A Brief Proof of the p-linearity of p-curvature

Daniel Schepler

2 March 2004

Fix the following notation: let S be a scheme of characteristic p, $X \to S$ smooth. Let $\mathcal{D}_{X/S}(1)$ denote the structure sheaf of the divided power neighborhood of the diagonal in $X \times_S X$, $\mathcal{D}^n_{X/S}(1)$ the structure sheaf of the nth infinitesimal divided power neighborhood, $\mathcal{P}_{X/S}$ the structure sheaf of $X \times_S X$ (all considered as $\mathscr{O}_X$-algebras). Let J denote the ideal of the diagonal in $\mathcal{P}_{X/S}$, and $\bar{J}$ the ideal of the diagonal in $\mathcal{D}_{X/S}$. Let $x_1, \ldots, x_r$ be a local coordinate system for X, and $\xi^{[k]}$ ($k \in \mathbb{N}^r$) the corresponding basis of $\mathcal{D}_{X/S}(1)$. Finally, let ϵ_i ($i = 1, \ldots r$) denote the canonical basis of $\mathbb{N}^r$.

Lemma 1. *Let $f, g : \mathscr{O}_X \to \mathscr{O}_X$ be PD differential operators of order $\leq m, n$, respectively. Then for any section x of $\mathscr{O}_X$, if we set $\eta = 1 \otimes x - x \otimes 1 \in \bar{J}$, then $(g \circ f)(\eta^{[m+n]}) = f(\eta^{[m]}) g(\eta^{[n]})$.*

Proof. By definition, we calculate $g \circ f$ by first taking $\delta^{n,m}(\eta^{[m+n]}) = \sum_{j=0}^{m+n} \eta^{[j]} \otimes \eta^{[m+n-j]} = \eta^{[n]} \otimes \eta^{[m]}$ in $\mathcal{D}^n_{X/S}(1) \otimes \mathcal{D}^m_{X/S}(1)$. We now apply $\mathrm{id} \otimes f$, which gives $\eta^{[n]}(1 \otimes f(\eta^{[m]}))$. However, since $\eta^{[n]} \in \bar{J}^{[n]}$ and $1 \otimes f(\eta^{[m]}) - f(\eta^{[m]}) \otimes 1 \in \bar{J}$, we see that $\eta^{[n]}(1 \otimes f(\eta^{[m]})) = f(\eta^{[m]}) \eta^{[n]}$ in $\mathcal{D}^{m+n}_{X/S}(1)$. Thus, applying g gives $f(\eta^{[m]}) g(\eta^{[n]})$, as desired. □

Lemma 2. *Let $\partial : \mathscr{O}_X \to \mathscr{O}_X$ be an $\mathscr{O}_S$-derivation, and let $D : \mathscr{O}_X \to \mathscr{O}_X$ be the unique PD differential operator of order ≤ 1 such that $D^\flat = \partial$. Then $D^p(\xi^{[k]}) = 0$ whenever $|k| \geq 2$, except that $D^p(\xi^{[p\epsilon_i]}) = D(\xi^{[\epsilon_i]})^p = (\partial x_i)^p$ ($i = 1, \ldots, r$).*

Proof. Since D^p is a PD differential operator of order $\leq p$, $D^p(\xi^{[k]}) = 0$ whenever $|k| > p$. If $2 \leq |k| \leq p$, but $k \neq p\epsilon_i$ for any i, then in fact $\xi^{[k]}$ is in the image of the natural map $\mathcal{P}_{X/S} \to \mathcal{D}_{X/S}(1)$, so the result follows from the fact that $(D^p)^\flat = \partial^{(p)}$ is an $\mathscr{O}_S$-derivation and thus a differential operator of order ≤ 1. Finally, an easy induction using the previous lemma shows that $D^r(\eta^{[r]}) = D(\eta)^r$ for all r, so in particular setting $r = p$ and $\eta = \xi^{[\epsilon_i]} = 1 \otimes x_i - x_i \otimes 1$ gives the desired result for $k = p\epsilon_i$. □

Corollary 3. *If $\nabla : \mathscr{E} \to \Omega^1_{X/S} \otimes \mathscr{E}$ is an integrable connection, then the p-curvature $\psi(\nabla) : \mathscr{H}om(\Omega^1_{X/S}, \mathscr{O}_X) \to \mathbf{F}_{X*}\mathscr{E}nd_{\mathscr{O}_X}(\mathscr{E})$ is p-linear.*

Proof. Since the statement is local, we may choose a local coordinate system. Let $\partial : \mathscr{O}_X \to \mathscr{O}_X$ be a derivation, D the PD differential operator of order ≤ 1 with $D^\flat = \partial$, and f a section of $\mathscr{O}_X$. Set $E = (fD)^p - f^p D^p$; then by the previous lemma, $E(\xi^{[k]}) = 0$ whenever $|k| \geq 2$. In particular, $E(\xi^{[p\epsilon_i]}) = (fD(\xi^{[\epsilon_i]}))^p - f^p D(\xi^{[\epsilon_i]})^p = 0$. Thus, E is a PD differential operator of order ≤ 1, so $\nabla(E) = \nabla(E^\flat)$. This means that

$$\nabla(fD)^p - f^p \nabla(D)^p = \nabla((f\partial)^{(p)}) - f^p \nabla(\partial^{(p)}), \text{ i.e.}$$
$$\nabla(f\partial)^p - \nabla((f\partial)^{(p)}) = f^p[\nabla(\partial)^p - \nabla(\partial^{(p)})].$$

Similarly, for additivity, let $\partial_1, \partial_2 : \mathscr{O}_X \to \mathscr{O}_X$ be two derivations, and let D_1, D_2 be the PD differential operators of order ≤ 1 with $D_i^\flat = \partial_i$. Set $E = (D_1 + D_2)^p - D_1^p - D_2^p$. Again, $E(\xi^{[k]}) = 0$ whenever $|k| \geq 2$, since for $k = p\epsilon_i$ we calculate $E(\xi^{[p\epsilon_i]}) = (D_1(\xi^{[\epsilon_i]}) + D_2(\xi^{[\epsilon_i]}))^p - D_1(\xi^{[\epsilon_i]})^p - D_2(\xi^{[\epsilon_i]})^p = 0$. Thus, E is a PD differential operator of order ≤ 1. Once more, $\nabla(E) = \nabla(E^\flat)$ implies

$$\nabla(D_1 + D_2)^p - \nabla(D_1)^p - \nabla(D_2)^p = \nabla((\partial_1 + \partial_2)^{(p)}) - \nabla(\partial_1^{(p)}) - \nabla(\partial_2^{(p)}), \text{ i.e.}$$
$$\nabla(\partial_1 + \partial_2)^p - \nabla((\partial_1 + \partial_2)^{(p)}) = [\nabla(\partial_1)^p - \nabla(\partial_1^{(p)})] + [\nabla(\partial_2)^p - \nabla(\partial_2^{(p)})].$$

□