An Introduction to the Algebraic Theory of Differential Equations

Taylor Dupuy

October 26, 2022

Contents

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 The issue is that they are a bit out of date and have a lot of dependencies within chapters that are not clearly marked so you almost have the read the books linearly.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 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. They are about differential field theory, differential algebraic groups, and applications of differential algebra to diophantine geometry.

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.

Also, in the spirit of these Differential Algebra books are the are the importants 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 and there is some folklore here that needs to be fleshed out in my humble opinion.

Once one gets into the Painlevé equations more algebraic geometry surfaces. The Japanese following Okomoto [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 Spectral Curves, Grassmannians, Jacobian Flows, and Krichever modules [Mul94]. We would be crazy not to mention the work of Sato and the book by Miwa, Jimbo, and Date about infinite dimensional Grassmannians.

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 holonomic 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] (mostly complex manifolds which are not covered in loc cit.), 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? What do people mean when they say classical asymptotic methods are "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 .pdf can be found here:

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

0.3 How do I cite these notes?

♠♠♠ Taylor: [WARNING: These notes have not been peer reviewed! Use at your own peril! (Also email me if you find mistakes: taylor.dupuy@gmail.com)]

```
@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 map a map of sets $\partial : R \to R$ that satisfied

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

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

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

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 $\partial : R \to R$ 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 $\partial_i : R \to R$.

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 $(\mathbb{C}[t], \dfrac{d}{dt})$
2. The ring of rational functions $(\mathbb{C}(t), \dfrac{d}{dt})$, 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 $\mathrm{Hol}(U)$ for some $U \subset \mathbb{C}^m$ is an example of a Δ-ring, $(\mathrm{Hol}(U), \{\dfrac{\partial}{\partial t_1}, \dots, \dfrac{\partial}{\partial t_m}\})$. 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 $\mathrm{Mer}(U)$. 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 $f : A \to B$ such that for each $\partial_i \in \Delta$ we have $f(\partial_i(a)) = \partial_i(f(a))$ for each $a \in A$.

1.1.3 Radicals of Differential Ideals

Let K be a ∂-field and let

$$K\{x\} = K\{x\}_\partial = K[x]_\partial = K[x, x', x'', \dots]$$

be the ring of ∂-polynomials.

For a subset A of $K\{x\}$ we will let

$$[A] = [A]_\partial$$

be the ∂-ideal generated by A. It is the smalled δ-ideal containing A. We will let

$$\{A\} = \{A\}_\partial = \sqrt{[A]}$$

be the smallest radical ideal containing A.

Lemma 1.1.3.1 (Radicals of Differential Ideals are Differential Ideals)**.** *Let A be a differential $\mathbb{Q}$-algebra. If I is a differential ideal then $\sqrt{I}$ is a differential ideal.*

Proof. Suppose $a \in \sqrt{I}$. Then there exists some natural number n such that $a^n \in I$. By the Power Lemma 5.4.2.1 we have that $(a')^n \in I$. This implies $a' \in \sqrt{I}$. □

Exercise 1.1.3.2. Show that the intersection of two radical ideals is radical.

Exercise 1.1.3.3. Give an example of an ideal I which is radical such that I^2 is not radical.

Chapter 2

Monodromy and Hilbert's 21st Problem

The statement of Hilbert's 21st problem to the 1900 International Congress of Mathematicians (ICM) is as follows.

"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, regular[1] 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

[1] Hilbert means holomorphic.

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.

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 $\mathbb{P}^1$.

2.1.1 Wronskians

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

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

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 $f_1, \ldots, f_n \in K$. Let $C = K^\partial$ be the constants. We have that $f_1, \ldots, f_n$ are linearly dependent over C if and only if $W(f_1, \ldots, f_n) = 0$.*

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

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

Taking derivatives gives

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

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

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

To get

$$\begin{aligned} 0 =& c_1 f_1^{(j+1)} + \cdots + c_n f_n^{(j+1)} + c_1' f_1^{(j)} + \cdots + c_n' f_n^{(j)} \\ =& c_2' f_2^{(j)} + \cdots + c_n' f_n^{(j)}. \end{aligned}$$

But $f_2, \ldots, f_n$ are linearly independent. This implies that $c_2' = \cdots = c_n' = 0$ which implies that $c_1, \ldots, c_n \in C$ which proves our result. □

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

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

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, tt turns out that every $n \times n$ matrix and we can obtain its formula from cofactor expansion. We have

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

where $\widetilde{A}_{ij}$ is the matrix obtains by deleting the ith row and jth 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 $\mathbb{Z}[x_{ij} \colon 1 \leq i, j \leq n]$.

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

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

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

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

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

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

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

[2]Named after Wilhelm Killing 1847–1923

We can now prove our result.

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

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

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

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

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

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

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

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 $\mathrm{Hol}(U) \to \mathrm{Hol}(U)$ and $\mathrm{Mer}(U) \to \mathrm{Mer}(U')$ given by restricting the domain of some $f(z)$ to U'. Both of these ring homomorphisms are injective by the analytic continutation principle (which holds in several variables as well as one variable).[3] The *stalk* at some $t_0 \in \mathbb{C}$ is

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

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

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

It is important to know that there is always a ring homomorphism $\mathrm{Hol}(U) \to \mathrm{Hol}_{t_0}$ and that any element of $\mathrm{Hol}(U)$ 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 $f_{i,j} : R_i \to R_j$ 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

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

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

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

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

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

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

$$\begin{cases} v_0' = v_1, \\ v_1' = v_2, \\ \ddots \\ v_{r-1}' = -a_{r-1}v_{r-1} - a_{r-2}v_{r-2} - \cdots - a_0 v_0. \end{cases}$$

which then can be written in matrix form

$$V' = AV$$

where

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

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.

Remark 2.1.3.1. In the case of linear differential equations there is a way of going backwards using cyclic vectors. That is, most first order linear systems of differential equations n variables in one derivative can be converted into a order n equation in n dependent variables.

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 $Y = (y_1, \ldots, y_n)$ (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 $\Phi \in \mathrm{GL}_n(R)$ 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 menas that if $\widetilde{\Phi}$ is another fundamental matrix there exists some $M \in \mathrm{GL}_n(R^\partial)$ such that

$$\widetilde{\Phi} = \Phi M. \tag{2.1.3}$$

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 $\widetilde{\Phi}$ are two fundamental matrices of a rank n linear system over a differential ring (R, ∂) then there exists some $M \in \mathrm{GL}_n(R^\partial)$ such that $\widetilde{\Phi} = \Phi M$.*

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 $Y_1, \ldots, Y_n \in K^n$ 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

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

But since $Y_2, \ldots, Y_n$ were assumed to be linearly independent we must have $c_2' = \ldots = c_n' = 0$ which proves the c_j's are constants. □

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-Kowalevski[4] 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 $K \subset \mathbb{C}\langle\langle t - t_0 \rangle\rangle$ 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 Siedenberg embedding theorem then tells us that $\widehat{K} \subset \mathbb{C}\langle\langle t - t_0 \rangle\rangle$, 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].

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 ith position). The solutions $Y_1, \ldots, Y_n$ are linearly independent over $K = \mathbb{C}\langle t - t_0 \rangle^{\oplus n}$ because $e_1, \ldots, e_n$ are linearly independent over K. Hence by Lemma 2.1.4.3 we get that the solutions are linearly independent over over $\mathbb{C}$. The matrix

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

is our fundamental system. $\square$

2.1.6 Restricting the Coefficient Matrix to a Lie Algebra

It will be conventient in the equation $Y' = AY$ to restrict the matrix A to a particular Lie algebra $\mathrm{Lie}(G)$ of some Lie group G. Here we recall

[4] This is the same as Cauchy-Kovaleskaya. Some people spell the Russian name differently.

type	Lie Group	Lie Algebra
complex	$\mathrm{GL}_n(\mathbb{C})$	$M_n(\mathbb{C})$
real	U_n unitary, $U^* = U^{-1}$	skew-adjoint $A^* = -A$
complex	$\mathrm{SL}_n(\mathbb{C})$, $\det(A) = 1$	trace free $\mathrm{tr}(B) = 0$

Table 2.1: Some common Lie groups and their Lie algebras.

that a Lie group (real or complex) is just a manifold (real of complex) with the structure of group. [5] The Lie algebra of such a Lie group $\mathrm{Lie}(G)$ can be described either as the group of tangent vector at the identity of G or as the globally invariant vector fields on G (obtained by propagating the tangent vector at the tangent space at the identity to any other point of the manifold by pushforward by multiplication-by-g). All Lie algebras come with a Lie bracket $(A, B) \mapsto [A, B]$ which is an infinitesimal version of the group multiplication. It turns out that this multiplication satisfies the so-called Jacobi identity. The fundamental property of Lie algebras of Lie groups are that if G is a matrix Lie group then

$$A \in \mathrm{Lie}(G) \implies e^A \in G.$$

Table gives some common Lie groups with their Lie algebras. As a sanity check not that if $A^* = -A$ then $(e^A)^* = e^{A^*} = e^{-A} = (e^A)^{-1})$ so skew-adjoint matrices give rise to unitary matrices after exponentiation.

For the uninitiated we mention that we often axiomatize the notion of a Lie algebra as an R-module (or a functor to R-modules) which has a Lie bracket and satisfies the Jacobi identity axiom. This is useful but not what we mean here. The collection of abstract vector fields on a space (scheme, complex manifold, real manifold) satisfy these axioms for example and here the ring R is a ring of functions on a space.

The main reason we mention Lie groups is because if we restrict our linear equations to have values in a Lie algebra, then the solution will be valued in a Lie group.

[5] There are also algebraic group or group schemes but readers familiar with those already know all of this, so I'm going to not say anything about that as it will take us two far afield.

Theorem 2.1.6.1. *Consider $Y'(t) = A(t)Y(t)$. If $A(t)$ is valued in* $\mathrm{Lie}(G)$ *then any fundamental matrix will be valued in G.*

Proof. This is a consequence of Theorem A.1.0.1 □

2.1.7 Monodromy of Holomorphic Linear Systems

Consider a holomorphic linear system

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

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 $\mathrm{GL}_n(\mathbb{C})$ which measures how much Φ changed once we take it around the look.

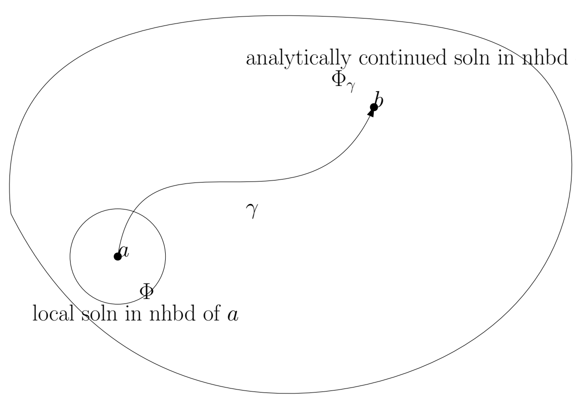

Figure 2.1: A picture of analytically continuing a local fundamental matrix along a path.

The group homomorphism

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

is called the *monodromy representation*. We will now explain what $M_\gamma \in \mathrm{GL}_n(\mathbb{C})$ 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

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

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.

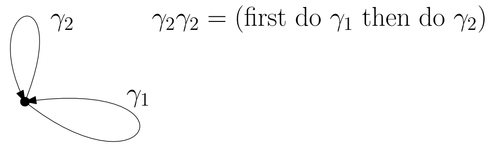

Figure 2.2: The convention we use for composition of paths. Other people use other conventions and it will mess up your formulas.

Remark 2.1.7.1 (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 $(\Phi_{\gamma_1})_{\gamma_2} = (\Phi M_{\gamma_1})_{\gamma_2} = \Phi_{\gamma_2}(M_{\gamma_1})_{\gamma_2} = \Phi M_{\gamma_2} M_{\gamma_2}$. 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 2.1.7.2 (Babymost Example)**.** In the rank one case we have a differential equations

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

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 to 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 issue 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 $M_1 \in \mathrm{GL}_n(\mathbb{C})$ 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

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

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

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

and we see that

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

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

Example 2.1.7.4 (Euler Systems)**.** In a punctured neighborhood around $0 \in CC$, 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

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

so the matrix $\Phi(t) = t^A$ is a local matrix solution of this equation. Since $\det(e^B) = e^{\mathrm{Tr}(B)}$ for any matrix B we have $\det(\Phi) = t^{\mathrm{Tr}(A)}$ 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 2.1.7.5. Every matrix $M \in \mathrm{GL}_n(\mathbb{C})$ can appear as the monodromy matrix of some system. (Hint: use the Euler system and show that for every $M \in \mathrm{GL}_n(\mathbb{C})$ 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 $|a_{n+1}|/|a_n| = n \to \infty$ as $n \to \infty$.

So what is the issue? The issue is that when we convert this equation into a first order system of differential equations is 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 $A(t) \in M_n(\mathrm{Hol}(\mathbb{P}^1 \setminus T))$ 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 at Fuchsian and Fuchsian at a point if there 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 $\mathrm{mult}_{t=t_0}(a_j(t)) \geq n - j$. 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

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

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:

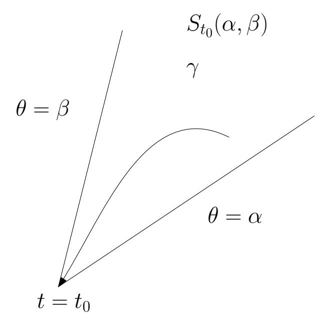

Figure 2.3: A sector used in the definition of regular singular.

In what follows we will let $S_{t_0}(\alpha, \beta) = \{t \in \mathbb{C} : \alpha < \arg(t - t_0) < \beta$ where $t_0 \in \mathbb{C}$, $\alpha, \beta \in [0, 2\pi]$ with $\alpha > \beta$ and arg 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 eminating 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 $Y_1(t), \ldots, Y_n(t)$ with $Y_j(t) = (y_{j1}(t), \ldots, y_{jn}(t))^T$ and $\lambda \in \mathbb{C}$ such that

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

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 soltion $Y(t)$ 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

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

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

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

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.

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

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 $\rho : \pi_1(\mathbb{P}^1 \setminus T) \to \mathrm{GL}_n(\mathbb{C})$ 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 Plemelj[6] 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.

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

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

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

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

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

the generators then can be taken to be homotopy classes of loops around each of the points $t_1, \ldots, t_{n-1}$. 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 $\pi_1(\mathbb{P}^1 \setminus T) \to \mathrm{GL}_n(\mathbb{C})$ are determined by tuples $(M_1, M_2, \ldots, M_{n-1}) \in \mathrm{GL}_n(\mathbb{C})^{n-1}$ modulo simultaneous conjugation by an element in $\mathrm{GL}_n(\mathbb{C})$. Here $M_j = \rho(\gamma_j)$.

[6]Nalini Joshi pronounces this "Plum-ell-i", I'm not sure how to pronounce this name

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

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

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

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

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

$$\widetilde{Y}' = \widetilde{A}Y, \qquad \widetilde{A} = \Phi^{-1}A\Phi - \Phi^{-1}\Phi'.$$

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

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

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

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

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

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

As an exercise one needs to compute

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

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 (Equivalence to Trivial Equation)**.** Consider a linear differential system $Y' = AY$ over a differential ring (R, ∂). Suppose that the system admits a fundamental matrix $\Phi \in \mathrm{GL}_n(R)$. Then setting $Y = \Phi\widetilde{Y}$ 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

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

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 $S = \{a_1, a_2, \ldots, a_m\} \subset \mathbb{P}^1 \setminus \{\infty\}$ takes the form

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

where $A_1, A_2, \ldots, A_m \in M_n(\mathbb{C})$ 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 $\mathbb{C}^n$ an get a category of complex manifolds. It turns out that the category of compact Riemann surfaces 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 exists 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 how 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 assocaited field is the field of meromorphic functions on X which we denote by $\mathrm{Mer}(X)$. 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 $\mathrm{Mer}(\mathbb{P}^1) = \mathbb{C}(t)$ 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 $f \in \mathrm{Mer}(\mathbb{P}^1)$ 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 $\mathrm{ord}_{t=\infty}(f(t)g(t)) = \mathrm{ord}_{s=0}(f(1/s)g(1/s))$. 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 $\mathrm{Mer}(X)$ is a finite algebraic extension of $\mathbb{C}(t)$. To give an idea of how different-looking $\mathrm{Mer}(X)$ and $\kappa(X)$ can be consider the case of an elliptic curve E. As a Riemann surface we like to describe this as $\mathbb{C}/\Lambda$ for some lattice $\Lambda \subset \mathbb{C}$. In this situation, we have $\mathrm{Mer}(E) = \mathbb{C}(\wp_\Lambda(t), \wp'_\Lambda(t))$ 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 $\mathbb{C}/\Lambda \to E$ 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 $f(t)dt$. [7] We can then 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

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

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 $\sum_{a\in X} \mathrm{res}_{t=a}(\omega) = 0$.*

[7] In 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.

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

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

□

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

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

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

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

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 $S = \{s_1, \ldots, s_m\}$ we will often assume that S takes the form $S = \{0, 1, \infty, s_4, \ldots, s_m\}$ 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 $\mathbb{P}^1$,*

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

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

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

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

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 $s_1, \ldots, s_m$ 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 $A_1, \ldots, A_m \in M_n(\mathbb{C})$ such that the components of

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

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 $\lim_{t \to s_j} B(t) = 0$ 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 $A(t)dt$. □

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 $S = \{s_1, \ldots, s_m\} \subset \mathbb{P}^1$ can be taken without loss of generality to be $S = \{0, 1, \infty, s_4, \ldots, s_m\}$ 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 $-s^2 \dfrac{dY}{ds} = s A_0 Y$ 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 criterian 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

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

It 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 $Y(t) = t^\rho \widetilde{Y}(t)$.

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 $R[\partial\colon \partial\in\Delta]$ where one has

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

One also has $\partial_1\partial_2 = \delta_2\delta_1$ for $\partial_1,\partial_2\in\Delta$.

The idea behind the formula $\partial a = a\partial + \partial(a)$ comes from looking a $a\in R$ 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 are for polynomial rings and refer to this particular Weyl algebra as *the* Weyl algebra. In this case they take $R = \mathbb{C}[x_1, \ldots, x_m]$ with $\Delta = \{\partial_{x_1}, \ldots, \partial_{x_m}\}$ and then only talks 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 A Bosonic Fock Space and Weyl Algebras

A frequently used physical perspective of D-modules that occurs is the following. Let

$$B = \mathbb{C}[x_1, x_2, \ldots][\partial_1, \partial_2, \ldots], \quad \partial_j = \frac{\partial}{\partial x_j}.$$

This is just a usual Weyl algebra but in countably many variables. In some portions of the algebraic theory of differential equations literature [SS83, MJD00], one uses the terminology of creation and annihilation operators,

$$a_n = \partial_n = \text{(annihilation operator)}, \quad a_n^* = x_n = \text{ (creation operator) },$$

and observes that these satisfy the rules

$$[a_n, a_m] = 0, \quad [a_m^*, a_n^*] = 0, \quad [a_m, a_n^*] = \delta_{mn}, \qquad m, n \in \mathbb{Z}.$$

From this viewpoint we think of a_n and a_n^* as acting on a space of functions which create and annihilate bosonic particles on the $\mathbb{Z}$ and the D-module (which is a B-module in this case) $\mathbb{C}[x_1, x_2, \ldots]$ of polynomials in infinitely many indeterminates is called the bosonic Fock space. The idea here is that particles along the lattice $\mathbb{Z}$ are built-up by acting by these creation and annihilation operators on $1 \in \mathbb{C}[x_1, x_2, \ldots]$ which is called the *vacuum state*.

None of this is really important at the moment and we just say this so that the reader can recognize in the literature that people talking about Bosonic Fock Spaces are really just talking about Weyl algebras.

Remark 2.5.2.1. For the uninitiated we mention that Bosons are particles that mediate the exchange of forces in physics.[8] There are a set of special Bosons for each of the fundamental forces: the photon the γ-boson for the electromagnetic force; the weak force for three bosons the W^+-boson, the W^--boson, and the Z-boson – the W-bosons carry a charge and the Z-boson does not; the strong force has six bosons called gluons.

In addition to there there is a Higgs boson which gives particles mass by the so called Higgs Mechanism.

Fermionic Fock Spaces

Recall that for every quadratic space (V, q) we can associated a Clifford algebra $\mathrm{Clf}(V, q)$ where $\mathrm{Clf}(V, q) = T(V)/I_q$ and I_q is generated by $v^2 = q(v)$ and $T(V) = \bigoplus_{n \geq 0} V^{\otimes n}$ is the tensor algbra (so v^2 really is $v \otimes v$ but we are dropping the $\otimes$ to make notation simple). For any such (V, q) there is an associated bilinear form $2B_q(v + w) = q(v + w) - q(v) - q(w)$ and we have the identity for $v, w \in V \subset \mathrm{Clf}(V, q)$ given by

$$\{v, w\} = vw + wv = 2B_q(v, w).$$

When 2 is invertible in your base ring the data of a bilinear form or quadratic form are equivalent.

We now perform our construction of the ring of so-called Grassmann numbers. Let V be a countably generated vector space with a basis v_i for $i \in \mathbb{Z}^*$.

[8] with the exception of the Higgs boson.

Let W be a countably generated vector sapce with a basis w_i for $i \in \mathbb{Z}^*$. We then give a bilinear form B_q on $V \oplus W$ given by

$$B_q(v_i, v_j) = 0, \quad B_q(w_i, w_j) = 0, \quad B_q(v_i, w_j) = \begin{cases} 0, & i+j \neq 0 \\ \frac{1}{2}, & i+j = 0 \end{cases}.$$

This gives a quadratic space and we have the description as a clifford algebra being given by

$$A = \mathrm{Clf}(V \oplus W, q).$$

For some reason, we then use the notation $v_n = \psi_n$ and $w_n = \psi_n^*$ for $n \in \mathbb{Z}^*$ and observe the identities:

$$\{\psi_m, \psi_n\} = 0, \quad \{\psi_m^*, \psi_n^*\} = 0, \quad \{\psi_m^*, \psi_n^*\} = \delta_{n+m,0}, \qquad m, n \in \mathbb{Z}^*.$$

Traditionally, the ψ_n's and ψ_n^*'s are indexed by $m, n \in \frac{1}{2} + \mathbb{Z}$ but we are going to forego this tradition because these are math notes and physics notation is silly. Note in particular that $\psi_n^2 = 0$ and $(\psi_n^*)^2 = 0$. Mathematically the ring generated by ψ_n for $n \in \mathbb{Z}^*$ is the exterior algebra of a countably generated free module and the same can be said for ψ_n^*.

Remark 2.5.2.2. For the uninitiated we mention that Fermions are particles like electrons on which forces act. Several behave very similarly to electrons mathematically and these are called leptons (12 in total). First every lepton has an anti-particle. For the electron this is a positron. These mathematically are pretty much identical but with time reversed (yes, weird, but mathematically simple). Then there are "heavier electrons" called muons and tauons which are like electrons but, well, heavier. Then there are the baby versions where are called neutrinos. There are electron neutrinos, muon neutrinos, and tauon neutrinos. These also have antiparticles.

The other six elementary Fermions in the standard model of 2022 are the quarks which are associated with the strong force. These are the cute sounding up, down, strange, charm, top, and bottom quarks.

2.5.3 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

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

we can define a D-module structure on $V = R^{\oplus n}$. Let $e_1, \ldots, e_n$ 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 2.5.3.1. 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 $v_1, \ldots, v_n$ and finds that

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

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

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

One then finds that the linear differential equation associated to the D-module is again the D-module with $v_1, \ldots, v_n$ 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

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

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 is gives rise to a gauge transformation of the associated linear system.

Exercise 2.5.3.2. Show that indeed a change of coordinates on the R-module induces a gauge tranformations of the linear differential equation.

2.5.4 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 2.5.4.1. 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

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

form a basis for V. This is probably the most important case. Before proving such cyclic vectors exist, lets 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 use $\partial(v_i) = v_{i+1}$ for $0 \leq i \leq n-2$ and then $\partial(v_{n-1}) = b_0 v_0 + b_1 v_1 + \cdots + b_{n-1} v_{n-1}$ for $b_i \in R$ and we get the linear systems

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

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 2.5.4.2 (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 $e_1, \ldots, e_n$ is the standard elementary basis for R^n.

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

Lemma 2.5.4.3. *Let $V = \mathbb{C}[[t]]^{\oplus n}$ be a D-module.*

1. *Let $h_0, \ldots, h_{n-1} \in V$ be horizontal (i.e. suppose $\partial(h_j) = 0$). Consider $v = \sum_{j=0}^{n} \dfrac{t^j}{j!} h_j$. The vector v is cyclic.*
2. *For each $v_0 \in V$ consider the system*

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

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. □

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, $\widetilde{e}_j := e^{-t\partial} e_j$ is also a basis for V since it is a basis modulo tV. Furthermore, by the Lemma $\partial(\widetilde{e}_j) = 0$. We now apply part one of Lemma 2.5.4.3 to get

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

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

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

gives a cyclic vector. □

2.5.5 Local Exponents

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

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

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 2.5.5.1. Let $R = \mathrm{Res}_{t=t_0}(A(t)) \in M_n(\mathbb{C})$ 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

$$\begin{array}{cccc}
s_1 & s_2 & \cdots & s_m \\
\hline\hline
\rho_1(s_1) & \rho_1(s_2) & \cdots & \rho_1(s_m) \\
\rho_2(s_1) & \rho_2(s_2) & \cdots & \rho_2(s_m) \\
\vdots & \vdots & \ddots & \vdots \\
\rho_n(s_1) & \rho_r(s_2) & \cdots & \rho_n(s_m)
\end{array}$$

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 $\rho - \mu \in \mathbb{Z}$.

2.5.6 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 2.5.6.1. In this problem $\partial = \partial_t$. Show that

1. $t^n\partial^n = \theta(\theta-1)\cdots(\theta-n+1)$
2. $\theta(t^\rho f) = t^\rho(\theta+m)f$.
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 2.5.6.2. 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 $b_1 = ta_1$ so that

$$M = t^2 a_0 + (ta_1 - 1)\theta + \theta^2.$$

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

Exercise 2.5.6.3. 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 $b_{i0} = b_0(0)$ and the indicial equation.

Theorem 2.5.6.4 (Indicial Equation). *If $t^\rho f(t)$ is a formal solution of L then ρ is a solution of*

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

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

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

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. □

Exercise 2.5.6.6. 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 $c_j = \lim_{t\to t_0}(t-t_0)^{n-j}a_j(t)$.
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 $c_j = \lim_{t\to\infty} t^{n-j}a_j(t)$.

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

Exercise 2.5.6.7. 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.

2.5.7 Fuchs' Relation

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

Theorem 2.5.7.1 (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 $\rho_1(a), \ldots, \rho_n(a)$ 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 mposes 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 2.5.7.2. 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

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

similarly for $a = \infty$ we have

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

Using the residue formula (Theorem 2.4.2.1) we have

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

which proves the result. □

2.5.8 Local Solutions of Exponent ρ

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

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

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 2.5.8.1. *Consider the formal Fuchsian system*

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

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.

We record the following for later use.

Theorem 2.5.8.2. ♠♠♠ Taylor: [Check and add. We might need to add more about formal power series solutions to do Stokes stuff later.] *More generally, in the non-resonant case one has a fundamental matrix of the form*

$$\Phi(t) = \Psi(t)t^{A_0}$$

where $\Psi(t)$ is a matrix of formal power series. The matrix $\Psi(t)$ is convergent if $A(t)$ is convergent.

2.5.9 Exponent Shifting

We record the following useful fact.

Lemma 2.5.9.1 (Exponent Shifting)**.** *The gauge transformation $Y = (t - t_0)^{\mu}\widetilde{Y}$ has the effect of $A(t)/(t-t_0) \mapsto \widetilde{A}(t)/(t-t_0)$ where $\widetilde{A}(t) = A(t) - \mu$.*

Proof. Without loss of generality we can suppose that $t_0 = 0$ since ∞ is invariant under the transfomation $t \mapsto t - t_0$. 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 $\theta(t^{\mu}\widetilde{Y}) = t^{\mu}(\theta + \mu)\widetilde{Y}$ and $A(t)t^{\mu}\widetilde{Y} = t^{\mu}A(t)\widetilde{Y}$ which gives the equation

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

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. $\square$

2.5.10 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 2.5.10.1. *Let $S = \{t_1, t_2, t_3\} \subset \mathbb{P}^1$ and fix a table of local exponents satisfying Fuchs' relation*

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

There exists a unique $a_1(t), a_2(t) \in \mathbb{C}(t)$ 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 $t_3 = \infty$. □

Theorem 2.5.10.2. *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

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

We next apply the exponent shifting lemma (Lemma 2.5.9.1). Making the gauge transformation $y = t^{-\alpha}(t-1)^{-\beta}\widetilde{y}$ to tranform the exponent table again to

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

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 2.5.7.1) forces $\beta' - \beta = c - a - b$. Since the exponents determine the equation (Theorem 2.5.10.1) the transformed equation must be a hypergeometric equation with the given exponents. $\square$

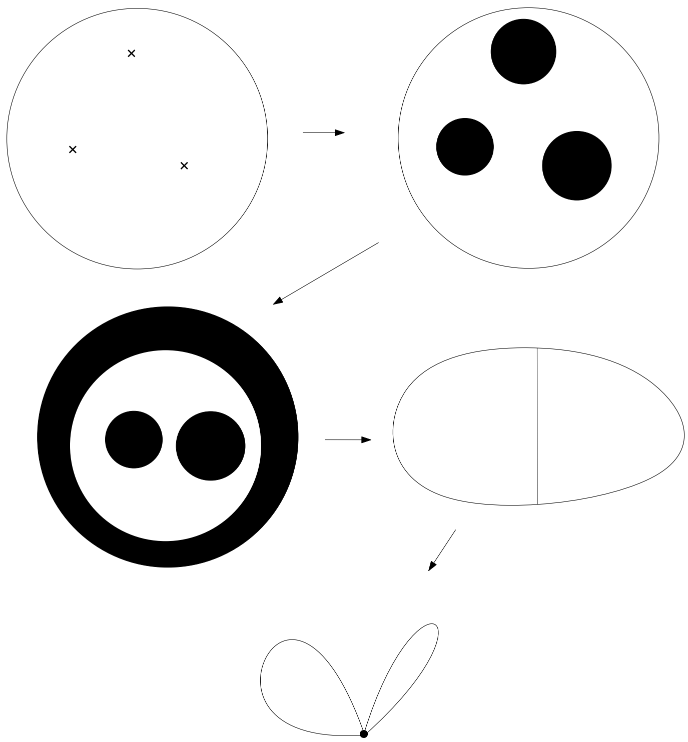

Figure 2.4: The figure shows $\mathbb{P}^1 \setminus \{0, 1, \infty\}$ being deformed into $S^1 \vee S^1$. The first step is the increase the size of the holes to make it look like a bowling ball. We then wrap one of the holes completely around to get a disc with two interior discs removed. This is then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

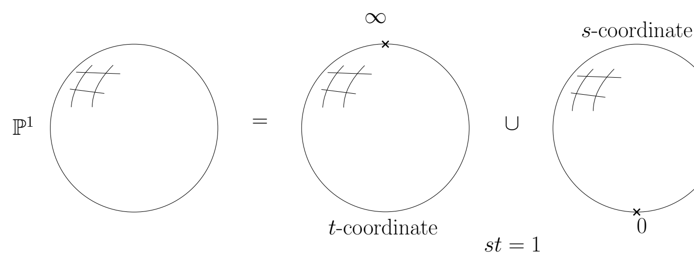

Figure 2.5: The projective line $\mathbb{P}^1$ is isomorphic to the Riemann sphere S^2 and is composed of two coordinate charts. The first chart we think of as the "usual" copy of $\mathbb{C}$ (which algebraic geometers upgrade to the affine line $\mathbb{A}^1$) which has coordinate t. Then when we want 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.

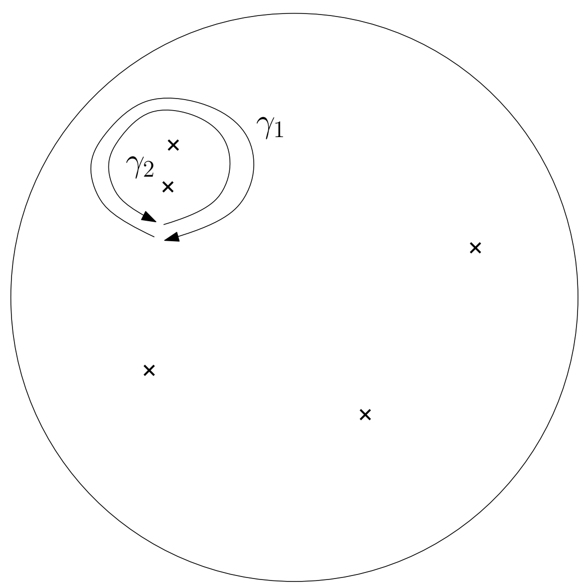

Figure 2.6: One uses the basic residue theorem on two simple integrals which are opposite of each other to prove the residue theorem on $\mathbb{P}^1$.

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

[1] There 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 lattitude, 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.

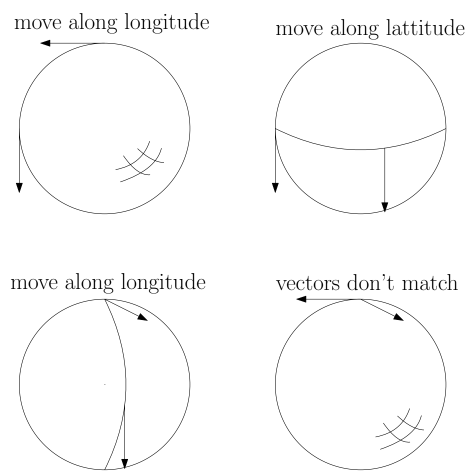

Figure 3.1: Curvature of the tangent bundle of S^2 associated to the Levi-Civita connection is evidenced by transporting tangent vectors around the the sphere.

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 two 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 $\psi : E \to R^{\oplus n}$ given by $v = f_1 v_1 + \cdots + f_n v_n \mapsto (f_1, \ldots, f_n)$ where $v_1, \ldots, v_n$ is a basis for E and $f_1, \ldots, f_n \in R$ 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 $\mathbb{P}^1 = U_0 \cup U_\infty$.

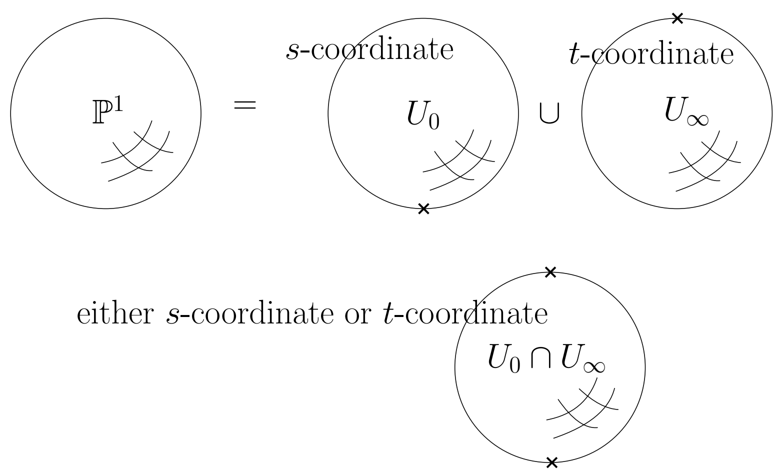

Figure 3.2: $\mathbb{P}^1$ is covered by U_0 and U_∞. Here U_0 uses the standard t-coordinate and U_∞ uses the coordinate at infinity s given by $s = 1/t$. On their intersection $U_0 \cap U_\infty$ you can use either coordinate.

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

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

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 bundles 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 $\mathcal{O}_X(U) = \mathrm{Hol}(U)$ 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 $\psi_U : E(U) \to \mathcal{O}_X(U)^{\oplus n}$.

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 $E_i|U_i \cap U_j \cong E_j|U_1 \cap U_j$ 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 $E = \Omega^1_{\mathbb{P}^1}$. 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

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

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

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

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

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 $\mathbb{P}^1$ and 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 $-s^{-d} \in \mathcal{O}^\times_{\mathbb{P}^1}(U_0 \cap U_\infty)$ as a representative of the cohomology class in $H^1(\mathbb{P}^1, \mathcal{O}^\times) = \mathrm{Pic}(\mathbb{P}^1)$ 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 have transition matrices which are diagonal of the form $\mathrm{diag}(t^{d_1}, t^{d_2}, \ldots, t^{d_n})$. 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 $d_1, d_2, \ldots, d_n \in \mathbb{Z}$.

Proof Reference and Sketch. This can be found in [HM82] which works algebraically over a general ring. By GAGA, that every algebraic vector bundle on $\mathbb{P}^1$ (considered as a scheme) is equivalent to proving every holomorphic vector bundles on $\mathbb{P}^1$ has this form.

The prove relies on a matrix factorization theorem of Birkoff. If $A \in \mathrm{GL}_n(\mathbb{C}[t, t^{-1}])$ then there $BAC = D$ where $C = C(t)$ and $B = B(t^{-1})$ have entries in $\mathbb{C}[t]$ and $\mathbb{C}[t^{-1}]$ respectively and D is diagonal with each entry of the form t^s for some integer s.

One works on two charts of $\mathbb{P}^1$ and then factors the transition data using this theorem. This is a Čech cocycle classifying the vector bundle and has the appropriate diagonal form. This proves the results. □

3.1.2 Systems vs Connections

Consider a vector bundle of rank n with connection (E, ∇) on complex manifold X. In the special case that $E \cong \mathcal{O}_X^{\oplus n}$ we often speak of a the connection as a *system* since there is really no extra global information. For non-compact Riemann surfaces all connections are really just systems.

Theorem 3.1.2.1 (Grauert-Röhrl Theorem)**.** *Every holomorphic vector bundle of rank n on a non-compact Riemann surface X is isomorphic to $\mathcal{O}_X^{\oplus n}$.*

Proof Reference and Sketch. This is [For81, Theorem 30.4]. The proof is by induction on the rank of the vector bundle. They show first that line bundles are trivial using the so-called Runge approximation theorem. Then they do an explicit computation with Čech cocycles after appling the inductive hypothesis to reduce the transition functions to unipotent matrices. They then reduce further arguing about an additive Čech cocycle. □

This will mean that we don't need to worry about global information coming from the vector bundle when trying to establish a Riemann-Hilbert correspondence for log-connections on the projective line minus a finite set of points.

3.1.3 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".

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

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 $R^n \to R^n$ whree $R = \mathrm{Hol}(U)$. 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 matric of differential 1-forms. This defined a map $\nabla : R^n \to \Omega^1_{\mathbb{P}^1}(U) \otimes R^n$. 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 $\Omega^1_{\mathbb{P}^1}(U)$ 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.

3.1.4 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 3.1.4.1. A *connection* on E is a $\mathbb{C}$-linear map

$$\nabla : E \to \Omega^1_X \otimes_{\mathcal{O}_X} E$$

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

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

This will play an important role in the Riemann-Hilbert correspondence in the case that ∇ is a so-called integrable connection.

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

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

3.1.5 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

$$E(U) = \mathcal{O}_X(U)s_1 + \cdots + \mathcal{O}_X(U)s_n,$$

[2] I'm really working with ringed spaces here.

$$T_X(U) = \mathcal{O}_X(U)\frac{\partial}{\partial t_1} + \cdots + \mathcal{O}_X(U)\frac{\partial}{\partial t_m},$$
$$\Omega_X(U) = \mathcal{O}_X(U)dt_1 + \cdots + \mathcal{O}_X(U)dt_m.$$

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 $\nabla(s_j) \in \Omega^1_X(U) \otimes E(U)$ must look like. It must have the form

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

The Christoffel symbols are just the structure "constants" for the connection.

Definition 3.1.5.1. The elements $\Gamma^i{}_{j\alpha} \in \mathcal{O}_X(U)$ are called the *Christoffel symbols* of ∇ with respect to the local basis $s_1, \ldots, s_n$ of E and local coordinates $t_1, \ldots, t_m$ of X.

The 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

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

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.

3.1.6 Covariant Derivatives

We now given the definition of a covariant derivative.

Definition 3.1.6.1. Let $\theta \in T_X(U)$ be a derivation. We define the *covariant derivative associated to* θ to be the operator

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

In the above expression the pairing $\Omega^1_X \times T_X \to \mathcal{O}_X$ (which we can write as $\Omega^1_X \otimes T_X \to \mathcal{O}_X$) is extended to $(E \otimes \Omega^1_X) \otimes T_X \to E \otimes \mathcal{O}_X = E$ 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 proceedding I will make some remarks on notation. First, derivatives $\dfrac{\partial}{\partial t^\beta}$ are "naturally lowered" in so Einstein notation and for convenience we often write them as $\partial_\beta = \dfrac{\partial}{\partial t^\beta}$. 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 3.1.6.2. Let $t^1, \ldots, t^m$ be local coordinates for X and let $s_1, \ldots, s_n$ be a local basis for E. Then if $\theta = \dfrac{\partial}{\partial t^\beta}$ 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

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

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

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

3.1.7 Connection 1-forms

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

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

where $\omega \in M_n(\Omega^1_X(U))$ 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 tje exterior algebra $\Omega^\bullet_X = \bigoplus_{d=0}^m \Omega^\bullet_X$ is a sheaf of skew commutative ring satisfying

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

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

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

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 dueal basis in $E^\vee(U)$ where $E^\vee$ is the dual vector bundle. It is defined by

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

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 3.1.7.1. 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 $\Omega^1_X \otimes \mathrm{End}(E)$ by

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

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

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 3.1.7.2 (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*

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

This is going to be used when computing our formulas for curvature (Theorem 3.1.8.6).

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

We now verify the claim about the description in local coordinates. Let U be an open subset for which $s_1, \ldots, s_n$ is a basis for $E(U)$. Let $\psi : E(U) \to \mathcal{O}_X(U)^{\oplus n}$ 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 $E(U) \otimes \Omega_X(U) \to \mathcal{O}_X(U) \otimes \Omega_X^1(U)$ by functorality of the tensor product and we will abusively also denote this isomorphism by ψ. We now have a square,

$$
\begin{tikzcd}
s = f^i s_i \arrow[rrr, mapsto] \arrow[ddd, mapsto] & & & \nabla(s) \\
& \rotatebox[origin=c]{135}{$\in$} & & \rotatebox[origin=c]{90}{$\in$} \\
& & E(U) \arrow[r, "\nabla"] \arrow[d, "\psi"] & \Omega_X(U) \otimes E(U) \arrow[d, "\psi"] \\
\begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix} & \in & \mathcal{O}_X(U)^{\oplus n} \arrow[r, "\nabla^\psi"] & \Omega_X(U) \otimes \mathcal{O}_X(U)^{\oplus n}
\end{tikzcd}
$$

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

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

where we have written

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

We summarize the above discussion with the following lemma.

Lemma 3.1.7.3. *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 $t^1, \ldots, t^m$ and local basis $s_1, \ldots, s_n$ then in local coordinates $\nabla_\alpha : \mathcal{O}_X(U)^{\oplus n} \to \mathcal{O}_X(U)^{\oplus n}$ takes the form*

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

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

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

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 3.1.7.4. The matrix $\omega = A_\alpha dt^\alpha \in M_n(\Omega_X(U))$ is called a *connection 1-form*.

3.1.8 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 3.1.8 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

$$\nabla_\alpha \begin{pmatrix} f^1 \\ \vdots \\ f^n \end{pmatrix} = 0, \tag{3.1.3}$$

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 $Y = (f^1, \ldots, f^n)$ 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.

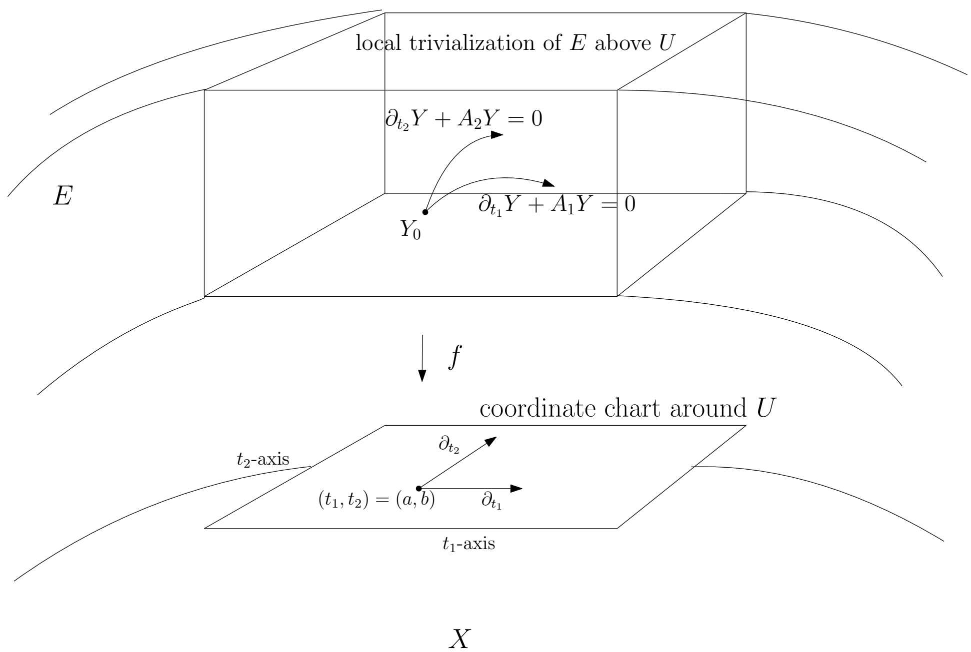

Figure 3.3: Movement in the base tells us how to move in the fibers by solving differential equations. The figure shows paths in the space being transported to the vector bundle locally by solving differential equations.

Problem 3.1.8.1 (Integrability Problem)**.** Is it possible to solve all of our equations (3.1.3) at once (i.e. simultaneously for $\alpha = 1, \ldots, m$)?

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

Problem 3.1.8.2 (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 concern that a solution of $\nabla_\alpha Y = 0$ may not also be a solution of $\nabla_\beta Y = 0$. The first equation is imposes a formula for $\partial_\alpha Y$ and the second equation imposes a formula 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 equations (or connection) is integrable.

As lifting paths is really about general derivations/vector fields. We now 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 3.1.8.3. 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:

$$[\nabla_{\theta_1}, \nabla_{\theta_2}] = \nabla_{[\theta_1, \theta_2]}? \tag{3.1.4}$$

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

Definition 3.1.8.4. The map $R_\nabla : T_X \otimes T_X \to \mathrm{End}(E)$ 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 $\nabla^2 : E \to \Omega^2_X \otimes E$. Hence ∇^2 can eat two tangent vectors θ_1 and θ_2 and a section s and spit out another section. The subsequet 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 between Ω^2_X and $T_X \wedge T_X$.

Exercise 3.1.8.5. 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 3.1.8.6 (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*
$$\frac{\partial A_\alpha}{\partial t^\beta} - \frac{\partial A_\beta}{\partial t^\beta} = -[A_\alpha, A_\beta].$$

More generally

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

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 $E(U)$ is 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.

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

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 3.1.7.2). 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. □

Exercise 3.1.8.7. 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

♠♠♠ Taylor: [I need to add my notes here, This has to do with gluing together local solutions.

]

3.3 Riemann-Hilbert Correspondences

By *a* Riemann-Hilbert correspondence we will mean a theorem which gives some equivalence between a category of differential equations of some flavor with a category of representations of fundamental groups of some flavor. This may not even be a functor but just a bijection of particular sets of some property.

The general strategy for a functorial version is to have some category of connections or differential equations on some space X which we will call $\mathrm{Conn}(X)$ (e.g. holomorphic connections, Fuchsian connections, holomorphic systems, Fuchsian systesm, holonomic D-modules), a category of local systems $\mathrm{LocSys}(X)$ (essentially solutions of the differential equations, perverse sheaves), and a category of representations of the fundamental group $\mathrm{Repn}(X)$.

The idea is then to pass from differential equations to representations through "local systems".

$$\mathrm{Conn}(X) \cong \mathrm{LocSys}(X) \cong \mathrm{Repn}(X). \tag{3.3.1}$$

Before going any further we describe what a local system is. This conversation continues in section 3.6 which the reader should feel free to skip immediately to on the first reading.

3.4 Local Systems and Representations of $\pi_1(X)$

In this section we are going to introduce the notion of a local system and then prove that the category of local systems on X is equivalent to the category of finite dimensional representations of the fundamental group.

Along the way we are going to work out what happens with the holomorphic differential equations on $\mathbb{P}^1 \setminus S$ where S is a finite collection of points.

3.4.1 Local Systems

In what follows we will let X be a topological space. For a ring or abelian group K we will let $\underline{K}_X$ be the sheaf associated to the presheaf constant presheaf defined by

$$U \mapsto F(U) = \begin{cases} K, & U \neq \emptyset, \\ 0, & U = \emptyset \end{cases}$$

The presheaf is not a sheaf because it is possible for there to be two disjoint open sets U_1 and U_2 which means that any element $a \in F(U_1)$ and $b \in F(U_2)$ agree on their restriction but don't lift to a common element in $F(U_1 \cup U_2)$. This is fixed by allowing for elements "like" $a \oplus b \in \underline{K}_X(U_1 \cup U_2)$ so that the sheaf axioms are satisfied. On topological spaces (as opposed to general sites) then $\underline{K}_X(U) = \mathrm{Cont}(U, K)$, the collection of continuous map from U to K where K is given the discrete topology.

Definition 3.4.1.1. Let X be a topological space and let K be a field. A K-*local system* over X is a sheaf L valued in finite dimensional K-vector spaces such that $L \cong \underline{K}_X$ locally.

This forms a category and morphisms are morphism of sheaves of $\mathbb{C}$-linear vector spaces.

Exercise 3.4.1.2. Let X be a topological space and let V be an R-module for some ring R. Show that the constant sheaf $\underline{V}_X$ on X is the same thing as the sheaf

$$U \mapsto \mathrm{Cont}(U, V), \quad U \subset X \text{ open}$$

where Cont denotes continuous maps. Here V is given the discrete topology.

The above exercise is important. Let X be a topological space and let $f : X \to Y$ be a continuous map to a discrete topological space Y. Then $f^{-1}(\{y\})$ is open for every $y \in Y$. This holds for every $y \in Y$ and in particular the map $f : X \to Y$ is *locally constant*.

The main example of a local systems comes from solutions of differential equations. In fact this example is why the definition even exists. Consider the first order holomorphic system defined on $U \subset \mathbb{C}$, given by

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

For every $V \subset U$ open define

$$L(V) = \{Y \in \mathrm{Hol}(U)^{\oplus n} \colon Y' = A(t)Y\}.$$

We know that at each $t_0 \in U$ we have an isomorphism

$$\{Y \in \mathbb{C}\langle t - t_0 \rangle^{\oplus n} \colon Y' = AY\} \cong \mathbb{C}^n$$

since solutions form a finite dimernsional $\mathbb{C}$-vector space. By considering representatives of each basis element in $\mathbb{C}\langle t - t_0 \rangle^{\oplus n}$ we know that there exists some $V \subset U$ containing t_0 where

$$L(V) \cong \mathbb{C}^{\oplus n}.$$

This proves that L is a local system. Note in particular that

$$L(V) = \Phi_V \cdot \mathbb{C}^{\oplus n}$$

where Φ_V is a fundamental matrix valid on $V \subset U$.

3.4.2 The Espace Étale: Pullbacks and Pushforwards

I was tempted to add the following exercise without comment:

Exercise 3.4.2.1. The category $\mathrm{LocSys}(X)$ the category of local systems of finite dimensional $\mathbb{C}$-vector spaces makes sense and for every morphism of topological spaces $f : X \to Y$ one a functor $f^{-1} : \mathrm{LocSys}(X) \to \mathrm{LocSys}(Y)$.

I thought about it a second time and concluded this isn't very nice.

We are going to need to talk about inverse images of pullback of local systems so I want to say a couple words about inverse images of sheaves in general. Let $\mathcal{G}$ be a sheaf on a topological space Y. For a morphism of topological spaces $f : X \to Y$ we want to define the inverse image sheaf $f^{-1}\mathcal{G}$ (sometimes denoted $f^*\mathcal{F}$) which is the left adjoint of f_*. Here f_* is the direct image sheaf and it turns sheaves $\mathcal{F}$ on X into sheaves $f_*\mathcal{F}$. These direct image sheaves are super easy to describe: given $V \subset Y$ we have $(f_*\mathcal{F})(V) = \mathcal{F}(f^{-1}(V))$. That is it. The sheaves $f^{-1}\mathcal{G}$, are no so simple. Harshorne, for example, defines $f^{-1}\mathcal{G}$ to be the sheaf associated to the presheaf

$$U \mapsto \varinjlim_{V \supset f(U)} \mathcal{G}(U).$$

Yuck! I think this definition sucks and is hard to work with.

For every sheaf $\mathcal{F}$ on a topological space X, I'm going to introduce a topological space $\overline{\mathcal{F}}$ and a morphism of topological spaces $\pi : \overline{\mathcal{F}} \to X$ that is going to make our life easier. This space is called the *espace étale* and has the important property that sections of π over an open set U correspond to elements of $\mathcal{F}(U)$. By a section over an open set U we mean

$$\Gamma_X(\mathcal{F})(U) := \{s : U \to \overline{\mathcal{F}} \colon \pi s = \mathrm{id}_U\}.$$

That is $\Gamma_X(\mathcal{F}) \cong \mathcal{F}$ as sheaves. I should mention that in the description s is just a continuous morphism of topological spaces. The situation is pictured in Figure 3.4.2.

Definition 3.4.2.2. Given a sheaf $\mathcal{F}$ on a topological space X we define the *espace étale* of $\mathcal{F}$ to be the topological space $\overline{\mathcal{F}}$ whose underlying set is

$$\overline{\mathcal{F}} = \coprod_{x \in X} \mathcal{F}_x = \{(x,t) \colon t \in \mathcal{F}_x\},$$

and whose topology is generated in the following way: for each $s \in \mathcal{F}(U)$ for $U \subset X$ open declare

$$W(U,s) = \{(x, s_x) \colon s_x \text{ germ of } s \text{ at } x \in U\ \}$$

to be open and take the topology on $\overline{\mathcal{F}}$ to be the smallest topology so that the $W(U,s)$ are open for every $s \in \mathcal{F}(U)$.

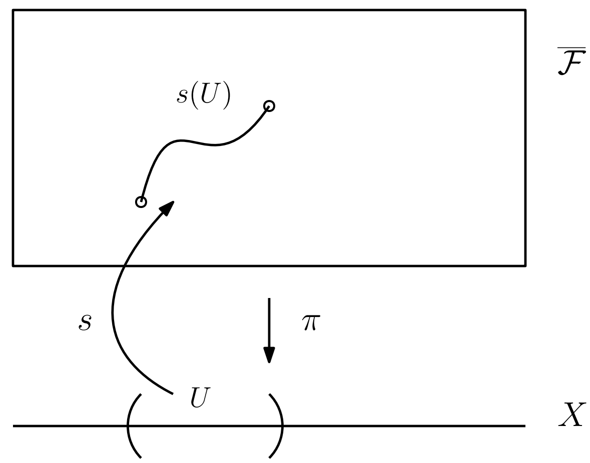

Figure 3.4: A picture of the espace étale for a sheaf $\mathcal{F}$. Pictures is a map s over an open set U such that $\pi s = \mathrm{id}_U$.

Exercise 3.4.2.3. Prove that $\Gamma_X(\mathcal{F})$ and $\mathcal{F}$ are isomorphism as sheaves.

We now turn to the point of introducing this construction. Given a morphism of topological spaces $f : X \to Y$ and a sheaf $\mathcal{G}$ on Y we define take define $f^{-1}\mathcal{G}$ to by its espace étale $\overline{f^{-1}\mathcal{G}}$ which is just the pullback of the $\overline{\mathcal{G}} \to Y$ to Y. That is

$$\overline{f^{-1}\mathcal{G}} = \overline{\mathcal{G}} \times_Y X,$$

where the fiber product is taken in the category of topological spaces. Similarly, $f_*\mathcal{F}$ for a sheaf $\mathcal{F}$ on X is the pushout

Exercise 3.4.2.4. Using this definition of $f^{-1}\mathcal{G}$ show that (f^{-1}, f_*) are an adjoint pair of functors.

3.4.3 Every Riemann Surface at Once

Using the espace étale construction in the case of the sheaf of holomorphic functions in the complex plane give something amazing. Here points are really (z_0, s_{z_0}) where $z \in \mathbb{C}$ and $s_{z_0} \in \mathbb{C}\langle z - z_0\rangle$ is a germ of a holomorphic function on z_0. This for example could be a germ of some branch of $\sqrt{z}$ at the points $z = 1$. Then this continues to all other branches of $\sqrt{z}$.

This branch determines a leaf or connected component of $\mathcal{X}$ where

$$\pi : \mathcal{X} = \overline{\mathcal{O}_{\mathbb{A}^{1,\mathrm{an}}}} \to \mathbb{C},$$

is the espace étale for the sheaf of holomorphic functions on $\mathbb{C}$.

The topology of the espace étale is crazy looking. For every germ of a holomorphic function over a point z_0 there is an element $\mathcal{X}$. So the fibers are uncountable. Also, for any element f of any fiber we can analytically continue germ around to get a connected component $\mathcal{X}(f) \subset \mathcal{X}$. This is actually a connected Riemann surface that maps to $\mathbb{C}$ and is the Riemann surface of the germ f. It may not be surjective onto $\mathbb{C}$ since not every germ admits an analytic continuation to every points of $\mathbb{C}$ (some "overconvergent" functions have a natural boundary) but it will give something.

So that is it. The espace étale $\mathcal{X} \to \mathbb{C}$ has as connected components all of the possible Riemann surfaces that come from analytic continuation of a germ, with many of them appearing with uncountable multiplicity.

3.4.4 Étale Morphisms, Coverings, and Locally Trivial Sheafs

The espace étale is very strange and to describe it geometrically we are going to recall the notions of étale and covering. The difference between these two notions is pictured in Figure 3.5. In particular we have drawn something a priori horrible (a posteriori not so horrible) corresponding to what we might thing an espace étale of a sheaf might look like.

Definition 3.4.4.1. In what follows we consider a morphism $f : Y \to X$ a morphism of topological spaces.

1. We say that f is *étale* if and only if

   (a) f is locally a homeomorphism.

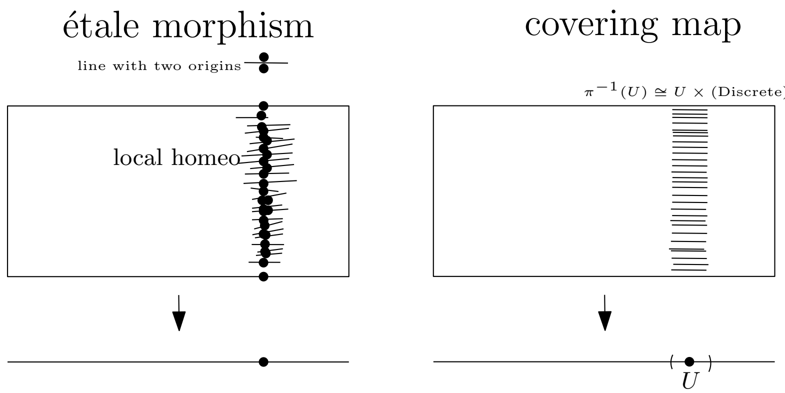

Figure 3.5: Both étale and covering morphisms are local homeomorphisms, it is just that étale morphisms can be weirder and include things in the fibers like lines with two origins which are not Hausdorff.

   (b) For all $x \in X$, $f^{-1}(x)$ has the discrete topology.

2. We say that f is a *covering map* if and only if it is étale and locally trivial.

Covering spaces are just like the covering spaces you have seen from Hatcher but perhaps the fibers are larger (and too large of fibers indicates in many cases that there needs to be multiple connected components). Étale morphisms are like coverings by includes things like the an arbitrary disjoint union of any collection of open subsets of X; the fibers don't need to have any particular uniform size or anything like that. Also there can be issues with separatedness where you can have things like the line with two origins appearing above an open set.

We now show the espace étale is indeed gives an étale morphism. The theorem below is actually a good exercise and you should try it.

Theorem 3.4.4.2. *If F is a sheaf of sets on a topological space X then the morphism $\pi : \overline{F} \to X$ from the espace étale of F to X is an étale morphism.*

Proof. The morphism π is given by $(x, s_x) \mapsto x$. By definition the $F_x = \varinjlim_{U \ni x} F(U)$ we know there exists some U such that $s \in F(U)$ restricts to

$s_x \in F(U)$. By definition of $\overline{F}$ we have $\{(x, s_x) \colon x \in U, s_x = (s)_x\}$ mapping homoemorphically onto U. We are using $(s)_x$ to denote taking the stalk of $s \in F(U)$ at x.

Given two points (x, t_x) and (x, s_x) corresponding sets above separate these points giving the fiber above x the discrete topology. □

If F is a sheaf of sets on X then $\pi : \overline{F} \to X$ is an étale morphism yet $\pi^{-1}(x) = F_x$. In general étale morphisms and covering morphisms can seem awful. This is horribly large. Here F_x is the stalk at the point. In the case when $F = \mathcal{O}_X$ where X is a complex manifold then $\mathcal{O}_{X,x}$ is like a ring of power series which is uncountable, so the fiber is an uncountable set in the discrete topology.

The following is Sabbah exercise 15.1 + a corollary around there. One needs to recall the Fundamental Theorem of Galois Theory for Covering Spaces: Given a connected space X, every subgroup of $\pi_1(X)$ comes from a connected covering of X Conversely given a covering, its connected components correspond to subgroups of $\pi_1(X)$. The construction is to quotient the universal cover by the group $\Pi \subset \pi_1(X)$ which acts discretely.

Theorem 3.4.4.3. *Let X be a connected topological space. Let F be a sheaf of sets on X.*

1. *The sheaf F is constant if and only if $\pi : \overline{F} \to X$ is trivial in the sense that $\overline{F} \cong X \times F_0$ for some fixed F_0 with the discrete topology.*
2. *The sheaf F is locally constant if and only if $\pi : \overline{F} \to X$ is a covering.*
3. *If X is simply connected then any locally constant sheaf is constant.*

Proof. Suppose that F is constant with $F(U) = F_0$ for all U nonempty. Then for all x and all y we have $F_x = F_y = F_0$. Then $\overline{F} \cong X \times F_0$. The map is given by $(x, s_x) \mapsto (x, s_x)$ since there is a canonical isomorphism between F_x and F_0.

Conversely, take sections and use the property that continuous maps to discrete things are constant (Exercise 3.4.1.2).

The second part follows from the first as we restrict to an open set. Locally constant implies locally trivial.

We now use the remarks proceeding the statement of the theorem about fundamental groups. If X is simply connected then $\pi_1(X) = 1$. This means that there are no non-trivial coverings. This means that every sheat is isomorphic to X and hence $\overline{F} \cong \coprod_{d\in D} X = X \times D$ where D is some space with the discrete topology. This is the constant sheaf of sets that we are after. $\square$

We now use this to disentangle the relationship beween (E, ∇) and E^{∇}.

3.4.5 Local Systems vs Vector Bundles

In this section we discuss the difference between a local system associated to some integrable (E, ∇) and the vector bundles E itself. In what follows we will let $L = E^{\nabla}$ and try to compare $\overline{L}$ and the physical vector bundle E.

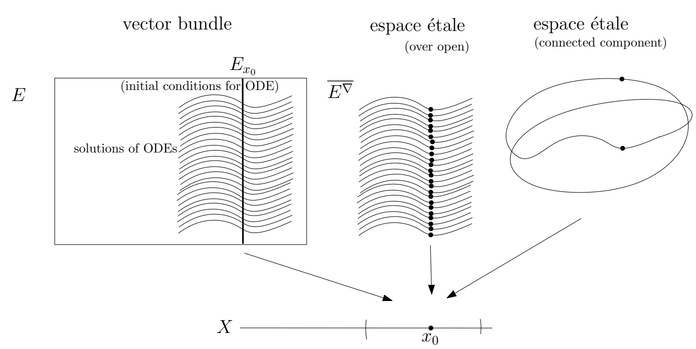

Figure 3.6: The shredding of the vector bundle E by the foliation, giving us the espace étale.

Since there is sometimes some confusion in notation I want to review the difference between stalks of locally free sheafs and fibers of vector bundles. As stated previous we often conflate $\mathcal{E}$ the sheaf of sections of a vector bundle E with E itself and just write E in place of $\mathcal{E}$. Technically, E is a geometric object with a morphism $f : E \to X$ a morphisms of spaces (schemes, complex manifolds, whatever). Also, technically, $\mathcal{E}$ is a locally free sheaf of $\mathcal{O}_X$-modules. These two things live in different categories and E "represents" $\mathcal{E}$.

With this conflation in mind there are two possible meanings for E_x given $x \in X$ that can be conflated. It can mean either the stalk

$$E_x = \varinjlim_{U \ni x} \mathcal{E}(U)$$

or the fiber

$$E_x = f^{-1}(x).$$

In the case it is the stalk we are thinking of E as a locally free sheaf and in the case of the fiber we are thinking about $E \to X$ as a physical vector bundle which is a space. The stalk description is huge and is the module over a local ring $\mathcal{O}_{X,x}$, which when working with complex manifolds is a ring of power series so an infinite dimensional vector space. In the case of a fiber E_x is a finite dimensional complex vector space. These are not the same. Not even close. One needs to reduce the stalk modulo the maximal ideal of $\mathcal{O}_{X,x}$ to get a comparison.

Now for a locally free sheaf, things start to get a little closer. The picture of the discussion that follows is Figure 3.6. In this discussion we restrict our attention to X being a complex manifold. If E a vector bundle of rank n (which we think of as a physical vector bundle) then E_x is a vector space of rank n. For the espace étale $\pi : \overline{L} \to X$ of a local system L we have $L_x = \overline{L}_x$, the fiber is the stalk. Also $L_x \cong \mathbb{C}^n$ so we have a vector space of rank n as the fiber of the espace étale. The big difference here is the topology! The topology of E_x is that of usual topological vector space. The topology of $\overline{L}_x$ discrete. What have we done? Well, the local system L parametrizes local solutions of of our differential equation. They are a linear combination of a basis of solutions. The connection defined a flow on the space E and the leaves of this flow are the solutions of the differential equation. What $\overline{L}$ is (at least locally), is the shredding of E by these leaves (again see Figure 3.6). So E has sort-of been ripped apart and reassembled initial condition by initial condition to give uncountably many fibers sitting discretely together.

3.5 Monodromy Representations and Local Systems

Let X be a topological space which is connected and locally path connected so that $\pi_1(X, x_0)$ makes sense. Let L be a local system on X and let $\gamma : [0,1] \to X$ be a continuous path contained in X. Then $\gamma^* L$ is a local system on $[0,1]$ which is trivial. In seeing this, Exercise 3.4.1.2 is important. Furthermore every germ $v \in (\pi^* L)_0$ extends uniquely to $(\gamma^* L)([0,1])$ via monodromy. In particular there is a morphism

$$M_\gamma : (\gamma^* L)_0 \to (\gamma^* L)_1.$$

Since $(\gamma^* L)_t \cong L_{\gamma(t)}$ for each t if γ is a closed path starting and ending at $x_0 \in X$ then $(\gamma^* L)_0 \cong L_{\gamma(0)} \cong L_{x_0} \cong L_{\gamma(1)} \cong (\gamma^* L)_1$. This then means that M_γ induces an automorphism of L_{x_0}. This defines the representation associated to the pair (L, x_0) consisting of a local system L and a point $x_0 \in X$

Definition 3.5.0.1. The *monodromy representation* associated to the (L, x_0) is

$$\rho_{(L,x_0)} : \pi_1(X, x_0) \to \mathrm{Aut}(L_{x_0}), \quad \gamma \mapsto M_\gamma,$$

as described above.

We now want to give a comparison between the representation above and the monodromy representation we had previously discussed in dimension one.

Recall that the category of K-representations of a group Π is the category of $K[\Pi]$-modules and when we talk about two representations being isomorphic we talk about them being isomorphic as $K[\Pi]$-modules. We are interested in the case when $\Pi = \pi_1(U, t_0)$ for some $U \subset \mathbb{C}$.

Theorem 3.5.0.2. *Consider a holomorphic system on $U \subset \mathbb{C}$ of the form*

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

Let L_A be the local system associated to this system. Let ρ_A be the monodromy representation associated to the fundamental matrix $\Phi(t)$ satisfying $\Phi(t_0) = I_n$. Then $\rho_{L_A,t_0} \cong \rho_A$ as reprentations of $\pi_1(U, t_0)$.

Proof. Our aim is to compute ρ_{L_A} (we will drop the base point t_0 from the notation for convenience). At a point $t_1 \in U$ along a curve γ_1 starting at t_0 and ending at t_1 we have

$$\Phi_{\gamma_1}\mathbb{C}^n \cong L_{A,t_1},$$

where $\Phi_\gamma \in \mathrm{GL}_n(\mathbb{C}\langle t - t_1\rangle)$ is the local fundamental matrix. The representation associated to the local system gives

$$\begin{tikzcd}
\mathbb{C}^n & \Phi\mathbb{C}^n \arrow[l, "\Phi^{-1}"'] \arrow[r] \arrow[d, equal] & \Phi_\gamma\mathbb{C}^n \arrow[r, "\Phi_\gamma^{-1}"] \arrow[d, equal] & \mathbb{C}^n \\
 & (\gamma^* L_A)_0 & (\gamma^* L_A)_1 &
\end{tikzcd}$$

Here $v = \Phi c \mapsto \Phi_\gamma c = \Phi M_\gamma c$ implies that $\Phi M_\gamma \Phi^{-1}$ is the action on the stalk. In triviallized coordinates we have

$$v \mapsto \Phi v \mapsto \Phi M_\gamma \Phi^{-1}\Phi v \mapsto \Phi_\gamma^{-1} M_\gamma v \mapsto M_\gamma^{-1}\Phi M_\gamma v.$$

$\square$

In higher dimensions this is how we are going to define the representation associated to a connection. That is to every vector bundle E with integrable connection $\nabla^2 = 0$ we have the associated local system E^∇ defined by

$$E^\nabla(U) = \{s \in E(U) \colon \nabla(s) = 0\}, \quad U \subset X \text{ open }.$$

The very definition of integrability is exactly so that the system of equations $\nabla(s) = 0$ has a local basis of solutions. More precisely if $t^1, \dots, t^m$ are local coordinates for X and $s_1, \dots, s_n$ are local basis for E then

$$\nabla(s) = 0 \quad \iff \quad \nabla_{\partial_j}(Y) = \frac{\partial Y}{\partial t^j} + A_j Y = 0, \quad 1 \le j \le m,$$

where $Y \in \mathcal{O}_X(U)^{\oplus n}$ is a presentation of $s \in E(U)$ under the trivialization given by the local basis and $A_j dt^j$ is the connection 1-form in local coordinates.

Definition 3.5.0.3. Let (E, ∇) module be a module with integrable connection. The construction $(E, \nabla) \mapsto E^\nabla$ is *local system associated to* (E, ∇)

Definition 3.5.0.4. Let (E, ∇) be an integrable connection on a complex manifold X. We define the *monodromy representation* of (E, ∇) to be the monodromy representation of E^{∇}.

This describes the functors

$$\mathrm{Conn}(X) \to \mathrm{LocSys}(X) \to \mathrm{Repn}(X).$$

The connection, goes to a local set of solutions (a local system), which by § 3.5 gives rise to a representation. Spoiler: for holomorphic vector bundles this will be an equivalence of categories. The map $\mathrm{LocSys}(X) \to \mathrm{Repn}(X)$ is always going to be an equivalence of categories.

3.5.1 Local Systems and Representations are Equivalent

In this section we show that for a general topological space the category of representations is equivalent to the category of local systems.

First some notation: let Π be a group. We will let $\mathsf{Mod}^{\mathrm{fin}}_{\mathbb{C}[\Pi]}$ denote the category of $\mathbb{C}[\Pi]$-modules which are finite dimensional as $\mathbb{C}$-vector spaces. We will often conflate a representation $\rho : \Pi \to \mathrm{GL}(V)$ with its underlying $\mathbb{C}[\Pi]$-module, which we also denote by V.

Now we give our theorem.

Theorem 3.5.1.1. *Let X be a topological space which admits a fundamental group and let $\Pi = \pi_1(X, x_0)$ for $x_0 \in \mathbb{C}$. The category of local systems on X is equivalent to the category of finite dimensional complex representations of the fundamental group*

$$\mathrm{LocSys}(X) \xrightarrow{\sim} \mathrm{Repn}(X), \quad L \mapsto \rho_{L,x_0},$$

Here $\mathrm{Repn}(X) = \mathsf{Mod}^{\mathrm{fin}}_{\mathbb{C}[\Pi]}$ *of finite dimensional complex representations of the fundamental group (=* $\mathbb{C}[\Pi]$*-modules).*

Proof Sketch. The quasi-inverse is given by the so-called suspension construction. From representaions to local systems is the map

$$V \mapsto \widetilde{X} \times_{\Pi} V = \overline{L},$$

which gives the espace étale of the local system L. Here, $\Pi = \pi_1(X, x_0)$, V is the underlying vector space of the Π-representation (it must be given the discrete topology), $\widetilde{X} \to X$ is the universal cover, and $\times_{\Pi}$ denotes the amalgamated product.

We give some more details. Let $\rho : \Pi \to \mathrm{GL}(V)$ be a finite dimensional complex representation of Π. Note that we have an action of Π on both $\widetilde{X}$ and on V and hence on the complex manifold $\widetilde{X} \times V$ where the action is given by

$$(\gamma, (\widetilde{x}, v)) \to (\gamma(\widetilde{x}), \rho(\gamma)(v)).$$

Note that since the action by deck transformations preserves fibers of f we have that Π also acts on $f^{-1}(U) \times V$ for every open subset $U \subset X$.

Define the constant sheaf $\widetilde{L} = \underline{V}_{\widetilde{X}}$. We claim the total space of this constant sheaf is $\widetilde{X} \times V$ provided we give V the discrete topology; that is, for every $\widetilde{U} \subset \widetilde{X}$ there is a bijection (see exercise 3.5.1.2)

$$\widetilde{L}(\widetilde{U}) \cong \{\text{sections of } \pi_1 : \widetilde{X} \times V \to \widetilde{X} \text{ above } \widetilde{U}\ \}. \tag{3.5.1}$$

Hence there is an action of Π on $(f_*\widetilde{L})(U)$ for every U subset X and it makes sense to define $L(U)$ by the formula

$$L(U) = (f_*\widetilde{L})(U)^{\Pi}.$$

We claim that this is the local system with the corresponding monodromy represenation. $\square$

Exercise 3.5.1.2. In this exercise M will be a topological space where fundamental groups make sense.[3]. We will let $\Pi = \pi_1(M, x_0)$ for $x_0 \in M$ some base point. Check (3.5.1) regarding the description of the total space of the local system.

[3] path connected, locally path connected

3.6 Holomorphic Riemann Hilbert Correspondence

Let X be a complex manifold. We have three categories.

$$\begin{aligned} \mathrm{Conn}(X) &= (\text{holomorphic vector bundles on } X \text{ with integrable connections}) \\ \mathrm{LocSys}(X) &= (\text{local systems on } X) \\ \mathrm{Repn}(X) &= (\text{finite dimensional complex representations of } \pi_1(X, x_0)\ ) \end{aligned}$$

The holomorphic Riemann-Hilbert correspondence states that these categories are all equivalent. To estabilish this we will show $\mathrm{Conn}(X) \cong \mathrm{LocSys}(X)$ and that $\mathrm{LocSys}(X) \cong \mathrm{Repn}(X)$. In fact this section part was already established for general topological spaces in §3.5.1.

The functor

$$\mathrm{Conn}(X) \to \mathrm{LocSys}(X), \quad (E, \nabla) \mapsto E^{\nabla}$$

was already described. Here E^{∇} is "the space of horizontal sections". It doesn't hurt to repeat that

$$E^{\nabla}(U) = \{s \in E(U) \colon \nabla(s) = 0\}, \quad U \subset X \text{ open }.$$

The key condition here is integrability, which, *by definition* is so that we have equatlity of mixed partials in the system of equations

$$\nabla(s) = 0 \quad \iff \quad \nabla_{\partial_j}(Y) = \frac{\partial Y}{\partial t^j} + A_j Y = 0, \quad 1 \leq j \leq m,$$

where $t^1, \ldots, t^m$ are local coordinates for X and $s_1, \ldots, s_n$ are a local basis for E giving rise to $s \mapsto Y$.

The converse construction is as follows.

$$\mathrm{LocSys}(X) \to \mathrm{Conn}(X), \quad L \mapsto (\mathcal{O}_X \otimes_{\mathbb{C}_X} L, \nabla_L),$$

where $\nabla_L(f \otimes v) = df \otimes v$ for $f \otimes v \in \mathcal{O}_X \otimes L$, and then the definition is given by extending linearly.

Readers can check as much detail as they want that this construction is a quasi-inverse of $(E, \nabla) \mapsto E^{\nabla}$. The main idea here is that a basis of solutions always gives a matrix Φ which makes the connection equivalent to the trivial connection (see exercise 2.3.2.2).

One of the major points here is that a connection is determined by its space of horizontal solutions. This allows us to define connections by declaring what their horizontal space is. When studying connections in characteristic p on schemes we can define a canonical connection of the pullback of a vector bundle by the Frobenius ∇^{can} by declaring the inverse image under the Frobenius to be horizontal.

Theorem 3.6.0.1 (Holomorphic Riemann-Hilbert). *We have the following equivalences of categories for X a complex manifold:*

$$\mathrm{Conn}(X) \cong \mathrm{LocSys}(X) \cong \mathrm{Repn}(X).$$

Later we will modify the category $\mathrm{Conn}(X)$ to get more information about the differential equations at the singularities.

3.7 Families of Differential Equations and Local Systems

Relative connections are the way we formalize families of differential equations depending on a parameters. This become relevant in section §4.1 when we first start to discuss Schlesinger's equations and isomonodromy. Here we seek a family of differential equations which when you deform a parameter you retain the same monodromy.

3.7.1 Relative Connections

Recall that given a morphism $\pi : X \to S$ we think of the fibers over $s \in S$ which we denote as X_s as a family. We will even call a morphism $\pi : X \to S$ a family. One way of thinking about relative connections are as families of differential equations (or more generally connections). Given morphism $\pi : X \to S$ of complex manifolds or schemes one can define a relative connection on a vector bundle E on S we make the following definition.

Definition 3.7.1.1. A *relative connection* $\nabla_{X/S}$ on E is an additive map

$$\nabla_{X/S} : E \to \Omega_{X/S} \otimes E$$

such that $\nabla_{X/S}(fs) = df \otimes s + f\nabla_{X/S}(s)$ for all $f \in \mathcal{O}_X$ and all $s \in E$. We say that the connection is *integrable* if $\nabla^2_{X/S} = 0$.

The fact that $df = 0$ for $f \in \pi^{-1}\mathcal{O}_S$ says that relative connections are $\pi^{-1}\mathcal{O}_S$-linear. Note that this makes sense: because we have a map $\mathcal{O}_S \to \pi_*\mathcal{O}_X$ there is an adjoint map $\pi^{-1}\mathcal{O}_S \to \mathcal{O}_X$ and it makes sense to talk about elements of $\pi^{-1}\mathcal{O}_S$ as elements of $\mathcal{O}_X$ (we are thinking about them via their image under the morphism just described).

Morphisms of relative connections are morphisms of vector bundles which are equivariant with respect to their connection. We denote the category of integrable relative connections for $\pi : X \to S$ by $\mathrm{Conn}(X/S)$.

Exercise 3.7.1.2. Given a relative connection check that $(E, \nabla)|_{X_s}$ makes sense as a module with integrable connection on X_s.

3.7.2 Relative Local Systems

Again we have a local system associated to an integrable relative connection

$$(E, \nabla_{X/S}) \to E^{\nabla_{X/S}}$$

where for $U \subset X$ open we have

$$E^{\nabla_{X/S}}(U) = \{s \in E(U) \colon \nabla_{X/S}(s) = 0\}.$$

Given a family $\pi : X \to S$ we make the following definition.

Definition 3.7.2.1. A *local system relative to* π (of rank n) is a sheaf L of $\pi^{-1}\mathcal{O}_S$-modules on X such that for all $x \in X$ there exists a $U \ni x$ open with

$$L|_U \cong (\pi^{-1}\mathcal{O}_S)^{\oplus n}|_U$$

as sheaves of $\pi^{-1}\mathcal{O}_S|_U$-modules.

The story is pretty much the same in the relative setting as it was in the absolute setting. We define morphisms of relative local systems as morphisms of $\pi^{-1}\mathcal{O}_S$-modules and denote the category of relative local systems for a morphism $\pi : X \to S$ by $\mathrm{LocSys}(X/S)$.

For the following exercises $\pi : X \to S$ will be a morphism of complex manifolds.

Exercise 3.7.2.2. Given an vector bundle E on X with a relative connection $\nabla_{X/S}$ which is integrable check that $E^{\nabla_{X/S}}$ is a relative local system.

Exercise 3.7.2.3. Check that relative local systems restrict to local systems on fibers.

Exercise 3.7.2.4. Check that relative connections restrict to connections on the fibers.

3.7.3 Relative Riemann-Hilbert Correspondence

From relative local systems to relative integrable connections we use

$$L \mapsto E_L = \mathcal{O}_X \otimes_{\pi^{-1}\mathcal{O}_S} L$$

and gives it the connection by extending $\pi^{-1}\mathcal{O}_S$-linearly the map

$$f \otimes v \mapsto df \otimes v.$$

Exercise 3.7.3.1. Check that the map

$$\mathrm{LocSys}(X/S) \to \mathrm{Conn}(X/S), \quad L \mapsto E_L$$

is an equivalence of categories with quasi-inverse $(E, \nabla_{X/S}) \mapsto E^{\nabla_{X/S}}$.

Chapter 4

Isomonodromic Deformations

The Painlevé VI equation takes the form

$$\frac{d^2q}{dx^2} = \frac{1}{2}\left(\frac{1}{q} + \frac{1}{q-1} + \frac{1}{q-x}\right)\left(\frac{dq}{dx}\right)^2 - \left(\frac{1}{x} + \frac{1}{x-1} + \frac{1}{q-x}\right)\frac{dq}{dx} + \frac{q(q-1)(q-x)}{x^2(x-1)^2}\left(\alpha + \beta\frac{x}{q^2} + \gamma\frac{x-1}{(q-1)^2} + \delta\frac{x(x-1)}{(q-x)^2}\right),$$

where $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ are constants. The aim of this section is to show that this equation is really about isomonodromic deformations of rank two Fuchsian equations with trace free entries on $\mathbb{P}^1 \setminus \{0, 1, \infty, x\}$ where x is a variable point.

4.1 Schlesinger's Equations

Consider the case of a Fuchsian differential equation of rank 2 on $\mathbb{P}^1$ with singular points at $\{0, 1, \infty, x\}$ for some variable $x \in \mathbb{P}^1$ not equal to $0, 1$ or ∞,

$$\begin{cases} \dfrac{dY}{dt} = A(t,x)Y \\ A(t,x) = \dfrac{A_0(x)}{t} + \dfrac{A_1(x)}{t-1} + \dfrac{A_2(x)}{t-x} \end{cases}$$

and define $A_\infty(x)$ by the equation $A_0+A_1+A_2+A_\infty = 0$. The matrices $A_j(x)$ we will assume depend holomorphically on the variable x in some unspecified domain (we think of x as varying a little bit around some x_0). We will often write $A_j = A_j(x)$ for short. We note that for a fixed $x = x_0$ the Fuchsian differential equation gives a monodromy representation. Also, as we vary $x \in \mathbb{P}^1\{0, 1, \infty\}$ the fundamental group does not change (see Figure 4.1). In

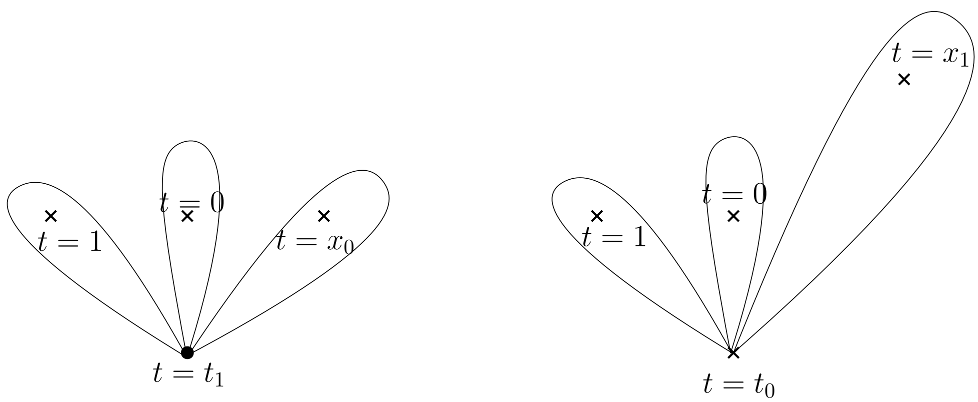

Figure 4.1: Varying x in $\mathbb{P}^1 \setminus \{0, 1, \infty, x\}$ does not change the fundamental group.

what follows we will let

$$\Pi = \langle \gamma_0, \gamma_1, \gamma_2 \colon \gamma_0\gamma_1\gamma_2 = 1\rangle.$$

For x_0 and x_1 in $\mathbb{P}^1$ with $x_0 \neq x_1$ we have

$$\Pi \cong \pi_1(\mathbb{P}^1 \setminus \{0, 1, \infty, x_0\}, t_0) \cong \pi_1(\mathbb{P}^1 \setminus \{0, 1, \infty, x_1\}, t_0).$$

Now fix $x = x_0$ and a fundamental matrix $\Phi_0(t)$ for our system at $x = x_0$ so that $\Phi_0'(t) = A(t, x_0)\Phi_0(t)$. This gives a monodromy representation

$$\rho_0 \colon \Pi \to \mathrm{GL}_2(\mathbb{C}), \quad \gamma \mapsto M_\gamma,$$

where $M_\gamma \in \mathrm{GL}_2(\mathbb{C})$ is the matrix so that $(\Phi_0(t))_\gamma = \Phi_0(t)M_\gamma$.

Problem 4.1.0.1 (Isomonodromic Deformations)**.** Find a parameter space X and a function

$$X \mapsto M_2(\mathbb{C})^3, \quad x \mapsto (A_0(x), A_1(x), A_2(x))$$

so that the fundamental matrix $\Phi(x,t)$ of the system

$$\frac{dY}{dt}(x,t) = \left(\frac{A_0(x)}{t} + \frac{A_1(x)}{t-1} + \frac{A_2(x)}{t-x}\right) Y(x,t)$$

satisfies

1. (Monodromy at x_0) $\Phi(t,x_0) = \Phi_0(t)$ (so that M_γ define the monodromy representation at that point).
2. (Isomonodromy) For each $x \in X$, and all $\gamma \in \Pi$,

$$\Phi(t,x)_\gamma = \Phi(t,x) M_\gamma,$$

   up to conjugation of the collection of M_γ.

If a function $x \to A(x)$ satisfies the isomonodromy problem we call it an *isomonodromic deformation* of the system at $x = x_0$.

4.1.1 Schlesinger's Theorem

In this section we are going to give some equations that determine when

$$x \mapsto A(x,t) = \frac{A_0(x)}{t} + \frac{A_1(x)}{t-1} + \frac{A_2(x)}{t-x}. \tag{4.1.1}$$

gives an isomonodromic deformation. We will make some simplifying assumptions.

1. Assume that $A_\infty(x)$ is a constant diagonal matrix.
2. (Non-resonance) Assume that for each x the eigenvalues of $A_j(x)$ for $j = 0, 1, \infty$ do not differ by an integer.

Under the conditions above we can give equations. Both of these conditations are used in the secret weapon of this theorem: the logarithmic Riemann-Hilbert correspondence.

Theorem 4.1.1.1. *Assume the assumptions and work in the notation of this section. The map $x \mapsto A(x)$ is isomonodromic if and only if*

$$\frac{\partial A_0}{\partial x} = \frac{[A_0, A_2]}{x}, \quad \frac{\partial A_1}{\partial x} = \frac{[A_1, A_2]}{x-1}, \quad \frac{\partial A_\infty}{\partial x} = 0.$$

The above Theorem is a special case of a more general formula we will give. It turns out these equations are equivalent to integrability conditions for a connection on a certain space. To state the connection to connections we will change our setup slightly. We will work with a rank n system on $\mathbb{P}^1$ and allow it to have singularities at m-points (and for simplicity we will exclusing ∞) we will let X be a parameter space of points

$$X = \{(x_1, \ldots, x_m) \in \mathbb{C}^m \colon x_i \neq x_j, \text{ for } i \neq j\}.$$

We will use the notation $x = (x_1, \ldots, x_m)$

Theorem 4.1.1.2 (Schlesinger's Equations)**.** *Consider the system,*

$$\frac{\partial Y}{\partial t}(t, x) = \sum_{j=1}^{m} \frac{A_j(x)}{t - x_j} Y(t, x)$$

on $\mathbb{P}^1 \times X$. This system is isomonodromic if and only if

$$\begin{cases} \dfrac{\partial A_j}{\partial x_i} = \dfrac{[A_j, A_i]}{x_j - x_i}, & i \neq j, \\ \dfrac{\partial A_j}{\partial x_j} = -\sum_{i \neq j} \dfrac{[A_i, A_j]}{x_i - x_j}. & \end{cases} \tag{4.1.2}$$

In these notes we will call (4.1.2) *Schlesinger's equations.* The commutators in the Schlesinger's equation should make the reader suspect that these equations are possibly an integrability/curvature condition on some connection. This is indeed the case and we will actually have a "logarithmic connection" on the space $X \times \mathbb{P}^1$.

Theorem 4.1.1.3. *The following are equivalent for a collection of matrix valued functions $A_j(x)$ on X.*

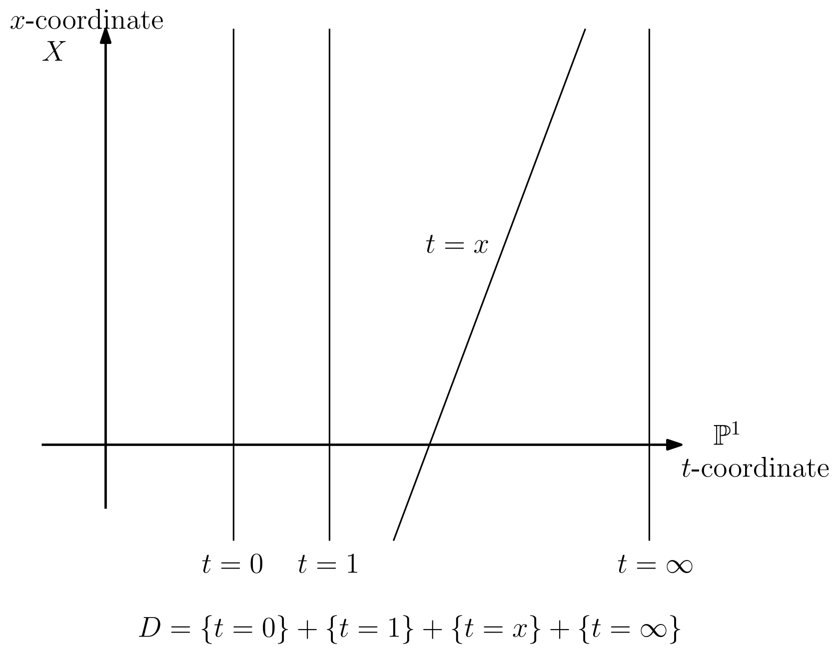

Figure 4.2: A picture of the $X \times \mathbb{P}^1$ and the associated divisor D in the case of two points and a single singular point x varying.

1. *The connection* $\nabla = d + \omega$ *on* $\mathcal{O}_{\mathbb{P}^1 \times X}$ *given by*

$$\omega = \sum_{j=1}^{m} A_j(x) \frac{d(t - x_j)}{t - x_j},$$

   is integrable.

2. *The matrices* A_j *satisfy*

$$dA_j - \sum_{i \neq j} [A_i, A_j] \frac{d(x_i - x_j)}{x_i - x_j} = 0, \quad 1 \leq j \leq m.$$

3. *Schlesinger's equations* (4.1.2) *hold.*

Proof. Also (2) are equivalent to Schlesinger's equations.

The computation of the equivalent between (2) and the integrability condition can be rough if you try to do too much or move to Einstein notation. Also, we should observe that Schlesinger's equation are the literal equations one obtains from integrability but can be derived from the integrability conditions.

The key is to not manipulate your equations too much and think about what you are doing. Integrability of the connection is equivalent to (see Theorem 3.1.8.6)

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

We compute.

$$d(\omega) = \sum_{i=1}^{m} \frac{dA_i}{t - x_i} \wedge dt - \sum_{i=1}^{m} \frac{dA_i \wedge dx_i}{t - x_i},$$

$$\omega \wedge \omega = \sum_{i,j} \frac{[A_i, A_j]}{t - x_j} \frac{dt \wedge d(x_i - x_j) + dx_i \wedge dx_j}{x_i - x_j}.$$

Collecting the terms with dt in them one finds

$$\sum_i \frac{dA_i}{t - x_i} = \sum_{i,j} \frac{[A_i, A_j]}{t - x_i} \frac{d(x_i - x_j)}{x_i - x_j} \tag{4.1.3}$$

The remaining terms give

$$\frac{dA_i \wedge dx_i}{(t - x_i)} = \sum_{i,j} \frac{[A_i, A_j]}{t - x_i} \frac{dx_i \wedge dx_j}{x_i - x_j}. \tag{4.1.4}$$

Taking residues of (4.1.3) along $t = x_k$ we find that

$$dA_k = \sum_{j \neq k} [A_k, A_j] \frac{d(x_k - x_j)}{x_k - x_j} = 0.$$

This proves (2).

To see the converse observe that multiplying (2) by $\frac{1}{t - x_j}$ and summing over j gives (4.1.3). To recover (4.1.4) we apply $- \wedge \dfrac{dx_i}{t - x_i}$ to (2) and sum over i. □

One can be a little more general with the points if one likes. The following is a description of

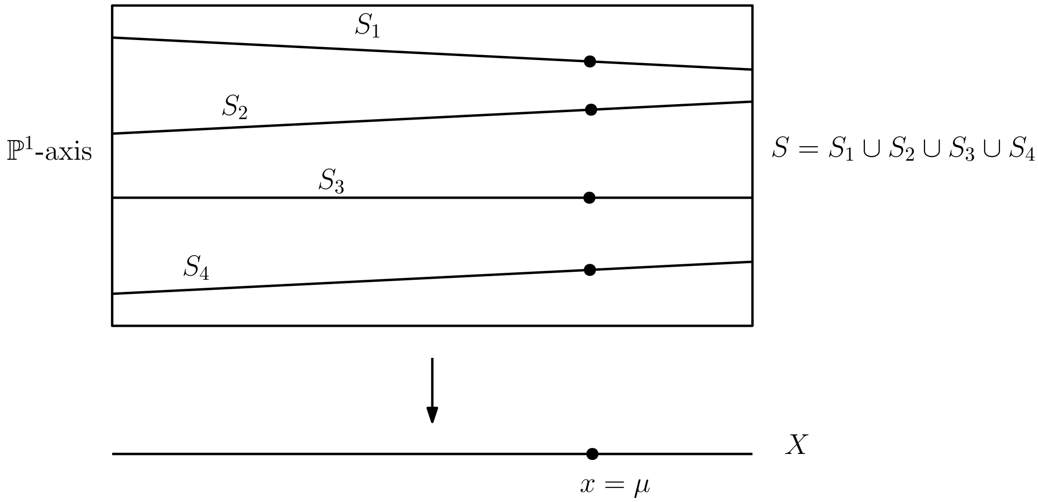

Figure 4.3: A picture of $S \subset \mathbb{P}^1 \times X$ in the case of four parametrized points.

4.1.2 The Geometry of the Connection

♠♠♠ Taylor: [LOTS OF NOTES NEED TO BE PUT IN HERE. They are mostly from Sabbah, with some parts from Deligne and Malgrange for the Riemann Hilbert Correspondence stuff.]

In [Sab07, VI, §1, pg 192] Sabbah works with an arbitrary complex manifold X some $S \subset X \times \mathbb{P}^1$ such that the map $f : S \to X$ induced by the projection onto X has degree m. This means that over $\mu \in X$ the we have a fiber S_μ

$$S_\mu = f^{-1}(\mu) = \{(\mu, x_1(\mu))(\mu, x_2(\mu)), \dots, (\mu, x_m(\mu))\}.$$

By assumptions on X (1-connectedness) the manifold S is trivializable: there exists maps $x_j : X \to \mathbb{P}^1$ such that S is the union of the graphs of these functions:

$$S = S_1 \cup \cdots \cup S_m$$

where S_j is the graph of the map x_j. In what we were doing above the maps x_j were the projections from $X \subset (\mathbb{P}^1)^m$ onto $\mathbb{P}^1$.

Note now that the fiber of $(X \times \mathbb{P}^1) \setminus S$ above $x = \mu \in X$ is $\mathbb{P}^1 \setminus S_\mu \cong (\mathbb{P}^1 \setminus S_\mu) \times \{\mu\}$. If we assume that $H_2(X, \mathbb{Z}) = 0$ or that $\pi_2(X) = 0$ then one can use the long exact sequence in homotopy groups to conclude that the inclusion of the fiber

$$\mathbb{P}^1 \setminus S_\mu \hookrightarrow (\mathbb{P}^1 \times X) \setminus S$$

induces an isomorphism of fundamental groups.

4.2 Representation and Character Varieties

To collection of representations of a fundamental group can be given the structure of a variety. This variety is the representation variety. To get representations up to isomorphism, we need to mod out by conjugation. After doing this we get the Character variety.

4.2.1 Representation Varieties

The term "representation variety" is non-standard, I think. Let Π be a finitely presented group. This means that

$$\Pi = \langle \gamma_1, \ldots, \gamma_\ell \colon R_1, \ldots, R_s \rangle$$

where the R_j are relations of the form

$$\gamma_{i_1}^{a_1} \gamma_{i_2}^{a_2} \cdots \gamma_{i_k}^{a_k} = 1,$$

for some $k \geq 0$, $\{i_1, \ldots, i_k\} \subset \{1, \ldots, \ell\}$ and $a_1, \ldots, a_k \in \mathbb{Z}$. A representation of $\Pi \to \mathrm{GL}_n(\mathbb{C})$ is determined by $\gamma_i \mapsto M_i$ where $M_i \in \mathrm{GL}_n(\mathbb{C})$ satisfy relations of the type

$$M_{i_1}^{a_1} M_{i_2}^{a_2} \cdots M_{i_k}^{a_k} = I_n$$

where I_n is the $n \times n$ identity matrix. Viewing the entries of M_i as variables we find get an algebraic variety $\mathrm{Repn}_{\mathrm{GL}_n}(\Pi)(\mathbb{C})$ whose points are in bijection with representations $\rho : \Pi \to \mathrm{GL}_n(\mathbb{C})$.

Note that the equations are actually defined over $\mathbb{Z}$ and the construction is quite general so $\mathrm{Repn}_{\mathrm{GL}_n}(\Pi)(\mathbb{C})$ is actually the $\mathbb{C}$-points of a scheme

$\mathrm{Repn}_{\mathrm{GL}_n}(\Pi)$ defined over $\mathbb{C}$. In fact, one can replace any GL_n with SL_n or any other algebraic group G to obtain a variety defined over any ring R to get a group scheme over R, $\mathrm{Repn}_G(\Pi)$. In fact if $G_1 \subset G_0$ then we have

$$\mathrm{Repn}_{G_1}(\Pi) = (G_1)^\ell \cap \mathrm{Repn}_{G_0}(\Pi).$$

The scheme $\mathrm{Repn}_{G_1}(\Pi)$ is actually a fiber product for the following diagram

$$\begin{tikzcd} & \mathrm{Repn}_{G_0}(\Pi) \arrow[d] \\ G_2^\ell \arrow[r] & G_0^\ell \end{tikzcd}.$$

We also want to mention that the construction is functorial in Π. If $\Pi_0 \to \Pi_1$ is a morphism of finitely presented groups then there is an induced morphism $\mathrm{Repn}_G(\Pi_1) \to \mathrm{Repn}_G(\Pi_0)$. If $w : \Pi_0 \to \Pi_1$ is given by sending a generator of Π_0 to a word in the generators of Π_1, $w\colon \gamma_i \mapsto w_i$ then $N = (N_0, \ldots, N_{\ell_1}) \in \mathrm{Repn}_G(\Pi_1)$ are mapped to $(w_1(N), w_2(N), \ldots, w_{\ell_0}(N)) \in \mathrm{Repn}_G(\Pi_0)$. Since the map w is a group homomorphism the matrices $w_1(N), w_2(N), \ldots, w_{\ell_0}(N)$ must necessarily satisfy the relations for Π_0.

Definition 4.2.1.1. Given Π a finitely presented group with ℓ generators and G a group scheme over a ring R the R-scheme $\mathrm{Repn}_G(\Pi) \subset G^l$ is called *representation scheme* (or *representation variety* of G and Π)). It is cut out by the relations imposed by the relations in the finite presentation.

When $\Pi = \pi_1(X)$ for some complex manifold these are also called *character varieties*. They are well-defined since the groups are well-defined up to conjugation. We use the notation

$$\mathrm{Repn}_G(\Pi) = \mathrm{Repn}_G(X).$$

This is again contravariant in X.

Example 4.2.1.2. In the case that $X = \mathbb{P}^1 \setminus \{0, 1, x, \infty\}$ then $\pi_1(X) = F_3$ the free group on three generators and the $\mathrm{Repn}_{\mathrm{GL}_2}(X) \subset \mathrm{GL}_2^4$ defined by the equation $M_0 M_1 M_2 M_\infty = I_2$ which is just isomorphic to GL_2^3.

Example 4.2.1.3. Let X be a genus two surface. Then

$$\pi_1(X) = \langle \alpha_1, \beta_1, \alpha_2, \beta_2 \colon \alpha_1\beta_1\alpha_1^{-1}\beta_1^{-1}\alpha_2\beta_2\alpha_2^{-1}\beta_2^{-1} = 1\rangle,$$

so $\mathrm{Repn}_{\mathrm{GL}_2}(X) \subset \mathrm{GL}_2(\mathbb{C})^4$ needs 16 variables and 4 equations.

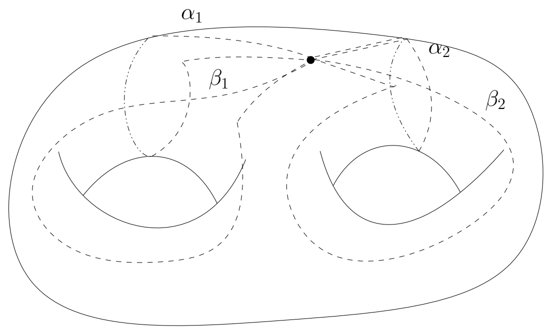

Figure 4.4: A picture of the generating cycles on a genus two compact real manifold of dimension two.

4.2.2 Character Varieties

t G be a group scheme. Two representations $\pi_1(X) \to G$ are equivalent if and only if they are conjugate. Hence we have an action by $B \in \mathrm{GL}_n$ on $(M_1, \dots, M_l)$ given by

$$B \cdot (M_1, M_2, \dots, M_l) = (BM_1B^{-1}, BM_2B^{-1}, \dots, BM_lB^{-1}).$$

We want to quotient by this action.

Definition 4.2.2.1. The algebraic stack

$$\mathrm{Char}_G(\Pi) := [\mathrm{Repn}_G(\Pi)/G]$$

is called the *character stack*.

When it is represented by a scheme, that scheme is unique up to isomorphism and we call it the *character scheme*. In the case that it is a variety we call it the *character variety*. We will abusively denote all of these things by $\mathrm{Char}_G(\Pi)$.

Remark 4.2.2.2. The point of using an algebraic stack here is not to be a jerk but to state that there is a serious mathematical issue that we need to contend with. It's very complicated and two complicated theories for which there exist entire books written about them are devoted to dealing with this issue. The first approach is Geometric Invariant Theory for which there is the famous book [MFK94]. The second would be the theory of Algebraic Stacks for which there is an entire Stacks Project. One can also use quotients in the category of sheaves too which is described here [Sta22, Tag 07S5].

As before, we let $\mathrm{Char}_G(X) = \mathrm{Char}_G(\Pi)$ for $\Pi = \pi_1(X)$ when X is connected topological space with fundamental group Π.

Example 4.2.2.3. In the case of $X = \mathbb{P}^1 \setminus \{0, 1, \infty\}$ and $G = \mathrm{SL}_2$ the character variety is $\mathbb{A}^3$:

$$\mathrm{Char}_{\mathrm{SL}_2}(\mathbb{P}^1 \setminus \{0, 1, \infty\}) \cong \mathbb{A}^3.$$

The representation space $\mathrm{Repn}_{\mathrm{SL}_2}(\mathbb{P}^1 \setminus \{0, 1, \infty\})$ is determined by three matrices in SL_2, M_0, M_1, M_∞ satisfying $M_0 M_1 M_\infty = I_2$. Using conjugation we can bring these into standard hypergeometric form with $(\theta_0, \theta_1, \theta_2)$ uniquely determining the monodromy matrices. This corresponds to the fact that the Gauss hypergeometric function has three parameters (a, b, c) see §2.5.10.

We have another example for which I won't give a proof. ♠♠♠ Taylor: [add reference, maybe Loray]

Example 4.2.2.4. In the case $X = \mathbb{P}^1 \setminus \{0, 1, \infty, x\}$ as before with $G = \mathrm{SL}_2$ then the representation variety $\mathrm{Repn}_{\mathrm{SL}_2}(X) \cong \mathrm{SL}_2^3$ since $\pi_1(\mathbb{P}^1 \setminus \{0, 1, \infty, x\}) \cong F_3$, the free group on three letters (in terms of monodromy matrices there are three M_0, M_1, M_2, M_∞ and they satisfy $M_0 M_1 M_2 M_\infty = 1$ so we think of M_∞ being determined by the first three).

It is a fact that the character variety $\mathrm{Char}_{\mathrm{SL}_2}(X)$ is a degree four hypersurface in $\mathbb{A}^7$ with coordinates (a, b, c, d, x, y, z) given by

$$a^2+b^2+c^2+d^2+x^2+y^2+z^2-(ab+cd)x-(ad+bc)y-(ac+bd)z+abcd+xyz-4 = 0.$$

4.3 The Painlevé Property

The Painlevé equations were not originally derived via the Schlesinger equations. Classically, mathematicians were looking for special functions which were solutions of differential equations. This is how they came up with Airy functions, Bessel functions, hypergeometric functions, trigonometric functions, etc.

Definition 4.3.0.1. An ordinary differential equation in one variable has the *Painlevé property* if and only if germ $f \in \mathbb{C}\langle t - t_0 \rangle$ admits a well-defined analytic continuation.

The main result concerning the

Theorem 4.3.0.2. *Consider a differential equation of the form*

$$\frac{d^2 q}{dx^2} = R(x, q, q').$$

If this equation has the Painlevé property then it is a specialization of the PVI.

An overview of the proofs can be found in [Shi03]. The first proof is in the back of Ince's book [Inc44]. The second proof is due to Malgrange

♠♠♠ Taylor: [I need to finish adding my notes here]

It then remains to determine if the functions obtained as solutions of the Painlevé differential equations are genuinely new. This was done by Umemura and others.

Chapter 5

Ordinary Differential Algebra

In this section we are going to develop more the theory of ∂-algebra done previously with an eye towards computation. In this section and the next we are following Kaplansky [Kap76], Ritt[Rit50], Kolchin [Kol73], and Boulier [Bou19, online version], Marker's notes [Mar00], as well as Buium and Cassidy's chapter in Kolchin's selected works [Kol99].

The goals of this section are two-fold.

1. Get to a point where complicated computations like the reduction of Schlesinger to Painlevé can be done routinely on a computer.

2. Give foundations for differential algebraic geometry.

A crash course in differential algebra by way of analogy is given below. In what follows we are going to let $\widehat{K}$

Polynomials In commutative algebra one works with polynomial rings $K[x_1, \ldots, x_n]$ over finitely many indeterminates and studies the ideals inside of it.

In differential algebra one works with *∂-polynomial rings*

$$K\{x_1, \ldots, x_n\} = K[x_i^{(j)} \colon 1 \leq i \leq n, j \geq 0],$$

with finitely many differential indeterminates $x_1, x_2, \ldots, x_n$. Elements of $K\{x_1, \ldots, x_n\}$ are called *∂-polynomials*. Note that the ring of ∂-polynomials has the unique differential operator ∂ extending the derivative on K such that $\partial(x_i^{(j)}) = x_i^{(j+1)}$. In ∂-algebra we study ∂-ideals (ideals closed under derivations) inside rings of ∂-polynomials.

Given $u \in K\{x_1, \ldots, x_n\}$ containing x_j nontrivially, define the *order* of u in the variable x_j:

$$\mathrm{ord}^{\partial}_{x_j}(u) = \max\{r \in \mathbb{Z}_{\geq 0} \colon \partial u / \partial x_j^{(r)} \neq 0\}.$$

If u does not involve x_j, we define $\mathrm{ord}^{\partial}_{x_j}(u) = 0$.

Varieties In algebraic geometry, we study subsets of K^n cut out by polynomial equations. They are called algebraic sets and they form a basis of closed sets for the Zariski topology.

In differential algebraic geometry, we study subsets of K^n cut out by differential polynomials. They are called *Kolchin closed subsets* and form a basis of closed sets for the Kolchin topology.

Nullstellensatz In algebraic geometry algebraic sets correspond to radical ideals.

In differential algebraic geometry, Kolchin closed subsets correspond to radical differential ideals.

Basis Theorems In commutative algebra every ideal in a polynomial ring over a Noetherian ring is finitely generated.

In differential algebra we only have a weaker statement. For all radical differential ideals I there exists a finite number of elements such that the radical of the differential ideal differentially generated by those elements is equal to I.

Both of these theorems imply noetherianity of their respective topologies.

Primary Decompositions In commutative algebra we prove that every ideal I in a Noetherian ring R can be written as a finite irredundant intersection of primary ideals. This corresponds geometrically to the decomposition of a scheme (or variety) into finitely many irreduble components.

In differential algebra we can show that every differential ideal I in $K\{x_1, \ldots, x_n\}$ can be written as a finite intersection of prime differential ideals. Note that we didn't say differential primary ideals. There exists modest improvements of this but they are harder to describe.

Transcendence Degrees In commutative algebra one has can consider the transcendence degree of a field extension. Geometrically this corresponds to the dimension of an irreducible variety.

In differential algebra, given a differential field extension $F \supset K$ one can discuss the *differential transcendence degree* of F over K which we denote by $\mathrm{trdeg}^{\partial}_K(F)$. This is the maximal n such that there exists an injection of rings $K\{x_1, \dots, x_n\} \hookrightarrow F$.

5.1 Differential Equations in One Variable

5.1.1 Envelopes and Separants

We are going to need to redo this computation later. Let R be a ∂-ring and let $f \in R\{x\}$. Given $f \in R\{x\}$, the *leader* ℓ_f of f is the highest derivative of x, $x^{(r)}$ appearing in f (i.e. such that $\partial f/\partial x^{(r)} \neq 0$.

Something that we want to describe now is the separant whose vanishing or non-vanishing on a solution has important geometric consequences.

Definition 5.1.1.1. The *separant* of f is $S_f := \partial F/\partial \ell_f$.

I want to point out that we treat $x, x', x'', \dots$ like indeterminates. In particular $\partial x^{(i)}/\partial x^{(j)} = \delta_{ij}$ where δ_{ij} is the Kronecker delta. For concreteness we give an example.

Example 5.1.1.2. If $f = 2x(x')^2 + x^3$ then $\ell_f = x'$, $I_f = 2x$ and $S_f = 2x'x$.

The following example is going to be our introduction into singular or envelope solutions of differential equations and how it interacts with the separant. We are going to have a family of solutions and then there is going to be a limiting family around them which algebraically stands out.

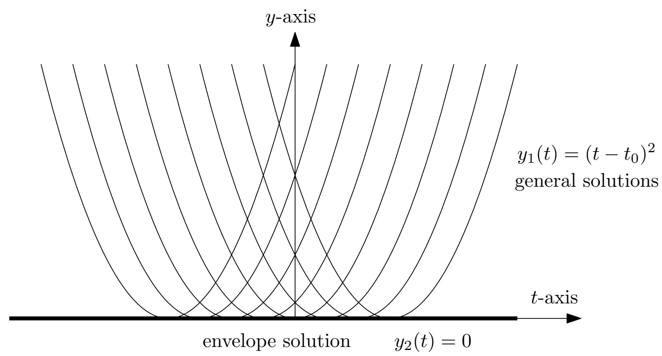

Figure 5.1: The above image plots solutions of the equation $\dot{y}^2 - 4y = 0$. The general solutions are the parabolas which are tangent to the t-axis labeled $y_1(t)$ for varying t_0. The envelope solution is $y_2(t) = 0$ and it is exactly where the separant vanishes.

Example 5.1.1.3 (Our First Envelope)**.** When we solve the ODE

$$f = (\dot{y})^2 - 4y = 0,$$

we are going to find out that there are two types of solutions "general solutions" and an envelope solution. The picture of the solutions is in Figure 5.1.

To do this lets observe that after taking a derivatives we get

$$0 = 2\dot{y}\ddot{y} - 4\dot{y} = 2\dot{y}(\ddot{y} - 2) = 0.$$

We can solve each of $\dot{y} = 0$ and $\ddot{y} - 2 = 0$ separately. These will give rise to two components of the ideal $\{f\}_\partial$ if $f = \dot{y}^2 - 4y$.

The solution fo $\ddot{y} - 2 = 0$ is $y_1(t) = t^2 + At + B$ for some constants A and B. Plugging this back into our original equation we get

$$0 = (2t + A)^2 - 4(t^2 + At + B) = A^2 - 4B$$

This then given $(t + A/2)^2 = t^2 + At + A^2/4 = t^2 + At + B$. So setting $t_0 = -A/2$ we see we get a collection of parabolas $y_1(t)$ hanging off the t-axis. These are going to be our "general solutions".

The solution of $\dot{y} = 0$ is $y_2(t) = C$ for some constant C. Plugging this back into our original equation gives

$$0 = 0^2 - 4C$$

which implies $C = 0$ which gives $y_2(t) = 0$. This solution is going to be our "envelope solution".

Before leaving, let's observe the relationship with the separant. We have $S_f = \dot{y}$. The general solutions $y_1(t)$ do not vanish on the separant. The envelope solutions *do* vanish on the separant.

Given a general ODE in one variable

$$\Sigma \colon f = 0,$$

defined by $f \in K[y]_\partial$ we can always decompose the space of solutions into two pieces. One of the pieces is where the separant S_f vanishes (the envelope piece) and the other is where the separant doesn't vanish (the general piece). We don't know that these pieces yet are irreducible components of a differential algebraic variety, but when $f \in K[y, y', y'', \ldots]$ is irreducible (as a polynomial in many variables) this will turn out to be the case.

Theorem 5.1.1.4. *For $g \in K[y]_\partial$ we have*

$$\{g\}_\partial = \{g, S_g\}_\partial \cap (\{g\}_\partial \colon S_g).$$

Proof. Let $E = \{g, S_g\}_\partial$ and $(\{f\}_\partial \colon S_g)$. Clearly $\{g\}_\partial \subset E \cap J$. Conversely, let $h \in E \cap J$. We have $h^2 \in \{g, S_g\}_\partial$ and hence $h \in \{hf, hS_g\}_\partial \subset hg, g\}_\partial \subset \{g\}_\partial$. Here we used the property of colon ideals. $\square$

Let's state our results that we can prove after introducing the division algorithm. One of the main thing that I wanted to talk about was the following application of pseudodivision.

Theorem 5.1.1.5. *Let $g \in K\{x\}_\partial$.*

1. *If $f \in [g]$ and $\operatorname{ord}^\partial(f) \leq \operatorname{ord}^\partial(g)$ then $g|f$.*

2. *If g is irreducible then $J = (\{g\}_\partial : S_g)$ is a prime ideal.*

The proof of this is given in §5.1.3. We remind the reader that for an ideal I in a ring R and an element $s \in R$ we have $(I : s) = \{r \in R\colon sr \in I\}$ is an larger ideal containing I (see [AM16]).

The second item in this theorem leads to two imporant ideals. The first idea is that simple field extensions and prime ideals in the differential setting are not so simple. Second, if we remember how prime and primary decompositions work from Commutative algebra, we see that this item is telling someting about the components of $\{g\}_\partial$. In particular by localizing (or specifying the inequation $S_g \neq 0$) we are picking out an irreducible component of the variety of solutions.

5.1.2 Univariate Pseudodivision Algorithm

In order to define our division algorithm (which is sort of crappy compared to the Groebner basis tools we know and love from usual commutative algebra) we need to introduce some terminology which will help us order our terms of our differential polynomials.

Definition 5.1.2.1. The *initial* of $f \in K\{y\}$ is the coefficient I_f of the top degree term of the leader. In other words if

$$f = a_0 + a_1\ell_f + \cdots + a_d\ell_f^d, \quad I_f = a_d$$

where if $\ell_f = x^{(r)}$ then $a_j \in R\{x, x', \ldots, x^{(r-1)}]$ is I_f. The d appearing in the formula about is the *leader degree* and we denote it by $\mathrm{ldeg}(f)$.

In what follows we will give $K\{x\}$ the term ordered induced by the lexicographic ordering such that $x^{(r)} \prec x^{(r+1)}$ for all $r \geq 0$. This is our first example of a ranking. This then defined an ordering on the collection of all polynomials where we say that $A \prec B$ if and only if $\mathrm{LT}(A) \prec \mathrm{LT}(B)$ where LT denotes the leading term. This ordering is completely determined by $f \mapsto \ell_f^{\mathrm{ldeg}(f)}$ where higher derivatives are larger, then degrees break ties.

Lemma 5.1.2.2. *Let $F \in K\{x\}$ be a non-constant ∂-polynomial.*

1. $\ell_{\partial^n(F)} = \partial^n(\ell_F)$

2. $I_{\partial^n(F)} = S_F$.

Proof. We can write $\ell = \ell_F$ and then expand F as $F = \sum_{j=0}^{d} a_j \ell^j$. We then have

$$\partial(F) = \left(\sum_{j=0}^{d} \partial(a_j)\ell^j\right) + \partial(\ell)\left(\sum_{j=0}^{d} a_j j \ell^{j-1}\right).$$

This proves that $I_{\partial(F)} = S_F$ and $\ell_{\partial(F)} = \partial(\ell_F)$. The formula $\ell_{\partial^n(F)} = \partial(\ell_F)$ is evident. Also, one can see from differentiating the formula repeatedly that $I_{\partial^n(F)} = S_F$. $\square$

We now get the following algorithm for pseudodivision. This is really the heart of the method of characteristics in several variables for determining ideal membership.

Theorem 5.1.2.3 (The Pseudodivision Algorithm)**.** *Fix some non-constant $F \in K\{x\}$. For all $A \in K\{x\}$ there exists integers $a, b \geq 0$, and some $\widetilde{A} \in K\{x\}$ with $\widetilde{A} \prec F$ such that*

$$I^a S^b A \equiv \widetilde{A} \quad [F],$$

where $I = I_F$ is the initial of F and $S = S_F$ is the separant of F.

Proof. We break the proof into cases.

- (case 1) A is lower than F.
- (case 2) $\mathrm{ord}_x^{\partial}(A) = \mathrm{ord}_x^{\partial}(F)$ but $\deg_\ell(A) > \deg_\ell(F)$.
- (case 3) $\mathrm{ord}_x^{\partial}(A) > \mathrm{ord}_x^{\partial}(F)$

For simplicity of notation we will let $\mathrm{ord}_x^{\partial}(A) = r_A$, $\mathrm{ord}_x^{\partial}(F) = r_F$, $\deg_{\ell_A}(A) = d_A$ and $\deg_{\ell_F}(F) = d_F$.

In case 1, we are done so we do nothing.

In case 2, we let $A_1 = I_F A_0 - I_A \ell^{d_A - d_F} F$. We have A_1 of lower degree than $A_0 = A$ in ℓ_A. If A_1 is lower than F we are done. If not, we repeat this process. Since the degree is lowered after every interation we eventually terminate.

In case 3, we let $A_0 = A$ and $F_0 = \partial^{r_A - r_F}(F)$. These two polynomials will have the same order and we can now apply case 2. Note that $I_{F_0} = S_F$ so we will eventually get

$$S_F^b A \equiv \widetilde{A} \mod [F],$$

with $\widetilde{A}$ lower than F_0. □

In formulas this reads,

$$I^a S^b A = (a_0 F + a_1 \partial(F) + \cdots + a_r \partial^a S^b A) + \widetilde{A}$$

with remainder $\widetilde{A}$. We get a contribution to a for every time case 1 is used in the proof and a contribution to b every times case 2 is used in the proof. There is no clean division here like in the division algorithm for polynomials.

Remark 5.1.2.4. In the case that we only want to lower the order and don't care about the degree, we only need to use separants. This plays a role in our decomposition into general and envelope components.

5.1.3 Application of Pseudodivision

Lemma 5.1.3.1 (Divisibility Lemma)**.** *Let $g \in K[x]_\partial$. If $f \in [g]$ and $\operatorname{ord}^\partial(f) \leq \operatorname{ord}^\partial(g)$ then $g|f$.*

Proof. This proof is a little tricky because it requires you to really treat elements of $K[x]_\partial$ as polynomials in weird looking symbols and nothing more. For me, this is psychologically difficult for some reason.

The relation $f \in [g]$ is equivalent to a formula

$$f = c_0 g + c_1 \partial(g) + \cdots + c_r \partial^r(g) \tag{5.1.1}$$

where $c_i \in K[x]_\partial$.

Let's write $\ell = \ell_g$, $S = S_g$ and observe that

$$g^{(j)} = S\ell^{(j)} + T_j, \qquad j \geq 1$$

where $T_j \prec \ell^{(j)}$. The key observation is that because $\mathrm{ord}^\partial(f) \leq \mathrm{ord}^\partial(g)$ only the right hand side of (5.1.1) can involve the indeterminates $\ell^{(j)}$ for $j \geq 1$. In particular, we can map them to whatever we want and still have an identity!

We start with using $\ell^{(r)} \mapsto T_r/S$ to get (after clearing denominators) a new equation

$$S^{e_r} f = d_0 g + d_1 \partial(g) + \cdots + d_{r-1}\partial^{r-1}(g).$$

We then proceed inductively getting

$$S^{e_r + e_{r-1} + \cdots + e_1} f = z_0 f$$

for some $z_0 \in K[x]_\partial$. We have that $g | S^e f$ by irreducibility of f. Since $S \prec g$ this implies that $g \nmid S$ and hence that $g|f$. $\square$

Lemma 5.1.3.2 (Irreducibility Lemma)**.** *If g is irreducible then $J = (\{g\}_\partial : S_g)$ is a prime ideal.*

Proof. Suppose that $f_1 f_2 \in J$. We will show that $f_1 \in J$ or $f_2 \in J$. For future reference we will let $S^{a_1} f_1 = \widetilde{f_1}$ and $S^{a_2} f_2 = \widetilde{f_2}$ where $\widetilde{f_i}$ have order less than or equal to that of g.

$$\begin{aligned} f_1 f_2 \in J \quad &\iff S f_1 f_2 \in \{g\}_\partial \\ &\iff (S f_1 f_2)^b \in [g]_\partial \\ &\implies S^{a_1 b} S^{a_2 b} S^{-1} (S f_1 f_2)^b = (S^{a_1} f_1 S^{a_2} f_2)^b \in [g]_\partial \\ &\implies (\widetilde{f_1}\widetilde{f_2})^b \in [g]_\partial \end{aligned}$$

By the divisibility lemma $\square$

5.2 Bigger Fish: Ritt-Raudenbaush, Noetherianity, Prime Decomposition

One thing that we have to contend with in differential algebra is non-Noetherianity. The rings are are dealing with have an infinite number of generators:

$$K[x, x', x'', x''', \ldots].$$

This ring is abstractly isomorphic (as rings) to $K[x_1, x_2, x_2, \ldots]$ which is not Noetherian since it contains the chain

$$\langle x_1 \rangle \subset \langle x_1, x_2 \rangle \subset \langle x_1, x_2, x_3 \rangle \subset \cdots .$$

There are still hopes.

The most naive thing to do would be to put the word "differential" in front of everything.

Problem 5.2.0.1.

1. Are all ∂-ideals in $K[x, x', x'', \ldots]$, ∂-finitely generated?
2. Do ∂-ideals in $K[x, x', x'', \ldots]$ satisfy the ascending chain condition?
3. Are these two properties equivalent?

In commutative algebra, if a ring R is Noetherian then so is $R[x]$ and being Noetherian is equivalent to ideals being finitely generated. This property allows us to conclude inductively that every ideal in $F[x_1, x_2, \ldots, x_n]$ for F a field is finitely generated. This is the famous Hilbert Basis Theorem.

It is a general theme in differential algebra that Naive generalizations don't work. Properties (1) and (2) above are false.

Exercise 5.2.0.2 (Counter Example to ACC)**.** Consider the chain of ∂-ideals in $K\{x\}$ given by

$$[x^2] \subset [x^2, (x')^2] \subset [x^2, (x')^2, (x'')^2] \subset \cdots ,$$

is an infinite ascending chain.

The examples given in the next exercise show that not only are all ideals not ∂-finitely generated but even products of ∂-finitely generated ideals are mostly not ∂-finitely generated. Here we say a ∂-ideal I is *∂-finitely generated* if and only if there exist some $f_1, \ldots, f_n \in R$ such that $I = [f_1, \ldots, f_n]_\partial$. Here is the counter-example.

Exercise 5.2.0.3 (Example of Infinite Generation)**.**

1. The ideal $\{xy\}_\partial$ is not ∂-finitely generated.

2. The ideals $[x]$ and $[y]$ in $K\{x, y\}$ are not ∂-finitely generated by
$$[x][y] = \langle x^{(i)}y^{(j)} \colon i, j \geq 0\rangle = \{xy\}_\partial = [x]_\partial \cap [y]_\partial.$$

To do this exercise it is useful to have Levi's Lemmas in hand which are given in §**??**.

So what *do* we get?

$$\text{(Basis Theorem)} \implies \text{(Noetherianity), (Prime Decomposition)}$$

5.2.1 Noetherianity and Finite Generation

The next example is one of the product of two ∂-finitely generated ideals which is not finitely generated.

Theorem 5.2.1.1. *The Ritt-Raudenbacher basis property implies any infinite ascending chain of radial differential ideals terminates.*

Proof. Consider for the sake of contradiction and infinite ascending chain of Δ-ideals
$$I_1 \subsetneq I_2 \subsetneq I_3 \subsetneq \cdots .$$
Let $I = \bigcup_{j\geq 1} I_j$. By the Ritt-Raudenbacher basis theorem there exists some $f_1, \ldots, f_n \in R$ such that $I = \{f_1, f_1, \ldots, f_n\}_\Delta$. Since $f_i \in I$ there exists some N_i such that $f_i \in I_{N_i}$. Let $N = \max\{N_1, \ldots, N_n\}$. We have $I \subset I_N$ and hence $I = I_N$ and the chain terminates. $\square$

Corollary 5.2.1.2. *The Kolchin topology on $\widehat{K}^n$ is Noetherian.*

5.3 Differential Transcendence

We introduce differential algebraic geometry by analogy with algebraic geometry. A big difference between algebraic geometry and differential algebraic geometry is how intersection theory works, as shown in Figure 5.2. ♠♠♠ Taylor: [I'm going to fill this in]

Let $F \supset K$ be a ∂-finitely generated extension of ∂-fields. One has the following important correspondence:

$$\operatorname{trdeg}_K^\partial(F) = 0 \iff \operatorname{trdeg}_K(F) < \infty.$$

Geometrically this corresponds to the fact that a differential algebraic variety Σ over K has differential dimension zero if and only if it has finite absolute dimension

$$\dim^\partial(\Sigma) = 0 \iff a(\Sigma) < \infty.$$

The number $a(\Sigma)$ is the *absolute dimension* of the differential algebraic variety Σ [Bui93, §2, pg 485].

In the case that Σ is irreducible $a(\Sigma)$ can be computed as the transcendence degree of the function field $K(\Sigma)$ over K. Equivalently it can be computed as the (Krull) dimension of the underlying jet scheme $\Sigma^{[\infty]}$ of Σ. By this we just mean the usual scheme-theoretic dimension of the scheme $\Sigma^{[\infty]}$. One issue we have to contend with is the fact that rings and ideals defining $\Sigma^{[\infty]}$ are not finitely generated, so the proofs of many theorems we would like to use do not apply directly.

The existence of the two notions of dimension means there are two ways to think about differential algebraic varieties: in terms of ∂-indeterminates $\{x_i\colon 1 \leq i \leq n\}$ and in terms of classical indeterminates $\{x_i^{(j)}\colon 1 \leq i \leq n, j \geq 0\}$ (see Figure 5.2). From the perspective of differential indeterminates, "neighborhoods" of intersections of ∂-dimension zero are described by finite dimensional schemes. Both perspectives are useful.

5.4 Appendix: Levi's Lemmas on $[y^n]$ and $[xy]$

??

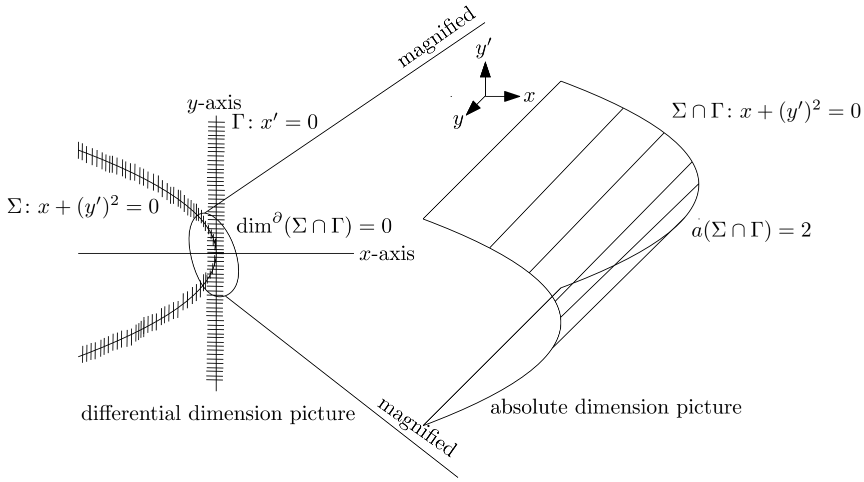

Figure 5.2: The above picture shows two ways to draw the intersection of the differential algebraic varieties $\Sigma\colon x' = 0$ and $\Gamma\colon x + (y')^2 = 0$. From the point of view of differential transcendence degrees, $\Sigma \cap \Gamma$ has dimension zero. From the point of view of transcendence degrees or absolute dimensions, $\Sigma \cap \Gamma$ has dimension 2.

5.4.1 Monomials in Ordinary Differential Polynomial Rings

It will be conveniant to introduce some notation for monomials in $K[y]_\partial$ where y is a single indeterminant. For $\alpha \in \mathbb{Z}_{\geq 0}[\partial]$ with $\alpha = a_0 + a_1\partial + \cdots + a_r\partial^r$ we will write

$$y^\alpha = y^{a_0}(y')^{a_1} \cdots (y^{(r)})^{a_r}.$$

We assign two gradings to $K[y]_\partial$. A grading by weight and grading by degree. We say that the *degree* of y^α (or just α) is

$$\deg(y^\alpha) = a_0 + a_1 + \cdots + a_r.$$

Similarly, we define the *weight* of y^α (or just α) is

$$\mathrm{wt}(y^\alpha) = a_1 + 2a_2 + \cdots + ra_r.$$

A basic observation is that terms with no derivatives have weight zero, terms with single derivatives have weight one, terms with only second derivatives have weight two etc. If we view α as a polynomial $\alpha(x) \in \mathbb{Z}[x]$ then $\deg(\alpha) = \alpha(0)$ and $\mathrm{wt}(\alpha) = \alpha'(0)$.

Exercise 5.4.1.1. Let $A \in K\{y\}$ be a differential monomial which is homogeneous in both degree and weight. Show that $\partial(A)$ is homogeneous in both degree and weight and that

$$\deg(\partial(A)) = \deg(A), \quad \mathrm{wt}(\partial(A)) = \mathrm{wt}(A) + 1.$$

In several ∂-indeterminates $x = (x_1, \ldots, x_n)$ we can again use a multi-index notation. We will take $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{Z}_{\geq 0}[\partial]^n$ and let

$$x^\alpha = (x_1, \ldots, x_n)^{(\alpha_1, \ldots, \alpha_n)} = x_1^{\alpha_1} \cdots x_n^{\alpha_n}$$

as where for each i we have $\alpha_i = a_{i0} + a_{i1}\partial + \cdots + a_{ir_i}\partial^{r_i}$ and

$$x_i^{\alpha_i} = x_i^{a_{i0} + a_{i1}\partial + \cdots + a_{ir_i}\partial^{r_i}}$$

as before.

This section is about nilpotents which are largely ignored in modern differential algebra due to the fact that we largely need to work with radical ideals to get nice finiteness properties. The two most ways zero divisors get introduces is through product relations $xy = 0$ and power relations $x^n = 0$ (nilpotents). This sectoin investigates the behaviour of differential ideals generated by these two relations.

Consider the differential ring $(\mathbb{Q}[x, y], \partial)$. In this section we are just going to work with a single derivative and use [] for differential ideas. We also retain the monomial notation $x^\alpha = x^\alpha = x^{a_0}\dot{x}^{a_1}\ddot{x}^{a_2} \cdots (x^{(r)})^{a_r}$ for $\alpha = a_0 + a_1\partial + \cdots + a_r\partial^r \in \mathbb{Z}_{\geq 0}[\partial]$. In [**?**] the following interesting questions were addressed:

Problem 5.4.1.2.

1. Consider the differential ideal $[x^n]$. Which monomials are in $[x^n]$?
2. Consider the ∂-ideal $[xy]$. Which monomials $x^\alpha y^\beta$ are in $[xy]$?
3. What about for $\{xy\}_\partial$?

5.4.2 Levi's Lemmas for $[x^n]$

The following is the most basic thing we can prove.

Lemma 5.4.2.1. *Let R be a differential $\mathbb{Q}$-algebra. Let I be a differential ideal. Suppose that $a \in R$ satisfies $a^n \in I$. Then $\dot{a}^n \in I$.*

Proof. If $a^n \in I$ then $\partial(a^n) = na^{n-1}\dot{a} \in I$. By the $\mathbb{Q}$-algebra hypothesis, $a^{n-1}\dot{a} \in I$. Taking derivatives again we get $(n-1)a^{n-2}\dot{a}^2 + a^{n-1}\ddot{a} \in I$. After multiplying by a' this implies $a^{n-2}\dot{a}^2 \in I$ since $a^{n-1}\dot{a} \in I$. This process continues. Knowing that $a^{n-j}\dot{a}^j \in I$ allows us to show that $a^{n-j-1}\dot{a}^{j+1} \in I$ by taking derivatives, multiplying by $\dot{a}$ and using the inductive hypothesis to get rid of a term. Setting $n-j-1=0$ we get $j = n-1$ which shows $\dot{a}^n \in I$. $\square$

We aim to give a much stronger version of the above in what follows culminating in Lemma 5.4.2.5.

Lemma 5.4.2.2. *A monomial y^α for $\alpha \in \mathbb{Z}_{\geq 0}[\partial]$ appears in $\partial^r(y^n)$ nontrivially if and only if it has degree n and weight r.*

Proof. The proof is by induction on r. The case $r=0$ is trivial. In the inductive step one needs to take a derivative of $y^{a_0}\dot{y}^{a_1}\cdots(y^{(r)})^{a_r}$ and observe that every monomial has an exponent which is a modification of $(a_0, a_1, \ldots, a_r, 0)$ by one of $(-1,1,0,\ldots,0,0), (0,-1,1,\ldots,0,0), \ldots, (0,0,0,\ldots,-1,1)$. Note that

$$\begin{aligned}\partial(y^{a_0}\dot{y}^{a_1}\cdots(y^{(r)})^{a_r}) =&\partial(y^{a_0})\dot{y}^{a_1}\cdots(y^{(r)})^{a_r}\\ &+ y^{a_0}\partial(\dot{y}^{a_1})\cdots(y^{(r)})^{a_r}\\ &+\cdots\\ &+ y^{a_0}\dot{y}^{a_1}\cdots\partial(y^{(r)a_r}).\end{aligned}$$

Since $\partial((y^{(j)})^{a_j}) = (y^{(j)})^{a_j-1}y^{(a_j+1)}$. This gives the contribution $-e_j+e_{j+1}$ in the exponent if e_j is the jth elementary basis vector. In the notation where exponents are $\mathbb{Z}_{\geq 0}[\partial]$ the new exponent is $-\partial^j + \partial^{j+1}$. $\square$

Before diving in to the next Lemma the reader may wish to consult the example found directly after the proof.

Lemma 5.4.2.3. *The lowest monomial of $\partial^m(y^n) \in K\{y\}$ with respect to its unique ranking is*

$$\mathrm{Low}(\partial^m(y^n)) := \partial^q(y)^{n-r}\partial^{q+1}(y)^r$$

where $q, r \in \mathbb{Z}_{\geq 0}$ are the unique integers with $0 \leq r \leq n$ in the Euclidean algorithm such that $m = qn + r$.

Proof of Lemma 5.4.2.3. First observe that

$$\deg(\partial^q(y)^{n-r}\partial^{q+1}(y)^r) = n,$$

$$\mathrm{wt}(\partial^q(y)^{n-r}\partial^{q+1}(y)^r) = q(n-r) + (q+1)r = qn + r = m,$$

so by Lemma 5.4.2.2, $\partial^q(y)^{n-r}\partial^{q+1}(y)^r$ appears as a monomial of $\partial^m(y^n)$. We now prove this is the lowest by induction on m. We will let y^β be a minimal term with $\beta \in \mathbb{Z}_{\geq 0}[\partial]$. If $m = 0$ then y^n is the lowest term which is when $q = 0$ and $r = 0$. We now do the inductive step. The lower term of $\partial^{m+1}(y^n)$ will be a term of $\partial(y^\gamma)$ where $\gamma = (n-r)\partial^q + r\partial^{q+1}$. This is because if $\beta \prec \beta'$ then $\beta - \partial^i + \partial^{i+1} \prec \beta' - \partial^i + \partial^{i+1}$. Now $\gamma - \partial^q + \partial^{q+1}$ is the smallest term. In the case $r = n-1$ we find that $m + 1 = (q+1)n$ and can check directly that $\partial(\partial^q(y)\partial^q(y)^{n-1})$ has the term $\partial^{q+1}(y)^n$. In the case $r < n-1$ we find that $m + 1 = n + r + 1$ and that $\gamma - \partial^q + \partial^{q+1} = (n-(r+1)) + (r+1)\partial^{q+1}$ is the lowest term. [Rit43, §21] $\square$

Here is an example of the above theorem.

Example 5.4.2.4. In the case $m = 10$ and $n = 3$ the proposition is saying that $\partial^{10}(y^3)$ has $\partial^3(y)^2\partial^4(y)$ as the lowest term in the ordering since $10 = 3(3) + 1$.

Levi's Criterion will give a recipe on the weight and degree of y^α for membership in the differential ideal generated by y^n.

Theorem 5.4.2.5 (Levi's Lemma)**.** *Let n be a non-negative integer. Suppose that y^α for $\alpha \in \mathbb{Z}_{\geq 0}[\partial]$ has weight w and degree d. We have the following*

$$w < f(n,d) \implies y^\alpha \in [y^{n+1}]$$

where $f(n,d) = q_d(q_d - 1)(n-1) + 2q_d r_d$ where q_d and r_d are the unique integers such that $d = q_d n + r_d$.

Proof. We will say a monomial $y^{a_0 + a_1\partial + \cdots + a_s\partial^s}$ is *dilute* (for n) if and only if for all j, $a_j + a_{j+1} < n$. If y^α is not dilute it will be called *concentrated* (for n). We will omit the "for n" from now on. A differential polynomial is called *concentrated* (resp *dilute*) if all of its monomials are. There are a series of claims:

1. If y^α is concentrated then there exists some j such that $L_j = \mathrm{Low}(\partial^j(y^n))$ divides y^α.

   Proof. Recall that $\mathrm{Low}(\partial^j(y^n)) = \partial^{q_j}(y)^{n-r_j}\partial^{q_j+1}(y)^{r_j}$ where $j = q_j n + r_j$ is the representation from the division algorithm. By definition $y^\beta = \partial^i(y)^a\partial^{i+1}(y)^b | y^\alpha$ for some a, b with $a+b \geq n$. Without loss of generality we can asume $a+b = n$. We also must have $\mathrm{wt}(y^\beta) = ia + (i+1)b = m$. To get the division we increase j until $q_j = a$ so that $j = an$. In this case $L_j = \partial^a(y)^n$. We can then increase j further to get $j = an + r$ for $0 \leq r < n$ and $L_j = \partial^a(y)^{n-r}\partial^{a+1}(y)^r$. □

2.

   Lemma 5.4.2.6. *If A is degree-homogeneous of degree d and weight-homogenous of weight w it is congruent to a dilute bihomogenous polynomial of the same degree and weight modulo $[y^n]$.*

   Proof. By the previous $y^\alpha \equiv 0 \mod [y^n]$ for every y^α concentrated. The result follows. □

   This shows that every A is congruent to an n-dilute ∂-polynomial modulo $[y^n]$. Conversely, we will show that the only n-dilute polynomial in $[y^n]$ is the zero polynomial.

3. If y^α is degree d and has weight less than $f(n,d)$ then y^α is n-concentrated.

4. By hypothesis, we have a concentrated polynomial which is equivalent to a dilute one.

We can peel off the concentrated terms, term-by-term and kill them. Suppose A is not dilute. Then let $A = bB + R$ where B is the lowest concentrated. We have $B = L_j H$ for some H and j. We have $\partial^j(y^n) = cL_j + \sum_{i=1}^s c_i P_i$ where $P_i \succ L_j$. Then

$$\begin{aligned} A &= bB + R \\ &= b\left(\frac{1}{c}\left[\partial^j(y^n) - \sum_{i=1}^s c_i P_i\right]\right) H + R \\ &\equiv \frac{-b}{c}\sum_{i=1}^s c_i P_i H + R \mod [y^n], \end{aligned}$$

and the terns $P_i H$ are higher that G in the ordering. We now repeat this process to kill higher and higher terms. Since there are only finitely many terms we can do this with, the process terminates. □

5.4.3 Levi's Lemmas for $[xy]$

♠♠♠ Taylor: [I need to add my notes here. In the meantime, Ritt's book is semi-readable and Levi's original paper is longwinded but readable. You can find both online. The main point here is that $[xy]$ has a TON of nilpotents.]

Chapter 6

Partial Differential Algebra

6.1 Monomials, Ranking, and Orders

Let R be a Δ-ring with $\Delta = \{\partial_1, \ldots, \partial_m\}$. Let $A = R[y_1, \ldots, y_n]_\Delta$. Let Θ be the collection of differential operators

$$\Theta = \{\partial^\alpha \colon \alpha \in \mathbb{Z}^n_{\geq 0}\}$$

where $\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_m^{\alpha_m}$.

Definition 6.1.0.1. A *ranking* is a total ordering $\prec$ on the set of differential variables $\{\theta(y_j) \colon \theta \in \Theta, 1 \leq j \leq n\}$ satisfying the following two axioms

1. $u \prec \theta u$,
2. $u \prec v \implies \theta u \prec \theta v$

For a given indeterminate $\theta(u)$ we define a filtration $(K[y_1, \ldots, y_n]_\Delta)_{\leq \theta(u)}$ to be the ring generated by variables of lower rank.

Just as in the ordinary case we make a sequence of definitions.

Definition 6.1.0.2. Let $f \in K[x_1, \ldots, x_m]_\Delta$ and fix a ranking on the variables. The *leader* ℓ_f of f is the higher rank element of $\{\theta x_j \colon 1 \leq j \leq n, \theta \in \Theta$ such that $\partial f / \partial \ell_f \neq 0$. If let $\ell = \ell_f$ and write

$$f = a_d \ell^d + a_{d-1}\ell^{d-1} + \cdots + a_0,$$

with $a_j \in K[x_1, \ldots, x_n]_{\Delta, \leq \ell}$ then

- The top coefficient $a_d = I_A$ is called the *initial* respect to the ranking.
- The partial derivative of f with respect to the leader $\partial f / \partial \ell$ is called the *separant* of f with respect to the ranking and we denote it by S_f.
- The degree d is the *leader degree* and we denote it by $\mathrm{ldeg}(f)$.

Note that to compare f and g we can use $\ell_f^{\mathrm{ldeg}(f)}$ and $\ell_g^{\mathrm{ldeg}(g)}$. The set of elements

$$\{\theta(x_j)^d \colon \theta \in \Theta, d \geq 1\}$$

is an ordered set called the set of ranks. We will set $\mathrm{rk}(f) = \ell_f^{\mathrm{ldeg}(f)}$ and call it the *rank* of f and write $f \prec g$ if and only if $\mathrm{rk}(f) \prec \mathrm{rk}(g)$.

Recall that a *term order* is total ordering on monomial that respects multiplication. That is, it is a total ordering $\prec$ for that for all monomial M,N, and L if $M \prec N$ then $LM \prec LN$. Given a ranking there is an induced term order on the collection of monomials given by taking lexicographic order on the differential variables. We will let $\prec$ also denote the term order induced by the ranking $\prec$.

Example 6.1.0.3. For a single variable y there is a unique term in $K\{y\}$. The monomials y^α are ordered first by order and then by degree. Ritt phrases this as saying $y^\alpha \prec y^\beta$ if and only if the greatest i such that $\alpha_i - \beta_i \neq 0$ we have $\alpha_i - \beta_i < 0$.

Theorem 6.1.0.4. *Fix a ranking on a ring of differential polynomials. If A is a differential polynomial and $\mathrm{LM}(A) = I\ell^m$ where I is the initial and ℓ^m is the leader to some power then $\mathrm{LM}(\partial(A)) = (Im\ell^{m-1})\partial(\ell)$ with new leader $\partial(\ell)$ and new initial $mI\ell^{m-1}$.*

Proof. We have $\partial(I\ell^m) = \partial(I)\ell^m + m\ell^{m-1}\partial(\ell)I$. We have $\partial(\ell) \succ \ell \succ I$. This means that $\partial(\ell)$ will precede any differential variable in any of the monomials of $\partial(I)$ by the second axiom of rankings. Hence $\mathrm{LT}(\partial(A)) = m\ell^{m-1}I\partial(\ell)$ and $I_{\partial(A)} = m\ell_A^{m-1}I_A$ and $\ell_{\partial(A)} = \partial(\ell_A)$. $\square$

6.2 Characteristic Sequences

Characteristic sequences $G = (g_1, \ldots, g_r)$ are things that characterize membership in a ∂-ideal I. There are like a crappy version of a Groebner basis for an ideal in that they can determine membership of the ideal but they actually don't generate the ideal. They are a valuable computational technique but also a source of many of our headaches in that one of the largest open problems in differential algebra, the Ritt Problem, is centered around them.

For the purpose of having something to focus on we state the main result of this section. Definitions will be developed as we go along.

Theorem 6.2.0.1. *Let I be a Δ-ideal in $K[x_1, \ldots, x_n]_\Delta$. Fix a ranking on $K[x_1, \ldots, x_n]_\Delta$. The following are equivalent for an ordered set $G = (g_1, \ldots, g_r)$ of elements from I,*

1. *The sequence G is a characteristic sequence of I.*

2. *$f \in I$ if and only if $\mathrm{red}_G(f) = 0$.*

In the above theorem, $\mathrm{red}_G(f)$ is the output of our pseudodivision algorithm analogous to the one we developed in the univariate case.

6.2.1 Chomp

The game chomp is a game that is useful for describe when certain processes terminate. This can be used in a proof of the Hilbert Basis Theorem or when we want to show that Buchberger's Algorithm terminates. A drawing of three moved in the game played in the plane is pictured in Figure 6.1.

The initial board is setup so there is a bead at each point in the first quadrant of the xy-plane including the lattice points on the axes. At each round a chomp is made. To make a chomp you select one bead is selected and all the beads above above and to the right are removed (so an infinite number of beads are removed at each stage). The game is over when there are no more chomps to be made. I really didn't describe a two player winning or

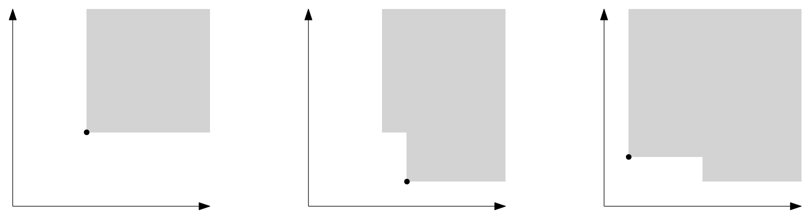

Figure 6.1: The moves in the game chomp are pictured. At each stage a vertex is selected in $\mathbb{Z}^2_{\geq 0}$ and every lattice point above and below it are removed. After a finite number of moves this game must terminate.

losing strategy, but that doesn't really matter. What matters is that the game terminates.

The idea is that each node (i,j) in the xy-plane represents a monomial $x^i y^j$ and the chomp represents all of the monomials that are divisible the node we selected.

The are obvious variants of this game in higher dimensions.

6.2.2 Well Orderings

In this section we will let $(S, \leq)$ be a partially ordered set. When we want a strict inequality we will use the $a < b$ to mean $a \leq b$ and $a \neq b$ for $a, b \in S$.

We recall that a partially ordered set $(S, \leq)$ is a *well ordered set* if and only every $S_0 \subset S$ has a minimal element.

Exercise 6.2.2.1. The following are equivalent for a partially ordered set $(S, \leq)$.

1. The set $(S, \leq)$ is a well ordering.
2. Every descending sequence $s_1 \geq s_2 \geq \dots$ in S terminates.

Proof. To show that the descending chain condition implies a well ordering we will argue by contrapositive. Suppose there is a set without a minimal element. Then form a descending chain.

To show that the well ordering property implies descending chains we again use a contrapositive. Given an infinite descending chain $a_0 > a_1 > a_2 > \cdots$ the set $S_0 = \{a_j\}_{j\geq 0}$ is a set without a minimal element. $\square$

6.2.3 Characteristic Sets

Let $\Delta = \{\partial_1, \ldots, \partial_m\}$ and let $\Theta = \{\partial^\alpha \colon \alpha \in \mathbb{Z}^m_{\geq 0}\}$ whre $\partial^\alpha = \partial_1^{\alpha_1} \cdots \partial_m^{\alpha_m}$.

Lemma 6.2.3.1. *There denote not exist an infinite sequence $\theta_1, \theta_2, \ldots$ where $\theta_j \in \Theta$ such that $\theta_j \nmid \theta_{j+i}$ for all $j \geq 1$ and $r > 1$.*

Exercise 6.2.3.2. Prove Lemma 6.2.3.1 using the idea that the game chomp from §6.2.1 terminates.

Definition 6.2.3.3. Fix a ranking on $K[x_1, \ldots, x_n]_\Delta$. Let $f, g \in K[x_1, \ldots, x_n]_\Delta$. We say that f is *reduced* with respect to g if and only if

1. No property derivative of the leader of g appears in f: for all $\theta \in \Theta$, $\partial f / \partial\theta(\ell_g) = 0$.
2. If the leaders are the same then g has bigger leader degree: $\mathrm{ldeg}(g) > \mathrm{ldeg}(f)$.

There exists a division algorithm similar for reducing f with respect to some g and the output of the algorithm is some $\mathrm{red}_g(f)$ which satisfies where

$$\mathrm{red}_g(d) = \begin{cases} \widetilde{f}, & \text{reduced} \\ f, & \text{if } f \text{ reduced} \end{cases}$$

where one has

$$sf \equiv \widetilde{f} \qquad [g]_\Delta$$

and s is in the multiplicative set generated by I_g and S_g.

Exercise 6.2.3.4. Building on the pseudodivision algorithm in the ODE case (Theorem 5.1.2.3) construct an algorithm which gives the desired output.

For sequences $G = (g_1, \ldots, g_r)$ of differential polynomials we will use the notation $|G| = r$.

Definition 6.2.3.5. A ordered sequence $G = (g_1, \dots, g_r)$ is *autoreduced* if and only if for all g_i, g_j with $i \neq j$ we have g_i reduced with respect to g_j.

We will always order our sequences with respect to a ranking for that $g_1 \prec g_2 \prec \cdots \prec g_r$.

Definition 6.2.3.6. We will say that f is reduced with respect to $G = (g_1, \dots, g_r)$ if and only if f is reduced with respect to g_j for each j.

For any $h_1 \prec h_2 \prec \cdots \prec h_r$ which is comparable, we will have $\mathrm{rk}(h_i) = \mathrm{rk}(g_i)$ for each i.

Theorem 6.2.3.7. *The do not exist infinite autoreduced sequences in* $K[x_1, \dots, x_n]_\Delta$.

Proof. Suppose there is some $g_1, g_2, \dots$ which is infinite and autoreduced. Then there is some x_j such that infinitely manu leader have the form θx_j for some $\theta \in \Theta$. By the Lemma 6.2.3.1, there is not infinite sequence $\theta_1, \theta_2, \dots$ in Θ with $\theta_i \nmid \theta_{i+j}$ for $i \geq 1$ and $j > 0$. This means that $\theta_3(\theta_1(x_j)) = \theta_2(x_j)$ for some $\theta_1 x_j$ and $\theta_2 x_j$ in the sequence. This implies the set is not autoreduced. $\square$

We now define a partial ordering on the collection of autoreduced sequences for the purposes of showing minimal ones exist. These will be the characteristic sequences.

Definition 6.2.3.8. ♠♠♠ Taylor: [double check] Let $G = (g_1, \dots, g_r)$ and $G = (h_1, \dots, h_s)$ be two autoreduced sequences with respect to some ranking. We say that $H \prec G$ if and only if

1. Reading from left to right we find some j such that $h_j \prec g_j$ (we will assume that this j is minimal and that for all of $i < j$ this does not hold)

2. They match all the way but $s > r$. In other words, H has more elements.

We now give the well-ordering property.

Theorem 6.2.3.9. *Any set of autoreduced sequences has a minimal element.*

Proof. Let $\mathcal{G}$ be a collection of autoreduced sequences. Form the sequences $\mathcal{G}_0 = \mathcal{G}$ and

$$\mathcal{G}_i = \{(g_1, g_2, \dots, g_i, \dots) \in \mathcal{G}_{i-1} \colon |G| \geq i,\ g_i \text{ minimal }\}, \qquad i \geq 1.$$

Note that minimality makes sense since the collection of ranks

$$\{\ell_{g_i}^{\mathrm{ldeg}(g_i)} \colon (g_1, \dots, g_i, \dots) \in \mathcal{G}_{i-1}\}$$

is either empty or has a minimal element. If for every i the set $\mathcal{G}_i$ is not empty then we could construct an infinite autoreduced sequence. This is impossible.

Hence there exists some $N \geq 0$ such that $\mathcal{G}_N \neq \emptyset$ and $\mathcal{G}_{N+1} = \emptyset$. At each stage g_i was constructed to be minimal. There will be no other autoreduced set below it by "breaking an characterizer" and there will be no autoreduced sets which are longer, any $G \in \mathcal{G}_N$ will be a minimal. □

Definition 6.2.3.10. Fix a Δ-ideal I in $K[x_1, \dots, x_n]_\Delta$ and some ranking. A *characteristic sequence* $G = (g_1, \dots, g_r)$ for I is a minimal (with respect to the ordering $\prec$ where longer sequences are "smaller") autoreduced sequence of elements from I.

6.2.4 Ideal Membership

Theorem 6.2.4.1. *Let I be a Δ-ideal in $K[x_1, \dots, x_n]_\Delta$. Fix a ranking on $K[x_1, \dots, x_n]_\Delta$. The following are equivalent for an ordered set $G = (g_1, \dots, g_r)$ of elements from I,*

1. *The sequence G is a characteristic sequence of I.*

2. *$f \in I$ if and only if $\mathrm{red}_G(f) = 0$.*

Proof. We will show that G is not minimal implies there exists some nonzero f reduced with respect to G. This is the contrapositive of the "membership characterization property" implying that G is a characteristic set. Suppose that $G = (g_1, \dots, g_r)$ is an autoreduced subset of I which is not minimal. Let $H = (h_1, \dots, h_s)$ be a characteristic set below G. The relationship $H \prec G$ can happen in two ways.

First there exists some i with $1 \le i \le r$ such that

$$\mathrm{rk}(h_1) = \mathrm{rk}(g_1), \quad \mathrm{rk}(h_2) = \mathrm{rk}(g_2), \quad \dots \quad \mathrm{rk}(h_{i-1}) = \mathrm{rk}(g_{i-1}), \quad h_i \prec g_i.$$

This h_i will be reduced with respect to G as reducedness with respect to H imply this for $g_1, \dots, g_{i-1}$ and being lower implies this for $g_i, \dots, g_r$.

The second way if for $\mathrm{rk}(h_j) = \mathrm{rk}(g_j)$ for $1 \le i \le r$ but $r < s$. In this case h_{r+1} produces our element which is reduced with respect to G.

Conversely suppose that G is a characteristic set for some ranking $\prec$ and, for the sake of contradiction. that $f \in I$ is reduced with respect to G and but not zero.

There are two cases. In the first case there exists some minimal j with $1 \le j \le r$ such that $f \prec g_j$. In this case we have

$$(g_1, \dots, g_{j-1}, f) \prec G,$$

which contradicts minimality.

In the second case we have $f \not\prec g_j$ for $1 \le j \le r$. We then construct

$$(g_1, \dots, g_r, f) \prec G$$

which again contradicts minimality. $\square$

6.3 Prime Decomposition

Chapter 7

The Poincaré-Fuchs Theorem

This section is motivated by the following problem.

Problem 7.0.0.1. 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.

A basic reference for the foliation theoretic proof is Pan and Sebastian [PS04]. There is also a great YouTube Lecture by Loray on the topic.

A basic reference for the differential algebraic proof is Matsuda's book [Mat80] and is based of his paper [Mat78]. I will follow the treatment from Buium's book [Bui86] where he further develops this theory in higher dimensions.

Both [PS04] and [Mat78] give with the stated goal of making Poincaré's original proof rigorous. See [Mat78] for details.

Chapter 8

Differential Galois Theory

This section is motivated by the following problem.

Problem 8.0.0.1. 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 9

The Poincaré Problem

This section is motivated by the following problem.

Problem 9.0.0.1. 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 10

The Kolchin Irreducibility Theorem

Chapter 11

Dimension Theory

This section is motivated by the following problem.

Problem 11.0.0.1. Let (K, Δ) be a Δ-field. Let $u_1, \ldots, u_n \in K\{x_1, \ldots, x_n\}_\Delta$ and consider the system of PDEs

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

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

Appendix A

Analytic Aspects of Differential Equations

A.1 Solutions of Homogeneous Differential Equations

The following is a useful formula.

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

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

provides a local fundamental matrix.

A.2 Cauchy-Kovalevskya (Cauchy-Kowalevski)

Appendix B

Manifolds

In this section we collect the results we need regarding smooth manifolds, schemes, and algebraic varieties.

B.1 Ehresmann's Theorem

The following Theorem will be used when discussing local systems. A consequence of this is that all elliptic curves are drawn topologically as donuts. In fact, all smooth complex algebraic varieties in families when viewed as smooth manifolds are topologically the same. They have the same hodge and betti numbers, fundamental groups, etc. To establish this we use Ehresmann's Theorem.

Theorem B.1.0.1 (Ehresmann's Theorem)**.** *Any proper submersions $f : X \to S$ of smooth manifolds is locally (diffeomorphically) trivial.*

We will remind the reader of the terms.

- A morphism $f : X \to S$ is *proper* if and only if the inverse image of a compact set is compact.[1]

[1]for those familiar with algebraic geometry, you know that you can rephrase this in terms of universal closedness and separatedness which can be formulated diagramatically.

- A morphism $f : X \to S$ is a *submersion* if and only if for every $x \in X$ the induced map $(TX)_x \to (TS)_{f(x)}$ is surjective.

- A morphism $f : X \to S$ is locally trivial if and only if for every $s_0 \in S$ there exists a $U \ni s_0$ open and an diffeomorphism ψ fitting into the diagram

$$\begin{tikzcd} f^{-1}(U) \arrow[rr, "\psi"] \arrow[dr, "f"'] & & U \times X_{s_0} \arrow[dl, "\mathrm{pr}_1"] \\ & U & \end{tikzcd} .$$

  In the above diagram $X_{s_0} = f^{-1}(s_0)$ is the fiber above s_0.

Sketch Proof. A doodle of the proof is found in Figure B.1. Without loss of generality we can assume that $U = \mathbb{R}^n$ with coordinates y^i for $1 \leq i \leq n$.

- We use the submersion property to show that there exists vector fields w_i on X lifting the vector fields $v_i = \frac{\partial}{\partial y^i}$. If $\phi^t_{v_i}$ and $\phi^t_{w_i}$ denote the flows (partially defined self maps $Y \to Y$ and $X \to X$ respectively) given by solving the flow ODE with the vector fields v_i and w_i respectively satisfying

$$\phi^t_{v_i} \circ f = f \circ \phi^t_{w_i}$$

  for some time t.

- We use the properness property to show that for each $s \in U$, the set

$$f^{-1}\{\phi^t_{v_i}(s) \colon t \in [a, b]\}$$

  is a compact set for $[a, b]$ where the map is defined. This allows us to argue that the flows $\phi^t_{w_i}$ need to be defined for just as many $t \in [a, b]$ flows $\phi^t_{v_i}$ are.

- We then define $\psi \colon f^{-1}(U) \to U \times X_{s_0}$ using two maps. The first is projection to U, and the second is flowing to X_{s_0} using the vector fields.

$\square$

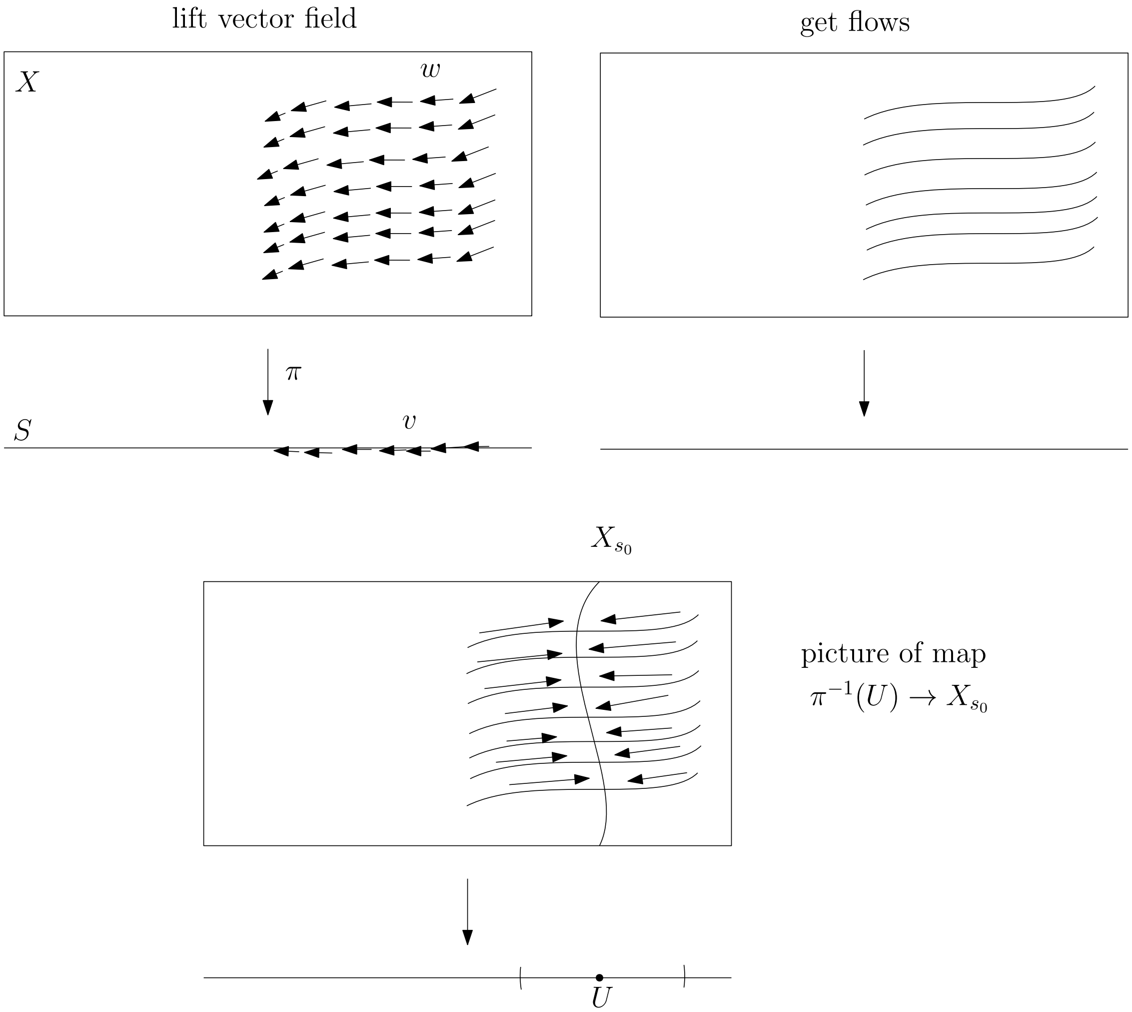

Figure B.1: A picture of the proof of Ehresmann's Theorem. The first frame shows us lifting the vector field. The second frame shows the flow-lines for the vector field. The third frame shows the morphism "go to the fiber" using the flow lines.

Some consequences of this are that the Betti and Hodge numbers of varieties in proper families are preserved.

Another consequence in that proper submersions $f : W \to S$ are actually fiber bundles with a common fiber F over an open subset of S where the fibers are smooth manifolds.

Corollary B.1.0.2 (Lefschetz Fibration)**.** *Let $\pi : W \to S$ be a proper submersion of connected smooth manifolds. Over a dense open subset of S the map π is a fiber bundle in the category of smooth manifolds with fiber W_{s_0}.*

Proof. The open subset of S is the set where W_s is a smooth manifold. We just need to show that the fibers are all diffeomorphic. If U and V are two trivializing opens with ψ_U and ψ_V trivializations to $U \times W_{s_0}$ and $V \times W_{s_1}$ then we have isomorphisms over $U \cap V$ (obtained from base changing our trivializations via $U \cap V \to U$ and $U \cap V \to V$) fitting into the diagram

$$\begin{tikzcd}
(U \cap V) \times W_{s_0} \arrow[dr, "\mathrm{pr}_1"'] & \pi^{-1}(U \cap V) \arrow[l, "\psi_U"'] \arrow[r, "\psi_V"] \arrow[d] & (U \cap V) \times W_{s_1} \arrow[dl, "\mathrm{pr}_1"] \\
& U \cap V &
\end{tikzcd}$$

Now the map $\psi_{U,V} = \psi_V \psi_U^{-1} : (U \cap V) \times W_{s_0} \to (U \cap V) \times W_{s_1}$ over $U \cap V$ which we can base change by any point. This shows that $W_{s_0} \cong W_{s_1}$. We then cover S by trivializing open sets and use the transitive property of diffeomorphism to get a common fiber. $\square$

♠♠♠ Taylor: [Check this smooth fiber condition]

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

[Bou19] François Boulier, *A differential algebra introduction for tropical differential geometry.*

[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

[Bui93] ———, *Effective bound for the geometric Lang conjecture*, Duke Math. J. **71** (1993), no. 2, 475–499. MR 1233446

[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

[For81] Otto Forster, *Lectures on Riemann surfaces*, Graduate Texts in Mathematics, vol. 81, Springer-Verlag, New York-Berlin, 1981, Translated from the German by Bruce Gilligan. MR 648106

[Hat02] Allen Hatcher, *Algebraic topology*, Cambridge University Press, Cambridge, 2002. MR 1867354

[HM82] Michiel Hazewinkel and Clyde F. Martin, *A short elementary proof of Grothendieck's theorem on algebraic vectorbundles over the projective line*, J. Pure Appl. Algebra **25** (1982), no. 2, 207–211. MR 662762

[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

[Kol99] Ellis Kolchin, *Selected works of Ellis Kolchin with commentary*, American Mathematical Society, Providence, RI, 1999, Commentaries by Armand Borel, Michael F. Singer, Bruno Poizat, Alexandru Buium and Phyllis J. Cassidy, Edited and with a preface by Hyman Bass, Buium and Cassidy. MR 1677530

[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

[Mar00] David Marker, *Model theory of differential fields*, Model theory, algebra, and geometry, Math. Sci. Res. Inst. Publ., vol. 39, Cambridge Univ. Press, Cambridge, 2000, pp. 53–63. MR 1773702

[Mat78] Michihiko Matsuda, *Algebraic differential equations of the first order free from parametric singularities from the differential-algebraic standpoint*, J. Math. Soc. Japan **30** (1978), no. 3, 447–455. MR 491639

[Mat80] ______, *First-order algebraic differential equations*, Lecture Notes in Mathematics, vol. 804, Springer, Berlin, 1980, A differential algebraic approach. MR 576060

[MFK94] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, third ed., Ergebnisse der Mathematik und ihrer Grenzgebiete (2) [Results in Mathematics and Related Areas (2)], vol. 34, Springer-Verlag, Berlin, 1994. MR 1304906

[MJD00] T. Miwa, M. Jimbo, and E. Date, *Solitons*, Cambridge Tracts in Mathematics, vol. 135, Cambridge University Press, Cambridge, 2000, Differential equations, symmetries and infinite-dimensional algebras, Translated from the 1993 Japanese original by Miles Reid. MR 1736222

[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

[PS04] Ivan Pan and Marcos Sebastiani, *Les équations différentielles algébriques et les singularités mobiles*, Ensaios Matemáticos [Mathematical Surveys], vol. 8, Sociedade Brasileira de Matemática, Rio de Janeiro, 2004. MR 2101092

[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

[Rit35] J. F. Ritt, *Jacobi's problem on the order of a system of differential equations*, Ann. of Math. (2) **36** (1935), no. 2, 303–312. MR 1503224

[Rit43] ______, *Bézout's theorem and algebraic differential equations*, Trans. Amer. Math. Soc. **53** (1943), 74–82. MR 8300

[Rit50] Joseph Fels Ritt, *Differential Algebra*, American Mathematical Society Colloquium Publications, Vol. XXXIII, American Mathematical Society, New York, N. Y., 1950. MR 0035763

[Sab07] Claude Sabbah, *Isomonodromic deformations and Frobenius manifolds*, french ed., Universitext, Springer-Verlag London, Ltd., London; EDP Sciences, Les Ulis, 2007, An introduction. MR 2368364

[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

[Shi03] Shun Shimomura, *Proofs of the Painlevé property for all Painlevé equations*, Japan. J. Math. (N.S.) **29** (2003), no. 2, 159–180. MR 2035538

[SS83] Mikio Sato and Yasuko Sato, *Soliton equations as dynamical systems on infinite-dimensional Grassmann manifold*, Nonlinear partial differential equations in applied science (Tokyo, 1982), North-Holland Math. Stud., vol. 81, North-Holland, Amsterdam, 1983, pp. 259–271. MR 730247

[Sta22] The Stacks project authors, *The stacks project*, `https://stacks.math.columbia.edu`, 2022.

[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