Erratum for the paper Model theory of difference fields, by Z. Chatzidakis and E. Hrushovski, Trans. Amer. Math. Soc. 351 (1999), pp. 2997-3071.

(6.1) The equation $\sigma(x) = x^2 + b$ (char. 0).

Let $E = acl_\sigma(E)$, and consider a solution a of $\sigma(x) = x^2 + b$, where $b \in E$. As we saw in (4.4), $tp(a/E)$ has a unique non-forking extension to any F containing E. We will show that the formula $\sigma(x) = x^2 + b$ is strongly minimal when $b^2 \neq -2\sigma(b)$; if $b^2 = -2\sigma(b)$, there are several possibilities, depending on the action of σ on the primitive roots of 1. We also study the non-orthogonality relation between such formulas, and show that its equivalence classes are least possible.

To determine whether or not $tp(a/E)$ is strongly minimal, we need to study the finite σ-stable extensions of the field $K =_{\text{def}} E(a)_\sigma$.

Suppose that L is a finite Galois extension of K, which is stable under σ, and write $L = K(\alpha)$. Let $i \in \mathbb{Z}$ be such that the minimal polynomial of α over K has its coefficients in $E(\sigma^i(a))$. We know that $\alpha \in K(\sigma(\alpha))$ and $\sigma(\alpha) \in K(\alpha)$. Hence, decreasing i if necessary, we may assume that $\sigma(\alpha) \in E(\sigma^i(a), \alpha)$ and $\alpha \in E(\sigma^i(a), \sigma(\alpha))$. Let $L_1 = E(a, \sigma^{-i}(\alpha))$; then L_1 and K are linearly disjoint over $E(a)$ since $[L_1 : E(a)] = [L : K]$, $\sigma(L_1) \subseteq L_1$, and $\sigma(L_1)(a) = L_1$. We will show that this implies that $b^2 + 2\sigma(b) = 0$.

Let $\mathcal{S}$ be the set of finite E-valuations on $E(a)$ which ramify in L_1; we know that $\mathcal{S}$ is non-empty. For $\alpha \in E$, we write $(a \to \alpha)$ for the valuation on $E(a)$ positive on $(a - \alpha)$. Observe that $(a \to \sigma^{-1}(b))$ is the only (finite) valuation ramifying in $E(\sigma^{-1}(a))$.

Assume that $(a \to \alpha)$ ramifies in L_1, for some $\alpha \neq \sigma^{-1}(b)$; since it does not ramify in $E(\sigma^{-1}(a))$, it has exactly two extensions to $E(\sigma^{-1}(a))$ and they both ramify in $\sigma^{-1}(L_1)$; these extensions can be described as: $(\sigma^{-1}(a) \to \pm\sqrt{\alpha - \sigma^{-1}(b)})$; taking images by σ, the two valuations $(a \to \pm\sqrt{\sigma(\alpha) - b})$ ramify in L_1.

Assume now that $(a \to \sigma^{-1}(b))$ ramifies in L_1, with ramification index $e \neq 2$; then the extension $\sigma^{-1}(L_1)$ of $E(\sigma^{-1}(a))$ is ramified at the unique extension of $(a \to \sigma^{-1}(b))$, i.e. at $(\sigma^{-1}(a) \to 0)$, with ramification index $\gcd(e, 2)$. Thus $(a \to 0)$ ramifies in L_1.

Moreover, if $(a \to \alpha)$ ramifies in L_1, then its restriction $(a^2 + b \to \alpha^2 + b)$ to $E(\sigma(a))$ ramifies in $\sigma(L_1)$ because $\sigma(L_1)$ and $E(a)$ are linearly disjoint over $E(\sigma(a))$. Applying σ^{-1}, we obtain that $(a \to \sigma^{-1}(\alpha)^2 + \sigma^{-1}(b))$ ramifies in L_1.

We define an oriented graph relation I on $\mathcal{T} =_{\text{def}} \{\alpha \mid (a \to \alpha) \in \mathcal{S}\}$ as follows: $\alpha I \beta \iff \sigma(\alpha) = \beta^2 + b$. If $\alpha I \beta$, then we call α a predecessor of β and β a successor of α. We have therefore shown that every element of $\mathcal{T}$ has a unique predecessor, and that if $\alpha \in \mathcal{T}$ and $\alpha \neq \sigma^{-1}(b)$, then α has two distinct successors: $\pm\sqrt{\sigma(\alpha) - b}$. If $\sigma^{-1}(b) \in \mathcal{T}$, then either $0 \notin \mathcal{T}$ or 0 is the unique successor of $\sigma^{-1}(b)$ in $\mathcal{T}$.

Let $\alpha \in \mathcal{T}$, and look at a maximal (oriented) path γ of $\mathcal{T}$ passing through α. Because $\mathcal{T}$ is finite and every element has a unique predecessor, this path must start with a loop. Moreover, if $\sigma^{-1}(b) \notin \gamma$, then γ has no terminal endpoint, and the uniqueness of predecessors then implies that γ must be a simple loop. If γ_0 is a maximal path containing $\sigma^{-1}(b)$, then either $0 \notin \mathcal{T}$, which implies that γ ends at $\sigma^{-1}(b)$, or $\sigma^{-1}(b) \neq 0 \in \mathcal{T}$ and then γ_0 must be a simple loop, or $\sigma^{-1}(b) = 0$ and then γ_0 consists of the vertex 0 (and we have $0I0$). Hence, there is at most one maximal path which is not a loop, the path ending with

$\sigma^{-1}(b)$, and all other maximal paths are simple loops; this implies that distinct maximal paths are disjoint, and therefore that $\mathcal{T}$ is the disjoint union of its maximal paths.

Let us denote by γ_0 the maximal path passing through $\sigma^{-1}(b)$ (which is empty by convention if $\sigma^{-1}(b) \notin \mathcal{T}$). Let $\gamma \neq \gamma_0$ be a maximal path in $\mathcal{T}$, and let $\alpha \in \gamma$. Then α has two distinct successors β and $-\beta$, and since γ is a simple loop, it can only contain one of them; then the maximal path containing the other intersects γ, which gives us a contradiction. Hence the only maximal path on $\mathcal{T}$ is γ_0.

Assume that $\alpha \in \gamma_0$, $\alpha \neq \sigma^{-1}(b)$. Then α has two successors: $\beta = \sqrt{\sigma(\alpha) - b}$ and $-\beta$. These two elements must be on γ_0, and they are both successors of α; hence one of them must equal α, which implies that that $\alpha I \alpha$, that γ_0 starts with α, and also that $\sigma(\alpha) = \alpha^2 + b$. This being true for all $\sigma^{-1}(b) \neq \alpha \in \gamma_0$, we obtain that γ_0 has at most two elements and is contained in $\{\sigma^{-1}(b), -\sigma^{-1}(b)\}$. If $|\gamma_0| = 1$, then $\sigma^{-1}(b)$ is the only element of γ, so that $\sigma^{-1}(b) I \sigma^{-1}(b)$; hence $b = \sigma^{-1}(b)^2 + b$, and therefore $b = 0$. If $|\gamma_0| = 2$, then γ_0 starts with $-\sigma^{-1}(b)$; hence $(-\sigma^{-1}(b)) I (-\sigma^{-1}(b))$, which implies that $\sigma^2(b) + 2b = 0$. In particular we have shown the first assertion: if $\sigma^2(b) + 2b \neq 0$, then the set $\mathcal{T}$ is empty and the formula $\sigma(x) = x^2 + b$ is strongly minimal.

Case 1. $b = 0$

Then $\mathcal{T} = \{0\}$. Recall that if L_1 is a Galois extension of degree e of the field of rational functions $E(a)$, then there are at least two E-valuations on $E(a)$ which ramify in L_1, and they ramify completely in L_1; moreover, if there are exactly two E-valuations ramifying in L_1, then L_1 is a radicial extension of $E(a)$. Thus $(a \to \infty)$ also ramifies in L_1, and $L_1 = E(\sqrt[e]{a})$ for some integer e, which is odd by linear disjointness of L_1 and K over $E(a)$. Clearly, any extension of this form is σ-stable.

Choose a set $\{c_e\}$ indexed by the odd integers e, and satisfying $c_e^e = a$, $c_{ef}^f = c_f$ for odd integers e, f. Let $L = K(c_e \mid e \text{ odd})$. Then L is a Galois extension of K, which is the union of finite σ-stable extensions of K, and every finite σ-stable extension of K is contained in L.

Also, $\mathcal{G}al(L/K)$ is isomorphic to $\prod_{p \neq 2} \mathbb{Z}_p$. There is an obvious extension σ_1 of σ to L, defined by setting $\sigma_1(c_e) = c_e^2$ for all e. The set of elements of $\mathcal{G}al(L/K)$ which commute with σ_1 is a closed subgroup H of $\mathcal{G}al(L/K)$, which is either infinite or trivial. If H is trivial, then all extensions of σ to L are conjugate by an element of $\mathcal{G}al(L/K)$ by (2.6).

Assume that $H \neq (1)$. Then the projection H_p of H onto some factor $\mathbb{Z}_p$ of $\mathcal{G}al(L/K)$ is non-trivial, and therefore of the form $p^n \mathbb{Z}_p$ for some integer n. Let L_p be the subextension of L with $\mathcal{G}al(L_p/K) = \mathbb{Z}_p$, and let τ be a generator of $\mathcal{G}al(L_p/K)$. Then the conjugates of σ_1 under the action of $\mathcal{G}al(L_p/K)$ are the elements $\tau^{-i} \sigma_1 \tau^i$ for $i = 0, \ldots, p^n - 1$. Moreover, if σ_2 is another extension of σ to K_p, then $\sigma_2 = \tau^j \sigma_1$ for some $j \in \mathbb{Z}_p$, and therefore H_p is also the set of elements of $\mathcal{G}al(L_p/K)$ which commute with σ_2. Hence, any extension of σ to L_p has only finitely conjugates under the action of $\mathcal{G}al(L_p/K)$. This implies that if M is a finite Galois extension of K, then any extension of σ to M will have infinitely many non-conjugate extensions to $L_p M$. From the characterisation of formulas, this implies that there is no strongly minimal formula implying the formula $\sigma(x) = x^2$.

Case 2. $b \neq 0$ and $b^2 + 2\sigma(b) = 0$.

Note that $-b/2$ satisfies $\sigma(x) = x^2$. Also, if c satisfies $\sigma(x) = x^2$, then $c + c^{-1}$ satisfies $\sigma(x) = x^2 - 2$, and $d = (c + c^{-1}) \sigma^{-1}(-b/2)$ satisfies $\sigma(x) = x^2 + b$. This shows the

non-orthogonality of the formulas $\sigma(x) = x^2$ and $\sigma(x) = x^2 + b$. However there is more to it.

Assume that $b^2 = -2\sigma(b)$ is non-zero; by assumption, a satisfies $\sigma(a) = a^2 + b$. Consider the element $e = (-a + \sqrt{a^2 + 2b})/\sigma^{-1}(b)$. Then $e^{-1} = (-a - \sqrt{a^2 + 2b})/\sigma^{-1}(b)$ because $-2b = \sigma^{-1}(b^2)$, and therefore $E(e)$ contains a, and $[E(a)_\sigma(e) : E(a)_\sigma] = 2$. Moreover, one computes that

$$e^2 = \frac{-\sigma(a) + a\sqrt{a^2 + 2b}}{b} \quad \text{and} \quad a^4 + 2a^2b = \sigma(a)^2 - b^2 = \sigma(a^2) + 2\sigma(b),$$

so that $E(a)_\sigma(e)$ contains the elements $-\sigma(a) \pm \sqrt{\sigma(a^2) + 2\sigma(b)}/b$, and is therefore closed under σ. Hence the two possible extensions of σ to $E(a)_\sigma(e)$ are given by : $\sigma(e) = e^2$ and $\sigma(e) = e^{-2}$. This shows that the set of realisations of $\sigma(x) = x^2 + b$ splits into two definable sets: one which is defined by $\exists z\, z^2 = x^2 + 2b \wedge \sigma(z) = zx$ and is related to the set defined by $\sigma(x) = x^2$, and the other defined by $\exists z\, z^2 = x^2 + 2b \wedge \sigma(z) = -zx$ and related to the set defined by $\sigma(x) = x^{-2}$.

We now are concerned with the triviality of such types when $b^2 \neq -2\sigma(b)$. It will follow from:

Lemma. Let E be a difference field, and $a \notin acl_\sigma(E)$ such that $\sigma(a) = a^2 + b$ for some $b \in E$. Assume that $c \in acl_\sigma(Ea) \setminus acl_\sigma(E)$ is such that $\sigma(c) = c^2 + d$ for some $d \in E$. Then

(1) $E(a, c)_\sigma$ is a finite σ-stable extension of $E(a)_\sigma$.
(2) If $b^2 \neq -2\sigma(b)$ and $d^2 \neq -2\sigma(d)$, then $c = e\sigma^k(a)$ for some $k \in \mathbb{Z}$ and e satisfying $\sigma(e) = e^2$, and $d = e^2\sigma^k(b)$.

Proof. (1) Replacing c by an appropriate transform, we may assume that $[E(a, c) : E(a)] = [E(a)_\sigma(c) : E(a)_\sigma]$. Then $[E(a, \sigma(c)) : E(a)] = [E(\sigma(a), \sigma(c)) : E(\sigma(a))] = [E(a, c) : E(a)]$, which implies that $E(a, c) = E(a, \sigma(c))$; from this we deduce that $E(a, c)_\sigma = E(a)_\sigma(c)$.

(2) By (1), we must have $E(a)_\sigma = E(c)_\sigma$, since $E(a)_\sigma$ and $E(c)_\sigma$ have no finite proper algebraic extension stable under σ. We will first show that $E(c) = E(\sigma^k(a))$ for some $k \in \mathbb{Z}$.

Let j be minimal such that $E(c) \supseteq E(\sigma^j(a))$; then $\sigma^{j-1}(a) \notin E(c)$ and therefore $E(\sigma^{j-1}(a)) \cap E(c) = E(\sigma^j(a))$. Replacing c by $\sigma^{-j}(c)$, we may assume that $j = 0$. Let k be minimal such that $E(\sigma^{-k}(a)) \supseteq E(c)$, and assume that $k > 0$. Then $E(c, \sigma^{-1}(a)) \subseteq E(\sigma^{-k}(a))$, and therefore $E(\sigma(c), a) \subseteq E(\sigma^{-k+1}(a))$; we also have: $[E(c) : E(\sigma(c))] = 2 = [E(a) : E(\sigma(a))]$ and $E(\sigma(c)) \cap E(a) = E(\sigma(a))$, which implies that $E(c) = E(a, \sigma(c))$; hence $E(c) \subseteq E(\sigma^{-k+1}(a))$, a contradiction. This implies that $k = 0$.

We may therefore assume that $E(a) = E(c)$; thus for some $\alpha, \beta, \gamma, \delta \in E$ such that $\alpha\delta - \beta\gamma \neq 0$, we have $c = \frac{\alpha a + \beta}{\gamma a + \delta}$. Applying σ, we then obtain:

$$\left(\frac{\alpha a + \beta}{\gamma a + \delta}\right)^2 + d = \frac{\sigma(\alpha)(a^2 + b) + \sigma(\beta)}{\sigma(\gamma)(a^2 + b) + \sigma(\delta)}.$$

These two functions must have the same poles and therefore $(\gamma a + \delta)^2 = \lambda\sigma(\gamma)a^2 + \lambda(\sigma(\gamma)b + \sigma(\delta))$ for some $\lambda \neq 0$. This implies $\gamma^2 = \lambda\sigma(\gamma)$, $\gamma\delta = 0$, and $\delta^2 = \lambda(\sigma(\gamma)b + \sigma(\delta))$. As

$b \neq 0$, we get $\gamma = 0$. Hence, dividing by δ, we may assume that $c = \alpha a + \beta$, with $\alpha \neq 0$. Thus

$$\alpha^2 a^2 + 2\alpha\beta a + \beta^2 + d = \sigma(\alpha)a^2 + \sigma(\alpha)b + \sigma(\beta),$$

which implies $\alpha\beta = 0$, whence $\beta = 0$, and $\alpha^2 = \sigma(\alpha)$, $d = b\alpha^2$.

Corollary. Fix b such that $b^2 \neq -2\sigma(b)$. Then:

(1) The type $\sigma(x) = x^2 + b$ is trivial.

(2) The type $\sigma(x) = x^2 + b$ is orthogonal to all types containing a formula $\sigma(x) = x^2 + d$ with $d^2 = -2\sigma(d)$.

(3) $(\sigma(x) = x^2 + b) \not\perp (\sigma(x) = x^2 + d)$ if and only if $d/\sigma^k(b)$ satisfies $\sigma(x) = x^2$ for some $k \in \mathbb{Z}$.

Proof. (1) Assume by way of contradiction that $\sigma(x) = x^2 + b$ is non-trivial. Then there is an algebraically closed difference field F containing E, and elements a_1, a_2, a_3 not in F and satisfying $\sigma(x) = x^2 + b$, such that a_1 and a_2 are independent over F and $a_3 \in acl_\sigma(F, a_1, a_2) \setminus (acl_\sigma(F, a_1) \cup acl_\sigma(F, a_2))$. By the lemma applied to $acl_\sigma(F, a_2)$, a_1 and a_3, there is an integer k and an element e satisfying $\sigma(x) = x^2$ such that $a_3 = e\sigma^k(a_1)$ and $b = \sigma^k(b)e^2$. Then $e^2 = b\sigma^k(b^{-1}) \in F$, which implies that $a_3 \in acl_\sigma(F, a_1)$ and gives us the desired contradiction.

(2) Clear by (1), because any non-algebraic type containing the formula $\sigma(x) = x^2 + d$ is non-orthogonal to one of $\sigma(x) = x^2$, $\sigma(x) = x^{-2}$, and hence non-trivial.

(3) The left-to-right implication is clear by the lemma and (2). For the converse, assume that $e \in E$ is such that $\sigma(e) = e^2$, and let $d = e\sigma^k(b)$. Then $\sigma(e\sigma^k(a)) = e^2\sigma^k(a^2 + b) = (e\sigma^k(a))^2 + e^2\sigma^k(b)$.