1} + \cdots + A_n Z_0,$$

and $\rho + n$ not being an eigenvalue puts $(A_0 - \rho - n)$ invertible. The fancy word for this is that $\rho + n$ is in the resolvent set of the operator A_0 (the resolvent set of a linear operator L is precisely the set of λ such that $\lambda - L_0$ is invertible). We will omit the proof of convergence. This has to do with estimating the operator norm of $(x - A_0)^{-1}$ for x in the resolvent set.

This proves the following.

Theorem 2.5.7.1. *Consider the formal Fuchsian system*

$$Y' = \frac{A(t)}{t}Y, \quad A(t) \in M_n(\mathbb{C}[[t]]). \quad (2.5.4)$$

Let $A_0 \in M_n(\mathbb{C})$ be the residue of $A(t)/t$ at $t = 0$ and let ρ be an eigenvalue such that $\rho + n$ is not an eigenvalue of A_0 for any integer $n \geq 1$. Then (2.5.4) admits a formal solution $Y(t) \in \mathbb{C}[[t]]^n$ of the form

$$Y(t) = t^\rho(Y_0 + Y_1 t + \cdots)$$

where Y_0 is an eigenvector of A_0 . Moreover the series is convergent is the series for $A(t)$ is.

Note that even the resonant case where there are eigenvalues ρ and μ with $\rho - \mu \in \mathbb{Z}$ then still one of these admits a solution of the type above. One just needs some eigenvalue such that there is no positive integer that gives another. If ρ and μ are equal then we don't need to worry about this. If $\rho - \mu$ is negative then we don't need to worry about this. If $\rho - \mu$ is positive then we can switch the role of ρ and μ and again not worry about this.

Lemma 2.5.7.2 (Exponent Shifting). *The gauge transformation $Y = (t - t_0)^\mu \tilde{Y}$ has the effect of $A(t)/(t - t_0) \mapsto \tilde{A}(t)/(t - t_0)$ where $\tilde{A}(t) = A(t) - \mu$.*

Proof. Without loss of generality we can suppose that $t_0 = 0$ since ∞ is invariant under the transformation $t \mapsto t - t_0$. The system has the form

$$\theta Y = A(t)Y'$$

where $A(t) = A_0 + A_1 t + \cdots \in M_n(\mathbb{C}[[t]])$ and $A_j \in M_n(\mathbb{C})$ for $j \geq 0$. Then $\theta(t^\mu \tilde{Y}) = t^\mu(\theta + \mu)\tilde{Y}$ and $A(t)t^\mu \tilde{Y} = t^\mu A(t)\tilde{Y}$ which gives the equation

$$\theta \tilde{Y} = (A(t) - \mu)\tilde{Y}.$$

One can check that $\sigma_p(A_0 - \mu) = \sigma_p(A_0) - \mu$ where $\sigma_p(B)$ denotes the eigenvalues of a matrix B . \square

2.5.8 Local Exponents Determine Equations in Hypergeometric Case

In this subsection we work in the Fuchsian case where $\#S = 3$ and rank two. In particular we work with single ordinary differential equations of the form

$$y'' + a_1(t)y' + a_0(t)y = 0$$

with $a_1(t), a_0(t) \in \mathbb{C}(t)$ since all rank two order one systems are equivalent to order two differential equations in one variable.

Theorem 2.5.8.1. *Let $S = \{t_1, t_2, t_3\} \subset \mathbb{P}^1$ and fix a table of local exponents satisfying Fuchs' relation*

$$\begin{array}{ccc} t_1 & t_2 & t_3 \\ \hline \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' \end{array}.$$

There exists a unique $a_1(t), a_0(t) \in \mathbb{C}(t)$ such that

$$y'' + a_1(t)y' + a_0(t)y = 0 \tag{2.5.5}$$

is a Fuchsian differential equation with polar locus S and exponents as given in the table.

Proof. The proof is a partial fraction expansion computation and follows [IKSY91, Chapter 2, Proposition 1.1.1] closely. We assume without loss of generality that $t_3 = \infty$. \square

Theorem 2.5.8.2. *The equation 2.5.5 reduces to the hypergeometric differential equation.*

Proof. Using a Möbius transformation we can transform (t_1, t_2, t_3) to $(0, 1, \infty)$ giving a new order two differential equation with $S = \{0, 1, \infty\}$. This gives a new exponent table

$$\begin{array}{ccc|ccc} t_1 & t_2 & t_3 & 0 & 1 & \infty \\ \hline \alpha & \beta & \gamma & \alpha & \beta & \gamma \\ \alpha' & \beta' & \gamma' & \alpha' & \beta' & \gamma' \end{array} \mapsto$$

We next apply the exponent shifting lemma (Lemma 2.5.7.2). Making the gauge transformation $y = t^{-\alpha}(t-1)^{-\beta}\tilde{y}$ to transform the exponent table again to

$$\begin{array}{ccc|ccc} 0 & 1 & \infty & 0 & 1 & \infty \\ \hline \alpha & \beta & \gamma & 0 & 0 & \gamma + \alpha + \beta \\ \alpha' & \beta' & \gamma' & \alpha' - \alpha & \beta' - \beta & \gamma' + \alpha + \beta \end{array} \mapsto$$

We then just relabel the exponents:

$$\begin{aligned} a &= \gamma + \alpha + \beta \\ b &= \gamma' + \alpha + \beta \\ 1 - c &= \alpha' - \alpha \end{aligned}$$

and finally Fuchs' relation (Theorem 2.5.6.1) forces $\beta' - \beta = c - a - b$. Since the exponents determine the equation (Theorem 2.5.8.1) the transformed equation must be a hypergeometric equation with the given exponents. \square

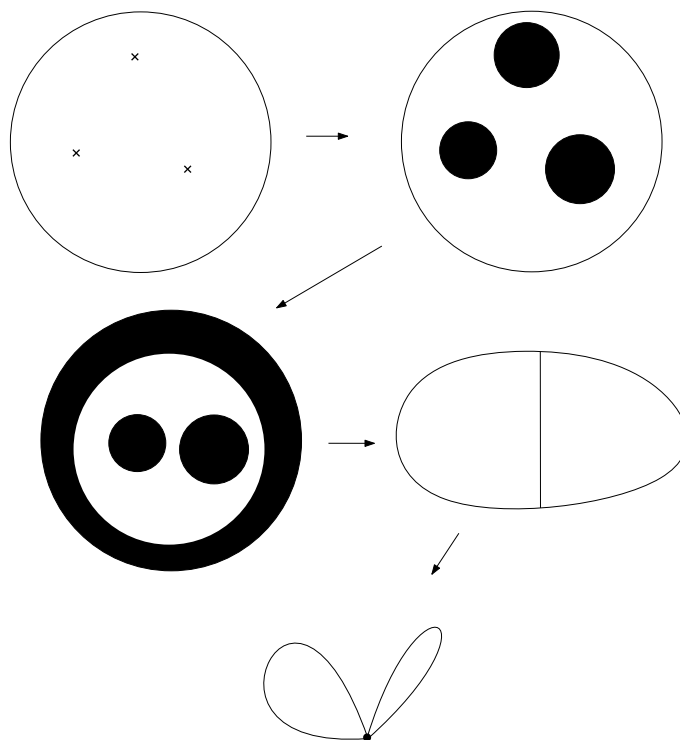


Figure 2.4: The figure shows $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ being deformed into $S^1 \vee S^1$. The first step is the increase the size of the holes to make it look like a bowling ball. We then wrap one of the holes completely around to get a disc with two interior discs removed. This is then seen to be equivalent to a circle with a line through it. After contracting the middle line one gets the bouquet of circles.

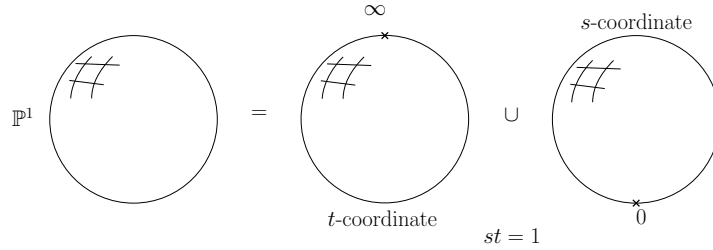


Figure 2.5: The projective line \mathbb{P}^1 is isomorphic to the Riemann sphere S^2 and is composed of two coordinate charts. The first chart we think of as the “usual” copy of \mathbb{C} (which algebraic geometers upgrade to the affine line \mathbb{A}^1) which has coordinate t . Then when we want to set $t = \infty$ we use another copy of \mathbb{C} with coordinate s where $s = 1/t$. The point $s = 0$ corresponds to the points $t = \infty$ and we use this s coordinate to do all of our computations at infinity.

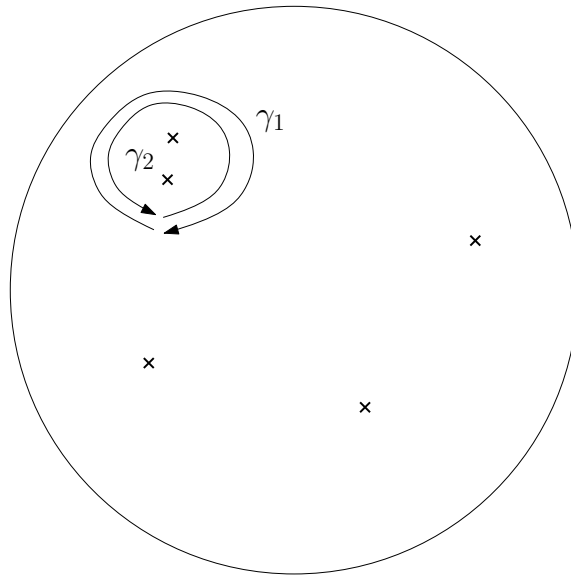


Figure 2.6: One uses the basic residue theorem on two simple integrals which are opposite of each other to prove the residue theorem on \mathbb{P}^1 .

Chapter 3

Connections and Riemann-Hilbert Correspondences

I started writing this section to present a proof of Plemelj's theorem, which at the time (and the 70 years that follows) was thought to be a solution to Hilbert's 21st problem in a modern form. To do this we need to introduce vector bundles with connections which then will later be used again for constructing isomonodromic deformations.

The basic strategy of Plemelj's proof is to glue a bunch of local equations together. We now understand this technique to be the part of what is called "descent theory". This is just a fancy word for "the theory of gluing things together". In order to glue things we need to define the things that we are gluing. These "things" are connections on vector bundles (E, ∇) over a Riemann surface (or complex manifold). The word "connections" is just a fancy word for "locally a D -module".

The idea of a connection leads to a bunch of interesting mathematics and physics including the notion of curvature. The basic motto is

$$\text{Force} = \text{Curvature}$$

Physically, all of the basic forces in the standard model of physics the weak, strong, electromagnetic are all Yang-Mills theories which involve this concept.

We will follow the chapters of Haeflinger and Malgrange the book on D -modules [BGK⁺87, Chapters III, IV] which largely follows [Del70]. There some nice material from the *Holomorphic Foliations and Algebraic Geometry* Summer School in Mathematics in 2019 which has excellent YouTube videos and notes. Viktoria Heu's Notes are brief and excellent. Also see Frank Loray's second lecture from the same Summer School here.¹

3.1 Vector Bundles and Connections

In this section we define vector bundles and connections. They are mainly a global language for linear differential equations and give a formalism in which we can talk about complicated changes of coordinates. The vector bundle encodes all possible changes of coordinates of the D -module and the connection is the derivation part of the D -module. More precisely it is the equation. Most importantly this global language allows us to glue together local information in order to solve (or show we can't solve) Hilbert's 21st problem.

Next given a vector bundle and a connection we get to talk about curvature and things like geodesics. A connection is a way to convert a derivative into a D -module structure. If we think of derivatives as vector fields on space, as one does with tangent bundles, connections are telling us how to associate a direction on our manifold to a direction in our vector bundle. In particular for each direction one gets a differential equation and solving this differential equations tells us how to move things around in the vector bundle. This is just solving an initial value problem in ODEs. One issue is that while moving around in our base space commutes, it doesn't necessarily translate to an commutative procedure for moving around in the vector bundle. This leads to a notion of curvature.

Consider Figure 3.1. In this picture we have our base space being a sphere S^2

¹There are actually many great videos of Frank Loray on YouTube if you do a quick search.

and the vector bundle being the tangent bundle itself – so note in particular that in this setup the tangent bundle is appearing twice: first as an object parametrizing directions and second as the object on which the directions act through covariant derivatives/assigning linear differential equations. One can see that if we move a vector from the north pole down a longitude, then along a latitude, then back up to the north pole around a longitude that we arrive with a vector which is displaced from the original one. This is what curvature is.

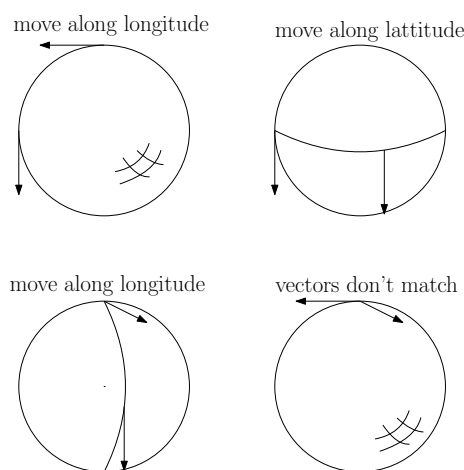


Figure 3.1: Curvature of the tangent bundle of S^2 associated to the Levi-Civita connection is evidenced by transporting tangent vectors around the sphere.

Finally, while curvature is an interesting concept in itself, the conditions for curvature give interesting differential equations. In fact all the gauge theories in particle physics (strong, weak, electro-magnetic), all of the fundamental forces are encoded by curvatures. Later, we are going to need curvature to vanish in order for an overdetermined collection of partial differential equations to be well-defined. This is needed for example in order to derive the Schlesinger equations for isomonodromic flows (this is an equation for equations!). These conditions are what give rise to the P_{VI} , Painlevé six. This is essentially a condition on equality of mixed partial derivatives for solutions of differential equations.

3.1.1 Vector Bundles (for the uninitiated)

There are two objects that we will often conflate: vector bundles and locally free sheaves. Locally free sheaves are essentially modules that we associate to open sets. They are modules over rings of holomorphic functions and they “glue” together nicely. Vector bundles are spaces over another space which have the property that local sections of the vector bundle form a locally free sheaf. Later we will conflate the two since for every vector bundle there is a locally free sheaf and conversely, to every locally free sheaf we can construct a vector bundle. If you know what these words mean, this section probably isn’t for you. I’m going to begin an introduction into these two fundamental objects with a discussion of coordinates of free modules. I will then extend this idea to describe the data of vector bundles given from two open sets. I will then describe the general definition and say a little bit about what it means to be a sheaf. After this section I’m going to assume everyone is familiar with these objects since a detailed discussion will lead us too far afield. We recommend [Vak17] for a more detailed discussion of vector bundles on schemes.

Let R be a commutative ring and let E be a free R -module of rank n . We have talked about coordinate isomorphisms $\psi : E \rightarrow R^{\oplus n}$ given by $v = f_1v_1 + \cdots + f_nv_n \mapsto (f_1, \dots, f_n)$ where v_1, \dots, v_n is a basis for E and $f_1, \dots, f_n \in R$ are called the coordinates. The map ψ is called a trivialization. We have also talked about how a change in choice of basis for E transforms the coordinates by some element of $\mathrm{GL}_n(R)$. The element of $\mathrm{GL}_n(R)$ transitions from one set of coordinates in one basis to another set of coordinates in another basis.

A vector bundle is sort of like this naive change of coordinates but we have a varying collection of R_i and E_i for i in some index set I and they need to satisfy compatibility conditions. The R_i are the functions on some open set of some space and the E_i are local sections of the vector bundle.

Example 3.1.1.1 (Vector Bundles With Two Charts). The data for this is some E_i R_i -modules for $i = 1, 2$ and an additional R_{12} -module E_{12} .

There are ring homomorphisms

$$R_1 \longrightarrow R_{12} \longleftarrow R_2$$

and morphisms of abelian groups

$$E_1 \longrightarrow E_{12} \longleftarrow E_2$$

which respects the module actions. Moreover these have the property that a basis for E_1 or E_2 induce a basis for E_{12} and hence trivializations for E_1 or E_2 induce trivializations for E_{12} . There are additional glueing properties, but I will state those when I state the official definition.

Example 3.1.1.2 (Vector Bundles on \mathbb{P}^1). In the case of \mathbb{P}^1 we have two open sets U_0 and U_∞ which cover $\mathbb{P}^1 = U_0 \cup U_\infty$.

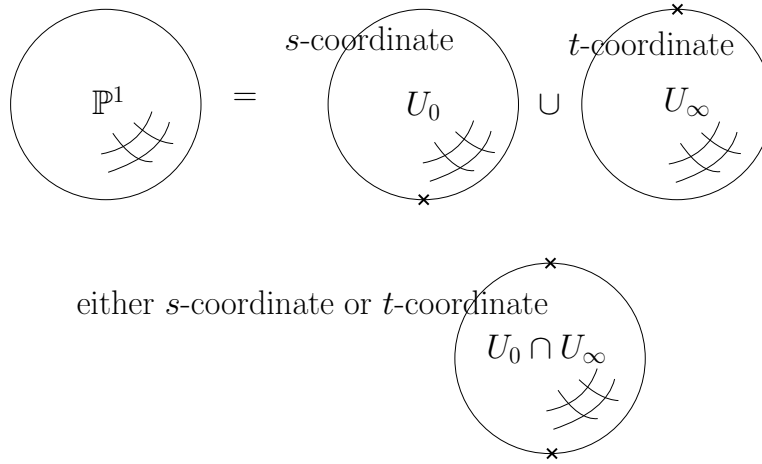


Figure 3.2: \mathbb{P}^1 is covered by U_0 and U_∞ . Here U_0 uses the standard t -coordinate and U_∞ uses the coordinate at infinity s given by $s = 1/t$. On their intersection $U_0 \cap U_\infty$ you can use either coordinate.

In the previous example (with indexing $0, \infty$ instead of $1, 2$) we have

$$R_0 = \text{Hol}(U_0) \longrightarrow R_{0\infty} = \text{Hol}(U_0 \cap U_\infty) \longleftarrow R_\infty = \text{Hol}(U_\infty)$$

where the maps are “restriction of the domain”. Then to specify a vector bundle one can specify three modules E_0 , E_∞ , and $E_{0\infty}$ which are free R_0 , R_∞ , and $R_{0\infty}$ -modules respectively.

While the basic data above is correct it is terrible to define vector bundles this way. We want a definition that doesn't depend on the choice of cover, works for all spaces, and allows us to glue local objects together. The technical thing we want to say is that a vector bundle E on X is a *locally free sheaf of \mathcal{O}_X -modules*. There are two categories in which we want to formulate this notion: the category of schemes and the category of complex manifolds.

- (\mathcal{O}_X -modules) In the category of complex manifolds for an open set U one has $\mathcal{O}_X(U) = \text{Hol}(U)$ the set of holomorphic functions on U and in the category of schemes $\mathcal{O}_X(U)$ is the structure sheaf.
- (Sheaves) For E to be a sheaf of \mathcal{O}_X -modules we need that $E(U)$ to be a $\mathcal{O}_X(U)$ -module for every $U \subset X$ and it needs to satisfy sheaf axioms. Elements of $E(U)$ are called sections over U . The first axiom says that if you have bunch of open sets and sections on those open sets that agree on the intersections then there exists a section over the union of the open sets that restricts to each of those sections. The second axiom says that such a lifting is unique: if you have two sections which agree on a collection of open sets that cover the set it is a section over then the two sections must be the same.
- (Locally free of rank n) Finally, for E to be locally free of rank n that means that for every $x \in X$ there exists some U open containing x and an isomorphism $\psi_U : E(U) \rightarrow \mathcal{O}_X(U)^{\oplus n}$.

The nice thing about vector bundles is that they satisfy effective descent. This is sort of like the sheaf axiom but for objects of the category themselves. If E_i are vector bundles over U_i and they $E_i|_{U_i \cap U_j} \cong E_j|_{U_i \cap U_j}$ and these isomorphisms satisfy some compatibility conditions, then there exists a vector bundle over $\bigcup_i U_i$. The correct way of talking about this now is to say that fibered category of vector bundles over the category of spaces you are considering is a stack. We aren't going to review this here, this would take an entire class. The take-away is that you can build up vector bundles from local data.

Let's do a simple example of a line bundle on \mathbb{P}^1 . A line bundle is just a vector bundle of rank one.

Example 3.1.1.3. Lets consider the sheaf of holomorphic differentials on \mathbb{P}^1 . This is the sheaf we denote by $E = \Omega_{\mathbb{P}^1}^1$. In our usual coordinates we have

$$E(U_0) = \mathcal{O}_{\mathbb{P}^1}(U_0)dt, \quad E(U_\infty) = \mathcal{O}_{\mathbb{P}^1}(U_\infty)ds.$$

There are trivializations

$$\psi_0: E(U_0) \rightarrow \mathcal{O}_{\mathbb{P}^1}(U_0), \quad f(t)dt \mapsto f(t)$$

$$\psi_\infty: E(U_\infty) \rightarrow \mathcal{O}_{\mathbb{P}^1}(U_\infty), \quad g(s)ds \mapsto g(s)$$

Both of these trivializations are valid on $E(U_0 \cap U_\infty)$ and over $U_0 \cap U_\infty$ we have

$$\psi_\infty \psi_0^{-1}(1) = \psi_\infty(dt) = \psi_\infty\left(\frac{-ds}{s^2}\right) = \frac{-1}{s^2} \psi_\infty(ds) = \frac{-1}{s^2}.$$

In the above example we see that the transition map was given by multiplication by $-1/s^2$. It turns out that every line bundle on \mathbb{P}^1 admits trivializations over U_0 and U_∞ with transition maps of the form $f \mapsto -s^{-d}f$ for some integer d . The integer d characterizes the line bundle up to isomorphism and we call the one with integer d , $\mathcal{O}_{\mathbb{P}^1}(-d)$. So for example, $\Omega_{\mathbb{P}^1} \cong \mathcal{O}_{\mathbb{P}^1}(-2)$. There are a couple interpretation of these line bundles, one being sheaves of meromorphic functions with poles at more order d at infinity. The notation is a little wonky too. If U is an open subset of \mathbb{P}^1 then to take sections of this vector bundle we write $\mathcal{O}_{\mathbb{P}^1}(d)(U)$ so the d has nothing to do with the open sets we were plugging in earlier.

Remark 3.1.1.4. Readers already familiar with algebraic or complex geometry will recognize $-s^{-d} \in \mathcal{O}_{\mathbb{P}^1}^\times(U_0 \cap U_\infty)$ as a representative of the cohomology class in $H^1(\mathbb{P}^1, \mathcal{O}^\times) = \text{Pic}(\mathbb{P}^1)$ given in terms of a Čech cocycle with two open sets.

The comment about line bundles extends to vector bundles. The obvious vector bundles we can think of are direct sums of line bundles. These turn out to be all of them.

Theorem 3.1.1.5 (Birkoff-Grothendieck). *All vector bundles on \mathbb{P}^1 are isomorphic to*

$$\mathcal{O}_{\mathbb{P}^1}(d_1) \oplus \cdots \mathcal{O}_{\mathbb{P}^1}(d_n)$$

for some $d_1, d_2, \dots, d_n \in \mathbb{Z}$.

3.1.2 Babymost case: Covariant Derivatives and Connections Associated to ODEs

Since connections can get abstract, before proceedings with the full definition, I'm going to explain what everything is for an ODE. The main idea is that we can convert “ Y is a solution of a differential equations’ into “ Y is horizontal for a connection”.

$$\frac{dY}{dt} = B(t)Y \iff \nabla_{\frac{\partial}{\partial t}}(Y) = 0.$$

The covariant derivative in this example is the \mathbb{C} -linear operator

$$\nabla_{\frac{\partial}{\partial t}} = \frac{\partial}{\partial t} - B(t).$$

If, say, $B(t)$ is a holomorphic on U some neighborhood of a points in \mathbb{P}^1 then this defined an operator $R^n \rightarrow R^n$ whree $R = \text{Hol}(U)$. The connection in this situation is a map

$$\nabla = d - B(t)dt$$

where d is the exterior differential acting on column vectors in R^n and $-B(t)dt$ is a matrix of differential 1-forms. This defined a map $\nabla : R^n \rightarrow \Omega_{\mathbb{P}^1}^1(U) \otimes R^n$. Explicitly

$$\nabla \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} dy_1 \\ dy_2 \\ \vdots \\ dy_n \end{pmatrix} - B(t) \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} dt.$$

Sometimes the matrix $-B(t)dt$ is called the connection 1-form and usually denoted by ω or $A(t)dt$ (so that $A(t) = -B(t)$). Before proceeding to the abstract theory we remark that ∇ and $\nabla_{\frac{\partial}{\partial t}}$ are related by the pairing

$$\nabla_{\frac{\partial}{\partial t}}(Y) = \langle \nabla(Y), \frac{\partial}{\partial t} \rangle.$$

Usually this pairing is just defined differential forms and derivations, i.e. between $\Omega_{\mathbb{P}^1}^1(U)$ and $T_{\mathbb{P}^1}(U)$ but if W is a vector of 1-forms given by we extend the pairing to W by pairing with each entry of W .

3.1.3 Connections

There exists definitions of connections for schemes but for concreteness we will work with complex manifolds.² Let X be a complex manifold and let E be a vector bundle on X .

Definition 3.1.3.1. A *connection* on E is a \mathbb{C} -linear map

$$\nabla : E \rightarrow \Omega_X^1 \otimes_{\mathcal{O}_X} E$$

satisfying

1. For all $f \in \mathcal{O}_X$ and all $s \in E$, $\nabla(fs) = df \otimes s + f\nabla(s)$.
2. For all $s_1, s_2 \in E$, we have $\nabla(s_1 + s_2) = \nabla(s_1) + \nabla(s_2)$.

We now can extract a more general definition of local system.

Definition 3.1.3.2. The space of *horizontal sections* of (E, ∇) is defined by

$$U \mapsto E^\nabla(U) = \{s \in E(U) : \nabla(s) = 0\}.$$

It is an exercise in familiarity with the definitions of vector bundles to check that this is in fact a local system.

Exercise 3.1.3.3. Check that E^∇ is a local system.

We remark that every connection ∇ can be extended to $\Omega_X^i \otimes E$. This means that given $\omega \otimes s \in \Omega_X^i \otimes E$ we define

$$\nabla(\omega \otimes s) = d(\omega) \otimes s + (-1)^i \omega \wedge \nabla(s).$$

²I'm really working with ringed spaces here.

3.1.4 Christoffel Symbols

Let X be a complex manifold of dimension n and let E be a vector bundle of rank n on X . Locally (on some open set $U \subset X$) we can fix coordinates t_1, \dots, t_m of X and a basis s_1, \dots, s_n of E . This means that on U we have

$$\begin{aligned} E(U) &= \mathcal{O}_X(U)s_1 + \dots + \mathcal{O}_X(U)s_n, \\ T_X(U) &= \mathcal{O}_X(U)\frac{\partial}{\partial t_1} + \dots + \mathcal{O}_X(U)\frac{\partial}{\partial t_m}, \\ \Omega_X(U) &= \mathcal{O}_X(U)dt_1 + \dots + \mathcal{O}_X(U)dt_m. \end{aligned}$$

In each of the above expression the sum are direct – so each $\mathcal{O}_X(U)$ -module is free with the given basis. Now to get the Christoffel symbols we can just write down what $\nabla(s_j) \in \Omega_X^1(U) \otimes E(U)$ must look like. It must have the form

$$\nabla(s_j) = \sum_{i=1}^n \sum_{\alpha=1}^m \Gamma_{j\alpha}^i dt_\alpha \otimes s_i. \quad (3.1.1)$$

The Christoffel symbols are just the structure “constants” for the connection.

Definition 3.1.4.1. The elements $\Gamma_{j\alpha}^i \in \mathcal{O}_X(U)$ are called the *Christoffel symbols* of ∇ with respect to the local basis s_1, \dots, s_n of E and local coordinates t_1, \dots, t_m of X .

There is a convenient way to write this down using Einstein notation. If we write t^α instead of t_α then we can write (3.1.1) in the simple form

$$\nabla(s_j) = \Gamma_{j\alpha}^i dt^\alpha \otimes s_i.$$

In Einstein summation notation a upper index followed by a repeated lower index implies summation over that variable. So, in this expression, there is an implied sum over α and over i . We will make use of Einstein notation freely.

3.1.5 Covariant Derivatives

We now give the definition of a covariant derivative.

Definition 3.1.5.1. Let $\theta \in T_X(U)$ be a derivation. We define the *covariant derivative associated to θ* to be the operator

$$\nabla_\theta: E(U) \rightarrow E(U), \quad s \mapsto \nabla_\theta(s) = \langle \nabla(s), \theta \rangle.$$

In the above expression the pairing $\Omega_X^1 \times T_X \rightarrow \mathcal{O}_X$ (which we can write as $\Omega_X^1 \otimes T_X \rightarrow \mathcal{O}_X$) is extended to $(E \otimes \Omega_X^1) \otimes T_X \rightarrow E \otimes \mathcal{O}_X = E$ by tensoring up to E . All tensors here are over \mathcal{O}_X . Note that this pairing just means that we pair each component of $E \otimes \Omega_X^1$ with a tangent vector and take the element of E obtained from the result. Also, since ∇ satisfies a product rule and sum rule we will have

$$\nabla_\theta(fs) = \theta(f)s + f\nabla_\theta(s), \quad \forall f \in \mathcal{O}_x, \forall s \in E,$$

$$\nabla_\theta(s_1 + s_2) = \nabla_\theta(s_1) + \nabla_\theta(s_2), \quad \forall s_1, s_2 \in E.$$

In local coordinates, we have an explicit expression for covariant derivatives using Christoffel symbols. Before proceeding I will make some remarks on notation. First, derivatives $\frac{\partial}{\partial t^\beta}$ are “naturally lowered” in so Einstein notation and for convenience we often write them as $\partial_\beta = \frac{\partial}{\partial t^\beta}$. In this notation, with these local coordinates, a general derivative θ is written as $\theta = a^\beta \partial_\beta$. Also, it is annoying to write $\nabla_{\frac{\partial}{\partial t^\beta}}$ all of the time. We will write

$$\nabla_\beta = \nabla_{\partial_\beta} = \nabla_{\frac{\partial}{\partial t^\beta}}$$

to simplify the notation.

Example 3.1.5.2. Let t^1, \dots, t^m be local coordinates for X and let s_1, \dots, s_n be a local basis for E . Then if $\theta = \frac{\partial}{\partial t^\beta}$ we have

$$\nabla_{\frac{\partial}{\partial t^\beta}}(s_j) = \langle \Gamma^i_{j\alpha} dt^\alpha \otimes s_i, \frac{\partial}{\partial t^\beta} \rangle = \Gamma^i_{j\alpha} \delta^\alpha_\beta s_i = \Gamma^i_{j\beta} s_i.$$

In particular if $s = f^j s_j \in E$ then

$$\begin{aligned} \nabla_\beta(f^j s_j) &= \partial_\beta(f^j) s_j + f^j \nabla_\beta(s_j) \\ &= \partial^\beta(f^j) s_j + f^j \Gamma^i_{j\beta} s_i \\ &= (\partial^\beta(f^j) + f^i \Gamma^j_{i\beta}) s_j \end{aligned}$$

A general θ we can write as $\theta = a^\beta \partial_\beta$ and

$$\nabla_\theta(s_j) = \nabla_{a^\beta \partial_\beta}(s_j) = a^\beta \nabla_\beta(s_j) = a^\beta \Gamma^i_{j\beta} s_i.$$

3.1.6 Connection 1-forms

The purpose of this section is to show that after fixing a local basis s_1, \dots, s_n for $E(U)$, that locally

$$\nabla = d + \omega, \quad \text{on } \mathcal{O}_X(U)^{\oplus n}$$

where $\omega \in M_n(\Omega_X^1(U))$ is a matrix of 1-forms. Before proceeding, I need to explain how to multiply matrices of one forms and how the exterior derivative d works.

Calculus of Connection 1-forms

For the purpose of iterating the connection later, we remark that the exterior algebra $\Omega_X^\bullet = \bigoplus_{d=0}^m \Omega_X^\bullet$ is a sheaf of skew commutative ring satisfying

$$\eta_1 \wedge \eta_2 = (-1)^{d_2} \eta_2 \wedge \eta_1, \quad \eta_j \in \Omega_X^{d_j}.$$

Matrices $\omega, \eta \in M_2(\Omega^\bullet)$ are then multiplied by using the formula

$$\omega \wedge \eta = \left(\sum_{l=1}^n \omega_{il} \wedge \eta_{lj} \right),$$

if $\omega = (\omega_{ij})$ and $\eta = (\eta_{ij})$.

One more notational remark before proceeding: If $s_i \in E(U)$ form a local basis we will let s^i denote the dual basis in $E^\vee(U)$ where E^\vee is the dual vector bundle. It is defined by

$$E^\vee(V) := E(V)^\vee = \text{Hom}_{\mathcal{O}_X(V)}(E(V), \mathcal{O}_X(V))$$

for V an open subset of X . The second \vee is just usual $\mathcal{O}_X(V)$ -module duality as defined by the last equality.

Exercise 3.1.6.1. Suppose that t^α are local coordinates of X and that s_i is a local basis for E . We can write ∇ as an element of $\Omega_X^1 \otimes \text{End}(E)$ by

$$\nabla = \Gamma_{j\alpha}^i dt^\alpha \otimes s_i \otimes s^j.$$

Here $\text{End}(E)$ is the sheaf of \mathcal{O}_X -linear maps from E to itself (endomorphisms) and $\text{End}(E) = E \otimes E^\vee$.

We now explain how to work with $\nabla = d + \omega$ as an element of a Weyl algebra. The thing to keep in mind here when working with $\nabla = d + \omega$ is that ω is a matrix over a non-commutative ring i.e. $\omega \in M_n(\Omega_X^\bullet)$ (so it is like super noncommutative) and d is an operator on this ring. This means we need to do computations in a very weird looking Weyl algebra $M_n(\Omega_X^\bullet)[d]$ where d here is the exterior derivative. If $\omega \in M_n(\Omega_X^\bullet)$ we need to understand how $d\omega$ acts on $\eta \in M_n(\Omega_X^\bullet)$ when ω is homogenous. The key thing to keep in mind is that ω act by wedge-matrix-multiplication:

$$(d\omega)(\eta) = d(\omega \wedge \eta) = d(\omega) \wedge \eta + (-1)^{\deg(\omega)} \omega \wedge d(\eta).$$

This proves the following:

Lemma 3.1.6.2 (Basic Weyl Algebra Rules). *In the Weyl algebra $M_n(\Omega_X^\bullet)[d]$ for $\omega \in M_n(\Omega_X^\bullet)$ homogeneous we have*

$$d \wedge \omega = (-1)^{\deg(\omega)} \omega \wedge d + d(\omega). \quad (3.1.2)$$

This is going to be used when computing our formulas for curvature (Theorem 3.1.7.6).

The formula: $\nabla = d + \omega$

We now verify the claim about the description in local coordinates. Let U be an open subset for which s_1, \dots, s_n is a basis for $E(U)$. Let $\psi : E(U) \rightarrow \mathcal{O}_X(U)^{\oplus n}$ be the trivialization given by $\psi(f^i s_i) = f^i e_i$ where e_i is the standard basis vector on $\mathcal{O}_X^{\oplus n}$ (say viewed as column vectors). The trivialization extends to $E(U) \otimes \Omega_X(U) \rightarrow \mathcal{O}_X(U) \otimes \Omega_X^1(U)$ by functoriality of the tensor

product and we will abusively also denote this isomorphism by ψ . We now have a square,

$$\begin{array}{ccc}
 s = f^i s_i & \xrightarrow{\quad} & \nabla(s) \\
 \downarrow & \in & \cap \\
 \begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix} & \in & \begin{array}{ccc} E(U) & \xrightarrow{\nabla} & \Omega_X(U) \otimes E(U) \\ \downarrow \psi & & \downarrow \psi \\ \mathcal{O}_X(U)^{\oplus n} & \xrightarrow{\nabla^\psi} & \Omega_X(U) \otimes \mathcal{O}_X(U)^{\oplus n} \end{array}
 \end{array}$$

where ∇^ψ denotes the connection in trivialized coordinates. Examining the diagram we see that we have

$$\psi(\nabla s) = \nabla^\psi \begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix},$$

so it remains to compute what $\psi(\nabla(s))$ is. We get

$$\begin{aligned}
 \nabla(s) &= df^i \otimes s_i + f^i \Gamma_{i\alpha}^j dt^\alpha \otimes s_j \\
 \mapsto \psi(\nabla(s)) &= df^i \otimes e_i + f^i \Gamma_{i\alpha}^j dt^\alpha \otimes e_j = \begin{pmatrix} df^1 \\ \vdots \\ df^n \end{pmatrix} + \omega \begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix}
 \end{aligned}$$

where we have written

$$f^i \Gamma_{i\alpha}^j dt^\alpha \otimes e_j = \omega \begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix}, \quad \omega = A_\alpha dt^\alpha, \quad A_\alpha = (\Gamma_{i\alpha}^j).$$

We summarize the above discussion with the following lemma.

Lemma 3.1.6.3. *If E is a rank n vector bundle on an m -dimensional complex manifold and $U \subset X$ is an open subset such that X has local coordinates*

t^1, \dots, t^m and local basis s_1, \dots, s_n then in local coordinates $\nabla_\alpha : \mathcal{O}_X(U)^{\oplus n} \rightarrow \mathcal{O}_X(U)^{\oplus n}$ takes the form

$$\nabla_\alpha = \partial_\alpha + A_\alpha,$$

where $A_\alpha = (\Gamma_{i\alpha}^j)$ and elements of $\mathcal{O}_X(U)^{\oplus n}$ are viewed as column vectors. In Einstein notation

$$\nabla_\alpha f^i = \partial_\alpha(f^i) + \Gamma_{j\alpha}^i f^j.$$

The matrix of 1-forms ω is called the connection 1-form and we record this in a definition environment for those browsing looking for the definition.

Definition 3.1.6.4. The matrix $\omega = A_\alpha dt^\alpha \in \Omega_X(U) \otimes \mathcal{O}_X(U)^{\oplus n}$ is called a *connection 1-form*.

3.1.7 Curvature

We keep our notation as in the previous section. We will let E be a rank n vector bundle on an m -dimensional complex manifold and let U be an open subset of X which admits local coordinate t^α for X and a local basis s_j for E .

Recall that ∇ tells us how to move on E given movement on the base: to flow from $s_* \in E(U)$ to another point s sufficiently near by along the direction of ∂_α . See figure 3.1.7 for a picture of paths in an open set $U \subset X$ being lifted to paths in the vector bundle by solving differential equations.

We describe these equations. If $s_* = f^j s_j$ then to we just solve

$$\nabla_\alpha f^j = 0, \quad j = 1, \dots, n \quad (3.1.3)$$

which is just a linear differential equation with only derivatives in t^α appearing. Note that this is just an equation of the form

$$\frac{\partial Y}{\partial t^\alpha} = -A_\alpha(t)Y,$$

which we are very familiar with by now. The two expressions are related by letting $Y = (f^1, \dots, f^n)$ and viewing it as a column vector.

There are some natural questions that come up when thinking about this transport of motion on the base to motion on the fiber.

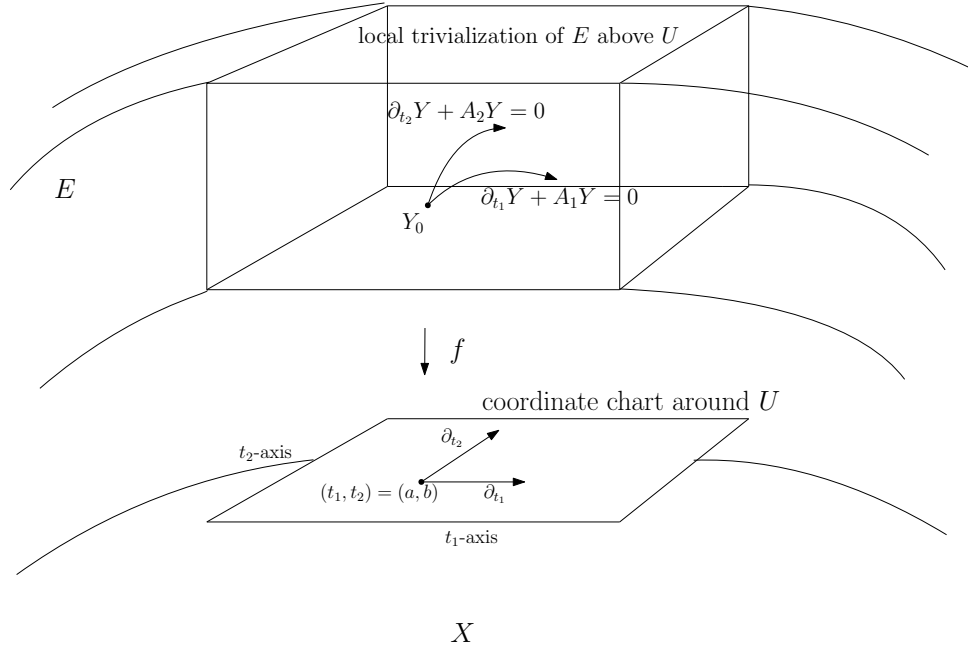


Figure 3.3: Movement in the base tells us how to move in the fibers by solving differential equations. The above figure shows paths in the space being transported to the vector bundle locally by solving differential equations.

Problem 3.1.7.1 (Integrability Problem). Is it possible to solve all of our equations (3.1.3) at once (i.e. simultaneously for $\alpha = 1, \dots, m$)?

If this is possible we call the connection *integrable*.

Problem 3.1.7.2 (Curvature Problem). Is moving the direction of t^α then moving in the direction of t^β the same as moving in the direction of t^β then moving in the direction of t^α ?

If this is the case we call the connection *flat*.

It turns out that flatness and integrability are really the same thing and that this condition is given $\nabla^2 = 0$ which is equivalent to vanishing the the curvature tensor. We will now explain.

The integrability issue is the is saying that a solution of $\nabla_\alpha Y = 0$ may not also be a solution of $\nabla_\beta Y = 0$. The first equation is an equation for $\partial_\alpha Y$ and the second equation is an equation for $\partial_\beta Y$. It is not guaranteed that $\partial_\alpha \partial_\beta Y = \partial_\beta \partial_\alpha Y$. When we can do this, the equation is integrable.

As lifting directions is really about general derivations, we formulate this commutation problem more generally. First recall that if $\theta_1, \theta_2 \in T_X$ then the commutator

$$[\theta_1, \theta_2] = \theta_1 \theta_2 - \theta_2 \theta_1,$$

is also a derivation. This is part of the fact that T_X is a sheaf of Lie algebras. The infinitesimal version of the curvature problem (taking the limit over small paths) leads to a “curvature zero” condition.

Problem 3.1.7.3. When does E have a well-defined structure of a T_X -module via ∇ . In other words, when is it the case that for all $\theta_1, \theta_2 \in T_X$ the following equation holds:

$$[\nabla_{\theta_1}, \nabla_{\theta_2}] = \nabla_{[\theta_1, \theta_2]}? \quad (3.1.4)$$

The failure of the commutation relation (3.1.4) is measured by the curvature tensor which we now define.

Definition 3.1.7.4. The map $R_\nabla : T_X \otimes T_X \rightarrow \text{End}(E)$ given by

$$R_\nabla(\theta_1, \theta_2)(s) = \nabla_{\theta_1}(\nabla_{\theta_2}(s)) - \nabla_{\theta_2}(\nabla_{\theta_1}(s)) - \nabla_{[\theta_1, \theta_2]}(s)$$

is called the *curvature tensor*.

When the context is clear we will just use $R = R_\nabla$ so that we don't have to keep writing the subscript ∇ . In what follows we will soon see that R as an alternative description in terms of ∇^2 . To see that this even makes sense note that $\nabla^2 : E \rightarrow \Omega_X^2 \otimes E$. Hence ∇^2 can eat two tangent vectors θ_1 and θ_2 and a section s and spit out another section. The subsequent theorem will prove that for all $s \in E$ and $\theta_1, \theta_2 \in T_X$ we will have

$$R(\theta_1, \theta_2)(s) = \langle \nabla^2(s), \theta_1 \wedge \theta_2 \rangle,$$

where the pairing $\langle -, - \rangle$ is the one between Ω_X^2 and $T_X \wedge T_X$.

Exercise 3.1.7.5. In local coordinates t^α and s_i find an expression for $R_\nabla(\partial_\alpha, \partial_\beta)(s_j)$ in terms of the Christoffel symbols.

We now can give our equations for integrability/flatness.

Theorem 3.1.7.6 (Integrability Conditions). *Let E be a vector bundle of rank n on a complex manifold X . Let ∇ be a connection on E . The following are equivalent:*

1. *Iterating ∇ twice is zero: $\nabla^2 = 0$.*
2. *The curvature tensor is identically zero: $R_\nabla = 0$.*
3. *For every set of local coordinate so that $\nabla = d + \omega$ locally with ω a connection one form with respect to these coordinates we have*

$$d(\omega) + \omega \wedge \omega = 0.$$

4. *For every set of local coordinate so that $\nabla = d + \omega$ with $\omega = A_\alpha dt^\alpha$ and $A_\alpha \in M_n(\mathcal{O}_X(U))$ one has*

$$\frac{\partial A_\alpha}{\partial t^\beta} - \frac{\partial A_\beta}{\partial t^\alpha} = -[A_\alpha, A_\beta].$$

More generally

$$d(\omega) + \omega \wedge \omega = G_{\alpha\beta} dt^\alpha \wedge dt^\beta, \quad G_{\alpha\beta} = \frac{1}{2} (\partial_\alpha A_\beta - \partial_\beta A_\alpha + [A_\alpha, A_\beta])$$

is a local expressions of the curvature 2-form.

Proof. We will first prove (2) if and only if (4). Let $U \subset X$ be an open set with coordinates t^1, \dots, t^m and where $E(U)$ has a basis s_1, \dots, s_n . Since every expression of $R(\theta_1, \theta_2)$ can be expressed in terms of $R(\partial_\alpha, \partial_\beta)$, we just need to compute $R(\partial_\alpha, \partial_\beta)$. The following computation is a Weyl algebra computation (meaning everything is viewed as an operator). Importantly, for a matrix A and a derivative θ we have $\theta A = A\theta + \theta(A)$ there $\theta(A)$ denotes

the application of θ to the entries of A and θA means the application of A and an operator followed by θ as an operator.

$$\begin{aligned}
R(\partial_\alpha, \partial_\beta) &= \nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha - \nabla_{[\partial_\alpha, \partial_\beta]} \\
&= \nabla_\alpha \nabla_\beta - \nabla_\beta \nabla_\alpha \\
&= (\partial_\alpha + A_\alpha)(\partial_\beta + A_\beta) - (\partial_\beta + A_\beta)(\partial_\alpha + A_\alpha) \\
&= (\partial_\alpha A_\beta + A_\alpha \partial_\beta) + A_\alpha A_\beta + \partial_\alpha \partial_\beta - ((\partial_\beta A_\alpha + A_\beta \partial_\alpha) + A_\beta A_\alpha + \partial_\beta \partial_\alpha) \\
&= [A_\alpha, A_\beta] + A_\alpha \partial_\beta - A_\beta \partial_\alpha \\
&\quad + A_\beta \partial_\alpha + \partial_\alpha(A_\beta) \\
&\quad + A_\alpha \partial_\beta + \partial_\beta(A_\alpha) \\
&= \partial_\alpha(A_\beta) - \partial_\beta(A_\alpha) + [A_\alpha, A_\beta].
\end{aligned}$$

This implies (in coordinates)

$$R(\partial_\alpha, \partial_\beta) = \partial_\alpha(A_\beta) - \partial_\beta(A_\alpha) + [A_\alpha, A_\beta].$$

Let's see that (3) and (4) are equivalent. We recall that locally $\omega = A_\alpha dt^\alpha$. We will compute $d\omega + \omega \wedge \omega$.

In what follows we are going to use the following trick: if $H_{\alpha\beta}$ is an antisymmetric tensor, e.g. there is some free R -modules V with a basis v^α such that $H_{\alpha\beta}v^\alpha \wedge v^\beta \in V \wedge V$ then $H_{\alpha\beta} = -H_{\beta\alpha}$ and

$$H_{\alpha\beta} = \frac{1}{2}(H_{\alpha\beta} - H_{\beta\alpha}).$$

This will be used when we compute coefficients of differential forms “coordinate tensor” is not alternating. You can just antisymmetrize. If you don't understand what this means now, it should become apparent in the following computation.

Write $\omega = A_\alpha dt^\alpha$ hence in local coordinates. We compute:

$$\begin{aligned}
\omega \wedge \omega &= A_\alpha dt^\alpha \wedge A_\beta dt^\beta = A_\alpha A_\beta dt^\alpha \wedge dt^\beta = \frac{1}{2}(A_\alpha A_\beta - A_\beta A_\alpha) dt^\alpha \wedge dt^\beta \\
d(\omega) &= d(A_\beta dt^\beta) = \frac{\partial A_\beta}{\partial t^\alpha} dt^\alpha \wedge dt^\beta = \frac{1}{2} \left(\frac{\partial A_\beta}{\partial t^\alpha} - \frac{\partial A_\alpha}{\partial t^\beta} \right) dt^\alpha \wedge dt^\beta
\end{aligned}$$

which gives our desired identity.

We will now show the computation $\nabla^2 = 0$ in (1) equivalent to $d\omega + \omega \wedge \omega = 0$ in (3). The key idea is to use the relations for the Weyl algebra of the exterior algebra (Lemma 3.1.6.2). As elements of $R[d]$ where $R = M_n(\Omega_X^\bullet)$ we have

$$\begin{aligned}\nabla^2 &= (d + \omega)(d + \omega) = d^2 + \omega \wedge d + d\omega + \omega \wedge \omega \\ &= \omega \wedge d + (d(\omega) - \omega \wedge d) + \omega \wedge \omega \\ &= d(\omega) + \omega \wedge \omega.\end{aligned}$$

This proves the desired equality. \square

Exercise 3.1.7.7. Photons are the force carrying particles for the electromagnetic force field (changes in force are mediated by the emission of light). In this exercise the speed of light will be $c = 1$. Space time is encoded by a manifold X and the vector bundle is rank one corresponding to the Lie algebra of $U(1)$. In local coordinates (some chart of spacetime) we let $E = (E_1, E_2, E_3)$ denote a electric field and $B = (B_1, B_2, B_3)$ denote a magnetic field. Both E and B are functions of (t, x, y, z) . They are encoded in the Faraday tensor $F \in \Omega_X^2$ given by

$$F = E_1 dx \wedge dt + E_2 dy \wedge dt + E_3 dz \wedge dt + B_1 dy \wedge dz + B_2 dz \wedge dx + B_3 dx \wedge dy.$$

If we order the variables $(x^0, x^1, x^2, x^3) = (t, x, y, z)$ then Faraday tensor is $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is or as an antisymmetric matrix is

$$(F_{\mu\nu}) = \begin{bmatrix} 0 & E_1 & E_2 & E_3 \\ -E_1 & 0 & -B_3 & B_2 \\ -E_2 & B_3 & 0 & -B_1 \\ -E_3 & -B_2 & B_1 & 0 \end{bmatrix}.$$

Since we are in rank one everything commutes and there is nothing really interesting to say about the curvature equations. Maxwell's equations become $dF = 0$.

I'm tempted to put an exercise about Yang-Mills equations here but haven't done so. Perhaps you should just google these now and see all of these curvature tensors appearing everywhere.

3.2 Plemelj's Construction

3.3 Holomorphic Riemann-Hilbert Correspondence

Chapter 4

The Poincaré-Fuchs Theorem

This section is motivated by the following problem.

Problem 4.0.0.1. Which ordinary differential equations of the form

$$P(t, y, y') = 0$$

with $P(u, v, w) \in \mathbb{C}[x, y, z]$ admit meromorphic solutions $y(t)$ which have the property that $y(t)$ has no-movable singularities?

The Poincaré-Fuchs theorem states that the only equations of this form are 1) Ricatti equations and 2) Weierstrass equations. We are going to give two proofs of this. The first is foliation theoretic and the second is differential algebraic.

4.1 Differential Equations from the Foliation-Theoretic Perspective

4.1.1 Exterior Differential Systems

4.1.2 Comparison to Δ -Schemes

Chapter 5

Differential Galois Theory

This section is motivated by the following problem.

Problem 5.0.0.1. Can the function $\int e^{x^2} dx$ be written in terms of elementary functions?

In order to address this we need to talk about what an elementary function is. We also need to talk about what a special function is.

Chapter 6

The Poincaré Problem

This section is motivated by the following problem.

Problem 6.0.0.1. Consider differential equations of the form

$$\frac{dy}{dx} = \frac{a(x, y)}{b(x, y)}, \quad a(x, y), b(x, y) \in \mathbb{C}[x, y].$$

For which $a(x, y), b(x, y) \in \mathbb{C}[x, y]$ does the equation admit algebraic solutions?

Chapter 7

Dimension Theory

This section is motivated by the following problem.

Problem 7.0.0.1. Let (K, Δ) be a Δ -field. Let $u_1, \dots, u_n \in K\{x_1, \dots, x_n\}_\Delta$ and consider the system of PDEs

$$u_1 = u_2 = \dots = u_n = 0.$$

How many constants of integration does a general solution of this equation need?

Appendix A

Analytcs Elements of Differential Equations

A.1 Solutions of Homogeneous Differential Equations

♠♠♠ Taylor: [This section needs to be finished]

Theorem A.1.0.1 (Solutions of Homogeneous Differential Equations). *If $Y' = A(t)Y$ is an ordinary differential equation then*

$$\Phi(t) = \exp\left(\int_{t_0}^t A(s)^T ds\right)^T$$

provides a local fundamental matrix.

A.2 Cauchy-Kovalevskya (Cauchy-Kowalevski)

Bibliography

- [AM16] M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*, economy ed., Addison-Wesley Series in Mathematics, Westview Press, Boulder, CO, 2016, For the 1969 original see [MR0242802]. MR 3525784
- [AvMV04] Mark Adler, Pierre van Moerbeke, and Pol Vanhaecke, *Algebraic integrability, Painlevé geometry and Lie algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 47, Springer-Verlag, Berlin, 2004. MR 2095251
- [Bea90] Arnaud Beauville, *Jacobiennes des courbes spectrales et systèmes hamiltoniens complètement intégrables*, Acta Math. **164** (1990), no. 3-4, 211–235. MR 1049157
- [BGK⁺87] A. Borel, P.-P. Grivel, B. Kaup, A. Haefliger, B. Malgrange, and F. Ehlers, *Algebraic D-modules*, Perspectives in Mathematics, vol. 2, Academic Press, Inc., Boston, MA, 1987. MR 882000
- [Bui86] Alexandru Buium, *Differential function fields and moduli of algebraic varieties*, Lecture Notes in Mathematics, vol. 1226, Springer-Verlag, Berlin, 1986. MR 874111
- [Bui92] ———, *Differential algebraic groups of finite dimension*, Lecture Notes in Mathematics, vol. 1506, Springer-Verlag, Berlin, 1992. MR 1176753

- [Bui94] ———, *Differential algebra and Diophantine geometry*, Actualités Mathématiques. [Current Mathematical Topics], Hermann, Paris, 1994. MR 1487891
- [Cox12] David A. Cox, *Galois theory*, second ed., Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2012. MR 2919975
- [Del70] Pierre Deligne, *Équations différentielles à points singuliers réguliers*, Lecture Notes in Mathematics, Vol. 163, Springer-Verlag, Berlin-New York, 1970. MR 0417174
- [Eva10] Lawrence C. Evans, *Partial differential equations*, second ed., Graduate Studies in Mathematics, vol. 19, American Mathematical Society, Providence, RI, 2010. MR 2597943
- [Hat02] Allen Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR 1867354
- [IKSY91] Katsunori Iwasaki, Hironobu Kimura, Shun Shimomura, and Masaaki Yoshida, *From Gauss to Painlevé*, Aspects of Mathematics, E16, Friedr. Vieweg & Sohn, Braunschweig, 1991, A modern theory of special functions. MR 1118604
- [Inc44] E. L. Ince, *Ordinary Differential Equations*, Dover Publications, New York, 1944. MR 0010757
- [Kap76] Irving Kaplansky, *An introduction to differential algebra*, second ed., Publications de l'Institut de Mathématique de l'Université de Nancago, No. V, Hermann, Paris, 1976. MR 0460303
- [Kat87] Nicholas M. Katz, *A simple algorithm for cyclic vectors*, Amer. J. Math. **109** (1987), no. 1, 65–70. MR 878198
- [Kol73] E. R. Kolchin, *Differential algebra and algebraic groups*, Academic Press, New York-London, 1973, Pure and Applied Mathematics, Vol. 54. MR 0568864
- [Lee13] John M. Lee, *Introduction to smooth manifolds*, second ed., Graduate Texts in Mathematics, vol. 218, Springer, New York, 2013. MR 2954043

- [Mag94] Andy R. Magid, *Lectures on differential Galois theory*, University Lecture Series, vol. 7, American Mathematical Society, Providence, RI, 1994. MR 1301076
- [Mat80] Michihiko Matsuda, *First-order algebraic differential equations*, Lecture Notes in Mathematics, vol. 804, Springer, Berlin, 1980, A differential algebraic approach. MR 576060
- [Mul94] Motohico Mulase, *Algebraic theory of the KP equations*, Perspectives in mathematical physics, Conf. Proc. Lecture Notes Math. Phys., III, Int. Press, Cambridge, MA, 1994, pp. 151–217. MR 1314667
- [Oka86] Kazuo Okamoto, *Studies on the Painlevé equations. III. Second and fourth Painlevé equations, P_{II} and P_{IV}* , Math. Ann. **275** (1986), no. 2, 221–255. MR 854008
- [Oka87a] ———, *Studies on the Painlevé equations. I. Sixth Painlevé equation P_{VI}* , Ann. Mat. Pura Appl. (4) **146** (1987), 337–381. MR 916698
- [Oka87b] ———, *Studies on the Painlevé equations. II. Fifth Painlevé equation P_V* , Japan. J. Math. (N.S.) **13** (1987), no. 1, 47–76. MR 914314
- [Oka87c] ———, *Studies on the Painlevé equations. IV. Third Painlevé equation P_{III}* , Funkcial. Ekvac. **30** (1987), no. 2-3, 305–332. MR 927186
- [Rit32] Joseph Fels Ritt, *Differential equations from the algebraic standpoint*, American Mathematical Society Colloquium Publications, vol. 14, American Mathematical Society, New York, 1932. MR 3822122
- [Rit50] ———, *Differential Algebra*, American Mathematical Society Colloquium Publications, Vol. XXXIII, American Mathematical Society, New York, N. Y., 1950. MR 0035763
- [Sch14] Wilhelm Schlag, *A course in complex analysis and riemann surfaces*, vol. 154, American Mathematical Society Providence, RI, 2014.

- [Sha13] Igor R. Shafarevich, *Basic algebraic geometry. 2*, third ed., Springer, Heidelberg, 2013, Schemes and complex manifolds, Translated from the 2007 third Russian edition by Miles Reid. MR 3100288
- [Ull08] David C. Ullrich, *Complex made simple*, Graduate Studies in Mathematics, vol. 97, American Mathematical Society, Providence, RI, 2008. MR 2450873
- [Ume11] Hiroshi Umemura, *Picard-Vessiot theory in general Galois theory*, Algebraic methods in dynamical systems, Banach Center Publ., vol. 94, Polish Acad. Sci. Inst. Math., Warsaw, 2011, pp. 263–293. MR 2905461
- [Vak17] Ravi Vakil, *The rising sea: Foundations of algebraic geometry*, 2017.
- [vdPS03] Marius van der Put and Michael F. Singer, *Galois theory of linear differential equations*, Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 328, Springer-Verlag, Berlin, 2003. MR 1960772