TOPICS IN NON COMMUTATIVE ALGEBRAIC GEOMETRY AND CONGRUENT ZETA FUNCTIONS (PART IV). THEORY OF CONNECTIONS.

YOSHIFUMI TSUCHIMOTO

Jujitsu? I'm going to learn Jujitsu?

Neo, ca 2199

In this part we present basic tools of "differential geometry on schemes." Since we mainly deal with local theories, we describe them in terms of rings and modules. Even so, students who are interested may note that the technique employed here (except for last few sections where the story is specific for characteristic $p \neq 0$) is also applicable for ordinary differential geometry. In fact, for any differentiable manifold M, we have

- The ring $B = C^\infty(M)$ of C^∞-functions on M plays the role of "affine coordinate ring". that is, if M is compact, the set $\operatorname{Spm} B$ of maximal ideals of B with the Zariski topology is homeomorphic to M. If M is not compact, $\operatorname{Spm}(B)$ is a bit larger, containing "ideal boundaries" of M, but even then B carries virtually all the information to study M.
- The Lie algebra of derivations $\operatorname{Der}_{\mathbb{C}}(B)$ is identified with the Lie algebra of C^∞ vector field on M.
- The module of 1-forms $\Omega^1_{B/\mathbb{C}}$ is identified with the module of C^∞-1-forms over M.

And so on. So the theory here gives results on differential manifold without serious modifications.

1. CONNECTION

1.1. **definition of connections on quasi coherent sheaves.** Let X be a separated scheme over S. Let $\mathcal{V}$ be a quasi coherent sheaf on X. Let $p_1^{(1)}, p_2^{(1)} : \Delta^{(1)}_{X/S} \to X$ be the canonical projections.

There are a several equivalent ways to give a "connection on $\mathcal{V}$".

One way is to provide an isomorphism

$$P : (p_1^{(1)})^* \mathcal{V} \cong (p_2^{(1)})^* \mathcal{V}$$

such that $P|_{\Delta_{X/S}} = \mathrm{id}$. (Since the structure sheaf of $\Delta^{(1)}_{X/S}$ is an extension of that of $\Delta_{X/S}$ by nilpotents, we may easily prove that the term "isomorphism" here may be safely replaced by a term "homomorphism".)

By the adjoint relation, we see that giving P is equivalent to giving an $\mathcal{O}_X$-linear homomorphism

$$D : \mathcal{V} \to (p_1^{(1)})_*(p_2^{(1)})^*\mathcal{V} = \mathcal{J}_1(\mathcal{V})$$

such that the composition

$$\mathcal{V} \xrightarrow{D} \mathcal{J}_1(\mathcal{V}) \to \mathcal{V}$$

is equal to identity.

Now let us call $\nabla = jet_1 - D$ "the covariant derivation". Then ∇ is $\mathcal{O}_S$-linear homomorphism

$$\nabla : \mathcal{V} \to \Omega^1_{X/S} \otimes \mathcal{V}.$$

In terms of the covariant derivation ∇, the $\mathcal{O}_X$-linear is expressed as the following identity.

(Co) $$\nabla(fm) = f\nabla(m) + df \otimes m \quad (f \in \mathcal{O}_X, m \in \mathcal{V})$$

Let us put it in terms of rings and modules. Let $X = \mathrm{Spec}(B)$ and $S = \mathrm{Spec}(A)$. $I = I_\Delta$ be the defining ideal of the diagonal $X \times_S X$.

The $\mathcal{O}_X$-linear homomorphism D corresponds to a B-module homomorphism

$$D : M \to ((B \otimes_A B)/I^2) \otimes_B M.$$

such that

$$1 \otimes m - D(m) \in (I/I^2) \otimes_B M (\cong \Omega^1_{B/A} \otimes_B M)$$

holds for all $m \in M$. The left hand side of the above formula is $\nabla(m)$. Namely,

$$\nabla(m) = jet_1(m) - D(m) = 1 \otimes m - D(m).$$

Let us verify the identity (Co).

$$\begin{aligned}
\nabla(fm) &= 1 \otimes fm - D(fm) \\
&= 1 \otimes fm - fD(m) \\
&= (1 \otimes f - f \otimes 1)m + f(1 \otimes m - Dm) \\
&= df \otimes m + f\nabla(m)
\end{aligned}$$

1.2. **direct sums, tensor products, homomorphisms.** The first definition in the previous subsection is important because one may easily recognize that the following type of lemma should hold.

LEMMA 1.1. *Let A be a commutative ring. Let B be a commutative A-algebra. Let M_1, M_2 be B-modules with connections*

$$\nabla_j : M_j \to \Omega^1_{B/A} \otimes_B M_j \qquad (j = 1, 2).$$

Then we may define a connection on the direct sum $M_1 \otimes M_2$, on the tensor product $M_1 \otimes_B M_2$, and on the module $\mathrm{Hom}_B(M_1, M_2)$. Namely,

$$\nabla(m_1, m_2) = (\nabla_1(m_1), \nabla_2(m_2))$$

$$\nabla(m_1 \otimes m_2) = \nabla_1(m_1) \otimes m_2 + m_1 \otimes \nabla_2(m_2)$$

$$\nabla(\phi)(s) = \nabla_2(\phi(s)) - (\mathrm{id}_{\Omega^1} \otimes \phi)(\nabla_1(s))$$

PROOF. Easy.

□

1.3. **functoriality.**

LEMMA 1.2. *Let C be a commutative ring. Let A, B be commutative C-algebras. Let $\varphi : A \to B$ be a C-algebra homomorphism. Assume we are given an A-module M with a connection*

$$\nabla : M \to \Omega^1_{A/C} \otimes_A M.$$

Then we may define a connection

$$\nabla_{(B)} : B \otimes_A M \to \Omega^1_{B/C} \otimes_B (B \otimes_A M) = \Omega^1_{B/C} \otimes_A M$$

on a pullback $B \otimes_A M$ by defining

$$\nabla_{(B)}(b \otimes m) = db \otimes m + b \cdot (d\varphi \otimes 1)\nabla(m)$$

for all $b \in B$, $m \in M$.

In fact, it is easy to see that the map $\nabla_{(B)}$ is well defined and that it satisfies the rule (Co) of connection. □

2. EXTERIOR ALGEBRAS

Let A be a commutative algebra. Let M be an A-module.
The tensor algebra $\otimes_A M$ is defined as

$$\otimes_A M = A \oplus (\bigoplus_{j=1}^{\infty} M^{\otimes j}).$$

It is a (non commutative) $\mathbb{N}$-graded algebra [1] over A. Let us define the following two sided ideal of $\otimes_A M$.

$$I_{\text{symmetric}} = (x \otimes y - y \otimes x; x \in M, y \in M).$$

$$I_{\text{exterior}} = (x \otimes x; x \in M).$$

Then we define

$$S_A M = (\otimes_A M)/I_{\text{symmetric}}$$

$$\wedge_A M = (\otimes_A M)/I_{\text{exterior}}$$

3. EXTERIOR DERIVATION

Let A be a commutative ring. Let B be a commutative A-algebra. We have already defined the exterior derivation

$$d : B \to \Omega^1_{B/A}.$$

We define

$$\Omega^\bullet_{B/A} = \wedge_B \Omega^1_{B/A}.$$

[2]

We would like to extend this to a map

$$d : \Omega^\bullet_{B/A} \to \Omega^\bullet_{B/A}$$

which satisfies the following rules.

(EXT1) $$d^2 = 0$$

(EXT2)
$$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge (d\beta) \qquad (\forall \alpha \in \Omega^k_{B/A}, \forall \beta \in \Omega^\bullet_{B/A})$$

It is easy to see that d is uniquely determined by the

3.1. exterior derivation on 1-forms.

Let us define an A-module homomorphism

$$\psi : B \otimes_A B \to \Omega^2_{B/A}$$

by the following formula.

$$\psi(b_1 \otimes b_2) = db_1 \wedge db_2.$$

Then we may easily see that

$$\psi((1 \otimes f - f \otimes 1)(1 \otimes g - g \otimes 1)) = 0.$$

---

[1] Note that an $\mathbb{N}$-graded algebra may be viewed as an $\mathbb{Z}$-graded or $\mathbb{Z}/2\mathbb{Z}$-graded algebra in a natural way.

[2] It seems that the definition here is not the best one for general schemes. Hochschild homology may be better. But since we have not done (co)homological arguments, we postpone this topic later.

Thus ψ defines a unique A-module homomorphism

$$d_1 : \Omega^1_{B/A} = I_\Delta/I_\Delta^2 \ni f dg \mapsto df \wedge dg \in \Omega^2_{B/A}$$

3.2. **exterior derivation on general forms.** We make use of the dual number ϵ. That means, we consider an algebra $A_\epsilon = A[\epsilon]/(\epsilon^2)$. We define

$$E = A_\epsilon \otimes_A (\wedge \Omega^1_{B/A}).$$

We assume ϵ is odd. That means, we equip an A-algebra structure on E by introducing the following commutation relations.

$$\epsilon\omega = -\omega\epsilon \quad (\forall \omega \in \Omega^1_{B/A}.)$$

(An equivalent and (probably) easier way to describe E is given by considering a free B-module $B\epsilon$. We then define

$$E = \wedge_B(B\epsilon \oplus \Omega^1_{B/A}).$$

This method is also valid when we deal with the interior derivation.)

Let us then define a map

$$\phi_0 : B \to E$$

by the following formula.

$$\phi_0(x) = (1 + \epsilon d)(x) = x + \epsilon dx.$$

We may easily see that the map ϕ_0 is an algebra homomorphism. We regard E as a B-algebra via this homomorphism. We then define

$$\phi_1 : \Omega^1_{B/A} \to E$$

by

$$\phi_1(\omega) = (1 + \epsilon d)\omega.$$

ϕ_1 is a B-module homomorphism.

So ϕ_0, ϕ_1 together defines an algebra homomorphism

$$\phi : \otimes_A \Omega^1_{B/A} \to E.$$

For any 1-form $\omega \in \Omega^1_{B/A}$, we have

$$\phi(\omega \otimes \omega) = (\omega + \epsilon d\omega) \wedge (\omega + \epsilon d\omega) = 0.$$

So ϕ factors through the exterior algebra and define an algebra homomorphism

$$\phi_{(\bullet)} : \wedge_A \Omega^1_{B/A} \to E.$$

We decompose the homomorphism above as $\phi_{(\bullet)} = 1 + \epsilon d$ and obtain the exterior derivation

$$d : \Omega^k_{B/A} \to \Omega^{k+1}_{B/A}$$

for any non negative integer k. It is easy to verify that the exterior derivation d satisfies the rules (EXT1) and (EXT2).

The theory of exterior derivation may of course be generalized to a theory of that on a separated scheme X over a scheme S.

4. CURVATURE OF CONNECTIONS

4.1. **connections extended.** Let X be a separated scheme over a scheme S. Let $\mathcal{V}$ be a quasi coherent sheaf over X with a connection

$$\nabla : \mathcal{V} \to \Omega^1_{X/S} \otimes_{\mathcal{O}_X} \mathcal{V}$$

on it. We extend ∇ to

$$\nabla : \Omega^\bullet_{X/S} \otimes_{\mathcal{O}_X} \mathcal{V} \to \Omega^\bullet_{X/S} \otimes_{\mathcal{O}_X} \mathcal{V}.$$

by defining

$$\nabla(\alpha \otimes v) = (d\alpha) \otimes v + (-1)^k \alpha \wedge (\nabla v) \qquad (\forall \alpha \in \Omega^k_{X/S}, \forall v \in \mathcal{V}).$$

In other words,

$$\nabla|_{\Omega^k_{X/S} \otimes \mathcal{V}} = d \otimes \operatorname{id} + (-1)^k \operatorname{id} \wedge \nabla.$$

It is easy to verify that ∇ is well-defined and satisfy

$$\nabla(\alpha \wedge x) = (d\alpha) \wedge x + (-1)^k \alpha \wedge (\nabla x) \qquad (\forall \alpha \in \Omega^k_{X/S}, \forall x \in \Omega^\bullet_{X/S} \otimes_{\mathcal{O}_X} \mathcal{V}).$$

4.2. **curvature.** Unlike the exterior derivation d, the square of ∇ may not be zero. Nevertheless, for any $\alpha \in \Omega^k_{X/S}$ and for any $x \in \Omega^\bullet_{X/S} \otimes_{\mathcal{O}_X} \mathcal{V}$, we have

$$\begin{aligned} &\nabla\nabla(\alpha x) \\ =&\nabla(d\alpha \wedge x + (-1)^k f \nabla(x)) \\ =&(-1)^{k+1} d\alpha \wedge \nabla(x) + (-1)^k d\alpha \wedge \nabla(x) + \alpha \nabla\nabla(x)) \\ =&\alpha \nabla\nabla(x). \end{aligned}$$

So we conclude that there exists

$$R \in \Omega^2_{X/S} \otimes_{\mathcal{O}_X} \operatorname{End}_{\mathcal{O}_X}(\mathcal{V}, \mathcal{V})$$

such that

$$\nabla\nabla(x) = R.x$$

holds.

DEFINITION 4.1. R is called the curvature tensor of ∇.

5. some linear algebra

5.1. **even and odd derivations.** Before proceeding further, let us do a little definition and computation on derivations.

Definition 5.1. Let A be a ring. Let B be a $\mathbb{Z}/2\mathbb{Z}$-graded A-algebra. Let M be a $\mathbb{Z}/2\mathbb{Z}$-graded M-module. An A-linear map

$$D : B \to M$$

is said to be

(1) an even derivation if it satisfies

$$D(B_i) \subset (M_i) \qquad (i = 0, 1)$$

and

$$D(fg) = (Df)g + f(Dg) \qquad (\forall f, g \in B)$$

holds.

(2) an odd derivation if it satisfies

$$D(B_i) \subset (M_{i+1}) \qquad (i = 0, 1)$$

and

$$D(fg) = (Df)g + (-1)^s f(Dg) \qquad (\forall f \in B_s,\ s = 1, 2,\ g \in B)$$

holds.

Following [3], let us denote by $\hat{\bullet}$ the "parity" of a homogeneous element $\bullet$. That means,

$$\hat{f} = s \qquad \text{if } f \in A_s, \hat{m} = s \qquad \text{if } m \in M_s,$$

$\hat{D} = 0$ for an even derivation D, and $\hat{D}' = 1$ for an odd derivation D'. Then the "Leibniz rules" of the definition above may be simply rewritten as

$$D(fg) = (Df)g + (-1)^{\hat{D}\hat{f}} f(Dg) \qquad (\forall f : \text{ homogeneous } \in B, \forall g \in B.)$$

For convenience, let us call a map D a **graded derivation** if it is either an even derivation or an odd derivation.

Definition 5.2. Let A be a ring. Let B be a $\mathbb{Z}/2\mathbb{Z}$-graded A-algebra Let $D_1, D_2 : A \to A$ be graded derivations. Then we define their **Lie bracket** by

$$[D_1, D_2] = D_1 D_2 - (-1)^{\hat{D_1}\hat{D_2}} D_2 D_1.$$

Caution: The bracket defined here differs from the ordinary "commutator" if (and only if) both D_1 and D_2 are odd. In such a case it would be better to write $[D_1, D_2]_+$ instead to avoid confusion.

PROPOSITION 5.3. *In the assumption of the definition above, the Lie bracket* $[D_1, D_2]$ *is an even or odd derivation with the parity* $\hat{D}_1 + \hat{D}_2$.

PROOF. We first compute

$$\begin{aligned}
&(D_1 D_2).(fg)\\
=&D_1((D_2.f)g + (-1)^{\hat{D}_2 \hat{f}} f(D_2.g))\\
=&(D_1.D_2.f)g + (-1)^{\hat{D}_1(\hat{D}_2+\hat{f})}(D_2.f)(D_1.g))\\
+&(-1)^{\hat{D}_2\hat{f}}(D_1.f)(D_2.g) + (-1)^{\hat{D}_2\hat{f}}(-1)^{\hat{D}_1\hat{f}} f(D_1.D_2.g).
\end{aligned}$$

Then by adding this equation with the one with D_1, D_2 interchanged, we obtain the required result.

□

The following easy lemma is frequently used.

LEMMA 5.4. *Let* A *be a ring. Let* B *be a* $\mathbb{Z}/2\mathbb{Z}$*-graded algebra. Let* M *be a* $\mathbb{Z}/2\mathbb{Z}$*-graded* B*-module. Then for any graded* A*-derivation* $D : B \to M$*, The kernel of* D *forms an* A*-subalgebra of* B.

□

5.2. **pairing of exterior algebras.** Let A be a commutative ring. Let M, N be A-modules. Assume there is a A-bi-linear pairing

$$\langle \bullet, \bullet \rangle : M \times N \to A.$$

Then we define a pairing of tensor algebras

$$\langle \bullet, \bullet \rangle_\wedge : \otimes_A M \times \otimes_A N \to A$$

defined by the following equation.

$$\langle v_1 \otimes v_2 \otimes \cdots \otimes v_s, w_1 \otimes w_2 \otimes \cdots \otimes w_t \rangle_\wedge = \begin{cases} \det((\langle v_i, w_j \rangle)_{ij}) & (\text{if } s = t) \\ 0 & (\text{if } s \neq t) \end{cases}$$

(The reader may prefer the expression

$$\begin{aligned}
&\langle v_1 \otimes v_2 \otimes \cdots \otimes v_s, w_1 \otimes w_2 \otimes \cdots \otimes w_s \rangle_\wedge \\
=&\sum_{\sigma \in \mathfrak{S}_s} \operatorname{sgn}(\sigma) \langle v_1, w_{\sigma(1)} \rangle \langle v_2, w_{\sigma(2)} \rangle \langle v_3, w_{\sigma(3)} \rangle \dots \langle v_{s-1}, w_{\sigma(s-1)} \rangle \langle v_s, w_{\sigma(s)} \rangle
\end{aligned}$$

rather than the expression above which uses a determinant.)

Then it is easy to see that the pairing above descends to define a pairing of exterior algebras

$$\langle \bullet, \bullet \rangle_\wedge : \wedge_A M \times \wedge_A N \to A.$$

5.3. **interior derivation.** We employ the same assumption of the previous subsection. For any $x \in N$, we define

$$i_x : \wedge_A M \to \wedge_A M$$

to be the unique A-linear map which satisfies the following properties.

(1) $i_x(f) = 0$ for any $f \in A$.
(2) $i_x(m) = \langle m, x \rangle$ for any $m \in M$.
(3) i_x is an odd derivation. That means,

$$i_x(\alpha \wedge \beta) = (i_x \alpha) \wedge \beta + (-1)^s \alpha \wedge (\imath_x \beta) \quad (\forall \alpha \in \wedge_A^s M,\ \forall \beta \in \wedge_A M).$$

The proof of the well-definedness of i_x is similar to the proof of well-definedness of d. We leave the detail to the reader.

Using the above definition we may also explicitly write down a formula for the interior derivations.

$$i_x(m_1 \wedge m_2 \wedge \cdots \wedge m_s) = \sum_{j=1}^{s} (-1)^j \langle m_j, x \rangle (m_1 \wedge m_2 \wedge \cdots \wedge \overset{\text{omit}}{\hat{m}_j} \wedge \cdots \wedge m_s).$$

This leads to an important adjunction relation

$$\langle i_x \alpha, \beta \rangle_\wedge = \langle \alpha, x \wedge \beta \rangle_\wedge \qquad (\forall \alpha \in \wedge M, \forall \beta \in \wedge N).$$

(We may verify the above equation by using the "determinant expansions by minors".)

In particular, we note that for any $x, y \in N$, an anti commutation relation

$$[i_x, i_y]_+ = 0$$

holds. Another useful equation is

$$i_{x_s} i_{x_{s-1}} \ldots i_{x_2} i_{x_1} \alpha = \langle \alpha, x_1 \wedge x_2 \wedge \cdots \wedge x_s \rangle_\wedge \qquad (\forall x_1, x_2, \ldots, x_s \in N, \forall \alpha \in \wedge_A^s M).$$

5.4. **Lie derivation.** Lie derivation appears in a wide contexts. It is based on the following observation. Let A be a commutative ring. Let B be a commutative A-algebra. Let ϵ be the dual number.

$$A_\epsilon = A[\epsilon]/(\epsilon^2), \quad B_\epsilon = B[\epsilon]/(\epsilon^2).$$

Then for any $x \in \mathrm{Hom}_B(\Omega^1_{B/A}, B) = \mathrm{Der}_A(B, B)$, we define

$$\mathrm{id} + \epsilon x : B_\epsilon \to B_\epsilon.$$

It is an A_ϵ-algebra automorphism of B_ϵ which reduces to the identity when we put $\epsilon = 0$. So any natural construction on algebras may be transformed by this map. Among such constructions is the modules of differential forms, modules of derivations, tensor products of them, and the dual of them.

We may also introduce

$$L_x : \Omega^\bullet_{B/A} \to \Omega^\bullet_{B/A}$$

to be the unique A-linear map which satisfies the following properties.

(1) $L_x(f) = x.f = \langle x, df \rangle$ for any $f \in B$.
(2) $L_x(df) = d(\langle x, df \rangle)$ for any $f \in B$.
(3) L_x is an even derivation. That means,

$$L_x(\alpha \wedge \beta) = (L_x \alpha) \wedge \beta + \alpha \wedge (L_x \beta) \quad (\forall \alpha \in \Omega^\bullet_{B/A},\ \forall \beta \in \Omega^\bullet_{B/A}).$$

Other useful Lie derivation is that for vector fields. Namely, Let x, y be two A-derivations from B to B. Then we define

$$L_x(y) = [x, y] (= xy - yx).$$

Lie derivations commute with contractions. Namely,

$$L_x \langle y, df \rangle = \langle L_x(y), df \rangle + \langle y, L_x(df) \rangle \quad (\forall x, y \in \mathrm{Der}_A(B), \forall f \in B).$$

We leave it the reader to do the detailed discussion.

5.5. **Relations of derivations.**

Lemma 5.5. *Let A be a ring. Let B be a commutative A-algebra. For any $x \in \mathrm{Der}_A(B)$, we have*

$$[L_x, d] = 0$$

Proof. This is based on the fact that d is naturally defined. A direct proof is obtained by using Lemma 5.4 and by noting the following facts.

(1)
$$(L_x d) f = L_x(df) = d(x.f) = d L_x f$$
(2)
$$(L_x d) dg = 0, \qquad (d L_x) dg = d(d L_x(g)) = 0$$

□

Lemma 5.6. *Let A be a commutative ring. Let B be a commutative B-algebra.*

(1) *For any $x \in \mathrm{Der}_A(B)$, we have*
$$[d, i_x]_+ = L_x.$$
(2) *For any $x, y \in \mathrm{Der}_A(B)$, we have*
$$[L_x, i_y] = i_{[x,y]}.$$

Proof. Same method as above works. □

5.6. **curvature revisited.** In this subsection we prove an important proposition on curvature. But before we do that, we do two things. Firstly, we prove the following lemma.

LEMMA 5.7. *Let A be a commutative ring. Let B be a commutative A-algebra. Then for any $X, Y \in \mathrm{Der}_A(B)$ and for any $\omega \in \Omega^1_{X/S}$, we have*

$$i_Y i_X(d\omega) = X.\langle\omega, Y\rangle - Y.\langle\omega, X\rangle - i_{[X,Y]}\omega.$$

PROOF.

$$\begin{aligned}
&i_Y i_X d\omega \\
=&i_Y(L_X - di_X)\omega \\
=&i_Y(L_X\omega) - i_Y d\langle\omega, X\rangle \\
=&L_X(i_Y\omega) - i_{[X,Y]}\omega - Y.\langle\omega, X\rangle \\
=&X.\langle\omega, Y\rangle - i_{[X,Y]}\omega - Y.\langle\omega, X\rangle
\end{aligned}$$

□

Secondly, we add a notation.

DEFINITION 5.8. Let A be a commutative ring. Let B be a commutative A-algebra. Let M be a B-module with a connection

$$\nabla : M \to \Omega^1_{B/A} \otimes_B M.$$

Then for any $X \in \mathrm{Der}_A(B)$,

$$(i_X \otimes 1)\nabla : M \to M$$

is an A-linear map. We shall denote it by ∇_X.

PROPOSITION 5.9. *Let A be a commutative ring. Let B be a commutative A-algebra. Let M be a B-module with a connection*

$$\nabla : M \to \Omega^1_{B/A} \otimes_B M.$$

Let R be the connection 2-form of ∇. Then for any $X, Y \in \mathrm{Der}_A(B)$, we have

$$\langle R, X \wedge Y\rangle_\wedge = [\nabla_X, \nabla_Y] - \nabla_{[X,Y]}$$

PROOF. Let v be an element of M. Since ∇v is an element of $\Omega^1_{B/A} \otimes_B M$, we may write it as :

$$\nabla v = \sum_j \omega_j \otimes v_j$$

for some $\omega_j \in \Omega^1_{B/A}, v_j \in M$. Then we compute

$$\nabla\nabla v = \sum_j d\omega_j \otimes v_j - \sum_j \omega_j \wedge \nabla v_j.$$

$$\begin{aligned} & i_Y i_X \nabla\nabla v \\ =& ((i_Y i_X)(d\omega_j)) \otimes v_j - i_Y \sum_j \langle \omega_j, X\rangle \nabla v_j + \sum_j \omega_j \nabla_X v_j \\ =& (i_Y i_X d\omega_j) \otimes v_j - \sum_j \langle \omega_j, X) \nabla_Y v_j + \sum_j \langle \omega_j, Y) \nabla_X v_j \end{aligned}$$

On the other hand, we have

$$\nabla_X(v) = (i_X \otimes 1)\nabla(v)) = \sum_j \langle \omega_j, X\rangle v_j$$

So we may proceed

$$\nabla\nabla_X(v) = \sum_j d\langle \omega_j, X\rangle v_j + \sum_j \langle \omega_j, X\rangle \nabla v_j$$

$$\nabla_Y \nabla_X(v) = \sum_j Y.\langle \omega_j, X\rangle v_j + \sum_j \langle \omega_j, X\rangle \nabla_Y v_j$$

Thus, together with the Lemma above, we see

$$i_Y i_X \nabla\nabla(v) + \nabla_Y \nabla_X(v) - \nabla_X \nabla_Y(v) = -\sum_j i_{[X,Y]}\omega_j \otimes v_j = -\nabla_{[X,Y]} v$$

□

COROLLARY 5.10. *If the curvature R of ∇ is equal to zero, then*

$$\nabla_\bullet : \mathrm{Der}_A \to \mathrm{End}_A(M)$$

is a Lie algebra homomorphism.

6. A FORMULA FOR p-POWERS

6.1. p-powers of derivations.

PROPOSITION 6.1. *Let p be a prime number. Let A be a (not necessarily commutative, but unital, associative) algebra over $\mathbb{F}_p$. Let B be a (not necessarily commutative, but unital, associative) algebra over A. Then for any A-derivation $D : B \to B$, its p-power D^p is also an derivation.*

PROOF. That D^p is A-linear is clear. Let $f, g \in B$. Then for any positive integer k, we have by using the Leibniz rule of D,

$$D^k(fg) = \sum_{j=0}^{k} \binom{k}{j} D^j(f) D^{k-j}(g).$$

In particular, when $k = p$, this means

$$D^p(fg) = D^p(f)g + fD^p(g).$$

Thus D^p also satisfies the Leibniz rule. □

6.2. **Jacobson's formula (p-powers and Lie brackets).** The treatment here essentially follows [1].

Let p be a prime number. Let A be a (not necessarily commutative, but unital, associative) algebra over $\mathbb{F}_p$. We may also regard A as a Lie algebra over $\mathbb{F}_p$, the bracket product being the ordinary commutator. We would like to obtain a formula for

$$(a + b)^p$$

for $a, b \in A$ with the help of Lie brackets. To do that, we first introduce an transcendent element (=variable) T which commutes with any element of A. Then we expand $(Ta + b)^p$ in terms of T. Namely,

$$(Ta + b)^p = T^p a^p + b^p + \sum_{j=1}^{p-1} T^j s_j(a, b) \tag{1}$$

where $s_j(a, b)$ is a non-commutative polynomial in a, b. Our task then is to find a nice formula for s_j.

A first thing to do is to differentiate the equation (1) by T.

$$\sum_{k=0}^{p-1} (Ta + b)^k a (Ta + b)^{p-1-k} = \sum_{j=1}^{p-1} jT^{j-1} s_j(a, b). \tag{2}$$

To compute the left hand side, we use a nice trick. For any element $x \in A$, we denote by $\lambda(x)$ (respectively, $\rho(x)$) an operator defined by the left multiplication (respectively, the right multiplication) of x. That is,

$$\lambda(x) : A \ni f \mapsto xf \in A,$$
$$\rho(x) : A \ni f \mapsto fx \in A.$$

It should be noted that for all $x, y \in A$, $\lambda(x)$ and $\rho(y)$ always commute. Indeed, we have

$$\lambda(x)\rho(y)z = x(zy) = (xz)y = \rho(y)\lambda(x).$$

So from an ordinary result on commutative algebra, we have

$$(\lambda(a)-\rho(a))^p = \lambda(a)^p - \rho(a)^p = \lambda(a^p) - \rho(a^p).$$

We define $\operatorname{ad}(a)$ to be

$$\operatorname{ad}(a) = \lambda(a) - \rho(a).$$

Then the equation above may be written as

$$(\operatorname{ad}(a))^p = \operatorname{ad}(a^p).$$

Another interesting formula is the following.

$$(\lambda(a)-\rho(a))^{p-1} = \sum_{j=0}^{p-1} \lambda(a)^j \rho(a)^{p-1-j}$$

(To verify that it holds, we notice that for any commutative variable T, U, an identity

$$(T-U)^p = T^p - U^p = (T-U)\sum_{j=0}^{p-1} T^j U^{p-1-j}$$

holds.) The above formula may then be rewritten as

$$(\operatorname{ad}(a))^{p-1}(b) = \sum_{j=0}^{p-1} a^j b a^{p-1-j}.$$

By suitable substitutions, we thus have

$$(\operatorname{ad}(Ta+b))^{p-1}(a) = \sum_{j=0}^{p-1} (Ta+b)^j a (Ta+b)^{p-1-j}. \tag{3}$$

comparing the equations (2) and (3), we see that each $s_j(a,b)$ belongs to the Lie sub algebra of A generated by a, b.

To sum up, we have obtained the following proposition.

Proposition 6.2 (Jacobson's formula). *Let p be a prime number. Let A be an algebra over $\mathbb{F}_p$ (which is not necessarily commutative, but unital, associative as we always assume.) Then for any elements $a, b \in A$, we have*

$$(a+b)^p = a^p + b^p + \sum_{j=1}^{p-1} s_j(a,b)$$

where s_j is a universal polynomial in a, b given by the following manner.

$$s_j(a,b) = (1/j)\operatorname{coeff}((\operatorname{ad}(Ta+b))^{p-1}a, T^{j-1}).$$

(Here, $\mathrm{coeff}(\bullet, T^j)$ *denotes the coefficient of* T^j *in* $\bullet$*.) In particular,* $s_j(a,b)$ *belongs to the Lie algebra generated by* a, b*.*

An important corollary is the following.

COROLLARY 6.3. *Let* p *be a prime number. Let* A, B *be algebras over a ring* R *of characteristic* p*. Let* L *be a Lie subalgebra of* A *(that means,* R*-submodule which is closed under commutators.) Let* $\phi : A \to B$ *be a* R*-linear map and assume that*

$$[\phi(x), \phi(y)] = \phi([x, y])$$

holds for any $x, y \in L$*. Let us define a map* $\psi : L \to B$ *by*

$$\psi(x) = \phi(x)^p - \phi(x^p).$$

Then the map ψ *is* p*-linear. That means,* ψ *satisfies*

$$\psi(fx + gy) = f^p\psi(x) + g^p\psi(y) \qquad (\forall x, y \in L, \forall f, g \in R).$$

PROOF. The only thing which needs to be verified is the additivity of ψ.

$$\begin{aligned}
&\psi(x+y)\\
=&(\phi(x)+\phi(y))^p - \phi((x+y)^p)\\
=&\left(\phi(x)^p + \phi(y)^p + \sum_j s_j(\phi(x), \phi(y))\right)\\
&-\left(\phi(x^p) + \phi(y)^p + \sum_j \phi(s_j(x,y))\right)\\
=&\psi(x) + \psi(y)\\
&+\left(\sum_j (s_j(\phi(x), \phi(y)) - \phi(s_j(x,y))\right)\\
=&\psi(x) + \psi(y)
\end{aligned}$$

(Note that each s_j commutes with ϕ because it is built only with commutators.) □

6.3. **appendix.** In this subsection we prove a formula which play a fairly important role in our theory.

PROPOSITION 6.4. *Let* p *be a prime number. Let* D *be a derivation on a commutative algebra* C *of characteristic* p*. Assume that there exists a non commutative algebra* A *which contains* C *as a subalgebra and that there exists an element* $x \in A$ *such that*

$$[x, f] = D(f)$$

holds for all $f \in C$. Then for any element f of C we have

$$(x+f)^p = x^p + D^{p-1}(f) + f^p$$

PROOF. We substitute $a = f$ and $b = x$ in the Proposition 6.2. We need to know $\operatorname{ad}(Tf+x)^{p-1}f$. To do that, first we see by induction that

$$\operatorname{ad}(Tf+x)^k f = D^k f$$

holds for any $k \in \mathbb{N}$. In particular,

$$\operatorname{ad}(Tf+x)^{p-1} f = D^{p-1} f$$

So

$$s_j(f,x) = (1/j)\operatorname{coeff}(\operatorname{ad}(Tf+x)^{p-1}f, T^{j-1}) = \begin{cases} D^{p-1}f & \text{if } j=1 \\ 0 & \text{otherwise.} \end{cases}$$

□

7. p-CURVATURE

In this section, we always assume p to be a prime number.

DEFINITION 7.1. Let A be a commutative ring of characteristic p. Let B be a commutative A-algebra. Let M be a B-module. Let

$$\nabla : M \to \Omega^1_{B/A} \otimes_B M$$

be a connection with zero curvature. Then the p-curvature of ∇ is defined as

$$\operatorname{curv}_p(\nabla)(X) = (\nabla_X)^p - \nabla_{(X^p)}$$

From the argument of the previous section we see that the p-curvature is p-linear.

A fairly good account on p-curvatures is given in [2]. Our treatment here is a bit different. It is not so general, but is easy using only arguments on rings and modules.

DEFINITION 7.2. Let X be a separated scheme over a scheme S. Let $\mathcal{F}$ be a quasi coherent sheaf on X.

$$\nabla : \mathcal{F} \to \Omega^1_{X/S} \otimes_{\mathcal{O}_X} \mathcal{F}$$

be a connection on $\mathcal{F}$. Then a local section s of $\mathcal{F}$ is said to be **(∇-)parallel** If $\nabla s = 0$. $\mathcal{F}$ is said to be **generated by parallel sections** if $\mathcal{F}$ is generated by the set

$$\{s \in \mathcal{F}(X); \nabla s = 0\}$$

of parallel sections as a $\mathcal{O}_X$-module.

$\mathcal{F}$ is said to be **locally generated by parallel sections** if there exists an open covering $\{U_\lambda\}$ of X such that $\mathcal{F}|U_\lambda$ is generated by parallel sections.

PROPOSITION 7.3. *Let S be a scheme on which $p = 0$. Let X be a smooth scheme over a scheme S of relative dimension n. Let $\mathcal{F}$ be a quasi coherent sheaf on X.*

$$\nabla : \mathcal{F} \to \Omega^1_{X/S} \otimes_{\mathcal{O}_X} \mathcal{F}$$

be a connection on $\mathcal{F}$.

Then the following conditions are equivalent.

(1) *$\mathcal{F}$ is locally generated by parallel sections.*
(2) *The curvature and the p-curvature of ∇ are both zero.*

Since the question is local, we may reduce the proposition to an lemma which we describe later. Before we do that, we need some preparation. First we note that when X is smooth of relative dimension n, we may locally choose a set of elements $\{x_1, x_2, \dots, x_n\}$ ("coordinates") such that $\Omega^1_{X/S}$ is freely generated by $\{dx_1, dx_2 \dots, dx_n\}$ over $\mathcal{O}_X$. Let U be an affine open subset of X on which such a local coordinate system exists. Then there exists vector fields

$$\partial/\partial x_1, \quad \partial/\partial x_2, \quad \partial/\partial x_3, \quad \dots, \quad \partial/\partial x_n$$

on U which are dual to $\{dx_j\}_j$. The derivations may also be characterized by the following formula.

$$(\partial/\partial x_j).x_i = \delta_{ij}.$$

Note also that from this observation we see

$$(\partial/\partial x_j)^p = 0.$$

Next, let us set some more notation. We employ the graded lexicographic order on an index set

$$\Lambda = \{0, 1, 2, 3, \dots, p-1\}^n.$$

We define an order reversing map $\Lambda \in K \mapsto \bar{K} \in \Lambda$ as follows

$$\overline{(j_1, j_2, \dots, j_n)} = (p-1-j_1, p-1-j_2, \dots, p-1-j_n)$$

Λ contains a distinguished element $0 = (0, 0, \dots, 0)$, elementary vectors e_i and $\bar{0} = (p-1, p-1, \dots, p-1)$.

Finally, for any $I = (i_1, i_2, \dots, i_n) \in \Lambda$, we define $I!$ by

$$I! = i_1! i_2! i_3! \dots i_n!$$

LEMMA 7.4. *Let A be a commutative ring of characteristic p. Let B be a commutative A-algebra. We assume that B contains elements $\{x_1, x_2, \dots, x_n\}$ ("coordinates") such that $\Omega^1_{B/A}$ is freely generated by $\{dx_1, dx_2 \dots, dx_n\}$ over B. Let M be a B-module. Let*

$$\nabla : M \to \Omega^1_{B/A} \otimes_B M$$

be a connection on M. Then the following conditions are equivalent.

(1) *M is generated by parallel sections.*
(2) *The curvature and the p-curvature of ∇ are both zero.*
(3) *For each element $m \in M$ there exists $\{m_I; I \in \Lambda\}$ such that*

$$m = \sum_{I \in \Lambda} x^I m_I$$

holds.

PROOF. Let M_0 be the set of parallel sections of M.

$$M_0 = \{m \in M; \nabla m = 0\}.$$

(1) $\implies$ (2): We have

$$\nabla\nabla M_0 = 0$$

$$((\nabla_D)^p - \nabla_{D^p})M_0 = 0 \qquad (\forall D \in \mathrm{Der}_A(B)).$$

Since we know that curvature is B-linear and p-curvature is p-linear over B, we deduce that the curvature and the p-curvature are zero.
(3) $\implies$ (1): obvious.
(2) $\implies$ (3): For any $I = (i_1, i_2, \dots, i_n) \in \Lambda$ and for any $v \in M$, we put

$$\partial_I v = (\nabla_{\partial/\partial x_1})^{i_1} (\nabla_{\partial/\partial x_2})^{i_2} (\nabla_{\partial/\partial x_3})^{i_3} \dots (\nabla_{\partial/\partial x_n})^{i_n} v.$$

Note that the condition (1) tells us that

$$\partial_I \partial_J v = \begin{cases} \partial_{I+J} v & \text{if } I + J \in \Lambda \\ 0 & \text{otherwise.} \end{cases}$$

Now we claim :

CLAIM 7.5. *For any $J \in \Lambda$, there exists elements $\{m_I\}_{I \le J}$ of M_0 such that*

$$\partial_{\bar{K}}(m - \sum_{I \le J} x^I m_I) = 0 \qquad (\forall K \le J).$$

Let us prove the claim above by induction on $J \in \Lambda$. For $J = 0$, it is easy to see that

$$\partial_{\bar{0}} m \in M_0.$$

So it is enough to put

$$m_0 = \partial_{\bar{0}} m.$$

Assume now that the claim holds for all $J' < J$. Since Λ is well-ordered set, there exists an index J_0 which is just before J. (That means, J_0 is the largest index which is smaller than J.)

$$\partial_{\overline{K}}(m - \sum_{I \leq J_0} x^I m_I) = 0 \qquad (\forall K < J_0)$$

Let us put

$$\tilde{m} = \partial_{\overline{J}}(m - \sum_{I \leq J_0} x^I m_I).$$

Then for any j, $J - e_j$ is smaller than J so that we have

$$\partial_j \tilde{m} = \partial_{\overline{J - e_j}}(m - \sum_{I \leq J_0} x^I m_I) = 0.$$

Thus we have $\tilde{m} \in M_0$.

Then we put

$$m_J = (J!)^{-1} \tilde{m}.$$

We may easily see that

$$\partial_{\overline{K}}(x^J m_J) = \begin{cases} 0 & \text{if } K < J \\ \tilde{m} & \text{if } K = J \end{cases}$$

holds and thus the claim holds for J. □

It is worthwhile to note that the coefficients $\{m_J\}_J$ in the condition (3) of the Lemma above is unique. Namely,

Corollary 7.6. *Under the same assumption as the Lemma, assume one (hence, all) of the condition of the Lemma holds. Then a map ψ defined by*

$$\psi : \bigoplus_{J \in \Lambda} M_0 \ni (m_J) \mapsto \sum_{I \in \Lambda} x^I m_I \in M$$

is bijective.

Proof. We have already shown that ψ is surjective. Let us prove that ψ is injective. Assume on the contrary that there exists a non-zero $(m_I) \in \operatorname{Ker}(\psi)$. Let J_0 be the maximal index such that $m_{J_0} \neq 0$. Then we have

$$0 = \partial_{J_0}(\psi((m_J))) = \partial_{J_0}(\sum_{I \in \Lambda} x^I m_I) = m_{J_0}.$$

This contradicts the assumption.

□

COROLLARY 7.7. *Let A be a commutative ring of characteristic p. Let B be a commutative A-algebra. We assume that B contains elements $\{x_1, x_2, \dots, x_n\}$ ("coordinates") such that $\Omega^1_{B/A}$ is freely generated by $\{dx_1, dx_2 \dots, dx_n\}$ over B. Let B_0 be the subalgebra of B defined by*

$$B_0 = \{b \in B; db = 0\}.$$

Then every element b of B is written uniquely as

$$b = \sum_{0 \leq j_1, j_2, \dots, j_n \leq p-1} b_{j_1 j_2 j_3 \dots j_n} x_1^{j_1} x_2^{j_2} x_3^{j_3} \dots x_n^{j_n} \qquad (b_{j_1 j_2 \dots j_n} \in B_0).$$

REMARK 7.8. The proof of Corollary 7.6 may be replaced by a direct computation.

To obtain that, first we recall Taylor expansion

$$f(s) = \sum_I \frac{1}{I!} f^{(I)}(t) \cdot (s-t)^I$$

of a polynomial f over (say) $\mathbb{C}$. putting $s = 0$, we obtain

$$f(0) = \sum_I \frac{1}{I!} f^{(I)}(t) \cdot (-t)^I.$$

By analogy we put

$$m_J = \sum_{I \in \Lambda} \frac{1}{I!} (-x)^I (\partial^{I+J} m)$$

Then we see that m_J is parallel.

Indeed, we have the following lemma.

LEMMA 7.9. *We employ the same assumption as the Corollary 7.6. Then for any $m \in M$, an element*

$$\sum_{I \in \Lambda} \frac{1}{I!} (-x)^I \partial^I m$$

is parallel.

PROOF.

$$\begin{aligned}
&\partial_j \sum_{I\in\Lambda} \frac{1}{I!}(-x)^I(\partial^I m)\\
&= -\sum_{I\in\Lambda} \frac{i_j}{I!}(-x)^{I-e_j}(\partial^I m)\\
&\quad + \sum_{I\in\Lambda} \frac{1}{I!}(-x)^I(\partial^{I+e_j} m)\\
&= -\sum_{I\in\Lambda, I\geq e_j} \frac{1}{(I-e_j)!}(-x)^{I-e_j}(\partial^I m)\\
&\quad + \sum_{I\in\Lambda+e_j} \frac{1}{(I-e_j)!}(-x)^{I-e_j}(\partial^I m)
\end{aligned}$$

If $I=(i_1,i_2,\dots,i_n)\in\Lambda+e_j\setminus\Lambda$, then we have $i_j=p$ so that $\partial^I m=0$. (We use the assumption $\mathrm{curv}_p(\nabla)=0$ here.) So we finally have

$$\partial_j \sum_{I\in\Lambda} \frac{1}{I!}(-x)^I(\partial^I m)=0$$

as required.

□

Finally we have

PROPOSITION 7.10. *We put*

$$m_J=\sum_{I\in\Lambda} \frac{1}{I!}(-x)^I(\partial^{I+J} m).$$

Then:

$$m=\sum_{J\in\Lambda} \frac{1}{J!}x^J m_J.$$

Proof.

$$\begin{aligned}
&\sum_{J\in\Lambda}\frac{1}{J!}x^J m_J\\
=&\sum_{J\in\Lambda}\frac{1}{J!}x^J\sum_{I\in\Lambda}\frac{1}{I!}(-x)^I(\partial^{I+J}m)\\
=&\sum_{I,J\in\Lambda}\frac{(-1)^{|I|}}{I!J!}x^{I+J}(\partial^{I+J}m)\\
=&\sum_{K\in\Lambda+\Lambda}\sum_{I,J\in\Lambda,I+J=K}\frac{(-1)^{|I|}}{I!J!}x^K(\partial^K m)
\end{aligned}$$

Since for $K\in\Lambda+\Lambda$, $(\partial_K m)=0$ unless $K\in\Lambda$, the last line is equal to

$$\begin{aligned}
&\sum_{K\in\Lambda}\sum_{I,J\in\Lambda,I+J=K}\frac{(-1)^{|I|}}{I!J!}x^K(\partial^K m)\\
=&\sum_{K\in\Lambda}\sum_{I,J\in\Lambda,I+J=K}\frac{K!(-1)^{|I|}}{I!J!}\frac{1}{K!}x^K(\partial^K m)
\end{aligned}$$

So the problem is reduced to computing the following combinatorial number:

$$c_K=\sum_{I,J\in\Lambda,I+J=K}\frac{K!(-1)^{|I|}}{I!J!}$$

From ordinary binomial theorem, we see easily that

$$\sum_{I,J\in\Lambda,I+J=K}\frac{K!x^I}{I!J!}=\left((1+x_1)^{k_1}(1+x_2)^{k_2}\dots(1+x_n)^{k_n}\right).$$

So we have

$$\begin{aligned}
c_K&=\left((1+x_1)^{k_1}(1+x_2)^{k_2}\dots(1+x_n)^{k_n}\right)|_{x_1=-1,x_2=-1,\dots,x_n=-1}\\
&=\begin{cases}1 & \text{if } K=0\\ 0 & \text{otherwise.}\end{cases}
\end{aligned}$$

This clearly gives the desired result. □

References

[1] N. Jacobson, *Lie algebras*, Interscience publishers, 1962.
[2] M. Katz, *Nilpotent connections and the monodoromy theorem: applications of a result of Turrittin*, Inst. Hautes Etudes Sci. Publ. Math. **39** (1970), 175–232.
[3] Yuri I. Manin, *Topics in non commutative geometry*, Princeton University Press, 1991.