FIELD AUTOMORPHISMS OF R AND $\mathbf{Q}_p$

KEITH CONRAD

1. Introduction

An *automorphism* of a field K is an isomorphism of K with itself: a function $f\colon K \to K$ that is a bijective field homomorphism (additive and multiplicative). For example, the identity function $f(x) = x$ is an automorphism of K. Here are two examples of non-identity automorphisms of fields.

Example 1.1. Complex conjugation is a field automorphism of $\mathbf{C}$. If $f\colon \mathbf{C} \to \mathbf{C}$ by $f(a+bi) = a - bi$ for $a, b \in \mathbf{R}$ then f is a bijection since it is its own inverse ($f(f(z)) = z$). It is additive and multiplicative since

$$\begin{aligned} f((a+bi)+(c+di)) &= f((a+c)+(b+d)i) \\ &= (a+c)-(b+d)i \\ &= (a-bi)+(c-di) \\ &= f(a+bi)+f(c+di) \end{aligned}$$

and

$$\begin{aligned} f((a+bi)(c+di)) &= f((ac-bd)+(ad+bc)i) \\ &= (ac-bd)-(ad+bc)i \\ &= (a-bi)(c-di) \\ &= f(a+bi)f(c+di). \end{aligned}$$

Example 1.2. The set $\mathbf{Q}[\sqrt{2}] = \{a + b\sqrt{2} : a, b \in \mathbf{Q}\}$ is a subfield of $\mathbf{R}$: it's clearly closed under addition and negation, it's closed under multiplication because

$$(a+b\sqrt{2})(c+d\sqrt{2}) = (ac+2bd)+(ad+bc)\sqrt{2}$$

and it's closed under inversion because

$$\frac{1}{a+b\sqrt{2}} = \frac{a-b\sqrt{2}}{(a+b\sqrt{2})(a-b\sqrt{2})} = \frac{a-b\sqrt{2}}{a^2-2b^2} = \frac{a}{a^2-2b^2} - \frac{b}{a^2-2b^2}\sqrt{2}.$$

The denominator $a^2 - 2b^2$ is not 0, since otherwise $2 = (a/b)^2$, which contradicts $\sqrt{2}$ being irrational.

On $\mathbf{Q}[\sqrt{2}]$ we have a conjugation operation: $f(a + b\sqrt{2}) = a - b\sqrt{2}$.[1] That f is an automorphism of $\mathbf{Q}[\sqrt{2}]$ is similar to complex conjugation being an automorphism of $\mathbf{C}$: f

[1] This is well-defined because the rational coefficients of a number in $\mathbf{Q}[\sqrt{2}]$ are unique: if $a+b\sqrt{2} = c+d\sqrt{2}$ then $a = c$ and $b = d$. Indeed, the equation implies $a - c = (d - b)\sqrt{2}$, and if $b \neq d$ then $\sqrt{2} = (a - c)/(d - b)$ would be rational, which is false, so $b = d$ and thus $a = c$.

is bijective since it is its own inverse ($f(f(a+b\sqrt{2})) = a+b\sqrt{2}$), and it is additive and multiplicative since

$$\begin{aligned} f((a+b\sqrt{2})+(c+d\sqrt{2})) &= f((a+c)+(b+d)\sqrt{2}) \\ &= (a+c)-(b+d)\sqrt{2} \\ &= (a-b\sqrt{2})+(c-d\sqrt{2}) \\ &= f(a+b\sqrt{2})+f(c+d\sqrt{2}) \end{aligned}$$

and

$$\begin{aligned} f((a+b\sqrt{2})(c+d\sqrt{2})) &= f((ac+2bd)+(ad+bc)\sqrt{2}) \\ &= (ac+2bd)-(ad+bc)\sqrt{2} \\ &= (a-b\sqrt{2})(c-d\sqrt{2}) \\ &= f(a+b\sqrt{2})f(c+d\sqrt{2}). \end{aligned}$$

As an illustration of the similarities between $\mathbf{R}$ and $\mathbf{Q}_p$, we will show the only automorphism of each field is the identity. In both cases this will be derived from the dense subset $\mathbf{Q}$ having that same property: its only automorphism is the identity. However, bear in mind that even when field operations are continuous, continuity is not a necessary property of a field automorphism. For example, since $\mathbf{Q}[\sqrt{2}]$ is inside of $\mathbf{R}$ we can talk about continuous functions on $\mathbf{Q}[\sqrt{2}]$ and the automorphism f in Example 1.2 is not continuous: there are rational numbers r_n that tend to $\sqrt{2}$, and $f(r_n) = r_n$ does not tend to $f(\sqrt{2}) = -\sqrt{2}$. On the other hand, complex conjugation on $\mathbf{C}$ is continuous.

2. Field automorphisms of $\mathbf{Q}$

Theorem 2.1. *The only field homomorphism $\mathbf{Q} \to \mathbf{Q}$ is the identity. In particular, the only field automorphism of $\mathbf{Q}$ is the identity.*

Proof. Let $f\colon \mathbf{Q} \to \mathbf{Q}$ be a field homomorphism. Since $1^2 = 1$ we get $f(1)^2 = f(1)$, so $f(1)$ is 0 or 1. If $f(1) = 0$ then for all $x \in \mathbf{Q}$ we have $f(x) = f(x \cdot 1) = f(x)f(1) = f(x) \cdot 0 = 0$, so f is identically 0. The zero function is not considered to be a homomorphism of fields, so $f(1) = 1$. Then by induction we get $f(n) = n$ for $n \in \mathbf{Z}^+$, so $f(-n) = -f(n)$ by additivity, and thus $f(n) = n$ for all $n \in \mathbf{Z}$. Finally, for any $r \in \mathbf{Q}$, writing $r = a/b$ with $a, b \in \mathbf{Z}$ implies from $br = a$ that $f(b)f(r) = f(a)$, so $f(r) = f(a)/f(b) = a/b = r$. $\square$

3. Field automorphisms of $\mathbf{R}$

Theorem 3.1. *The only field automorphism $\mathbf{R} \to \mathbf{R}$ is the identity.*

Proof. Let $f\colon \mathbf{R} \to \mathbf{R}$ be a field automorphism.

Step 1: $f(r) = r$ for all $r \in \mathbf{Q}$.

The same reasoning as in the proof of Theorem 2.1 shows $f(1) = 1$ and then from this $f(r) = r$ for all $r \in \mathbf{Q}$.

Step 2: f preserves inequalities.

The key point is that being positive can be described algebraically: $x > 0$ if and only if x is a nonzero square in $\mathbf{R}$. Therefore if $x > 0$, writing $x = y^2$ implies $f(x) = f(y^2) = f(y)^2$, so $f(x) > 0$ since $f(y) \neq 0$. (A field automorphism has $f(0) = 0$ and f is injective, so $y \neq 0 \Longrightarrow f(y) \neq 0$.) If $x > x'$, then $x - x' > 0$ so $f(x - x') > 0$. Since $f(x - x') = f(x) - f(x')$ we get $f(x) - f(x') > 0$, so $f(x) > f(x')$.

Step 3: $f(x) = x$ for all $x \in \mathbf{R}$.

For each $x \in \mathbf{R}$ there are rational numbers r_n and s_n such that $r_n < x < s_n$ for all n and $r_n \to x^-$ and $s_n \to x^+$. Since f preserves inequalities, $f(r_n) < f(x) < f(s_n)$, so $r_n < f(x) < s_n$ since r_n and s_n are rational. Letting $n \to \infty$, $r_n < f(x) \implies x \le f(x)$ and $f(x) < s_n \implies f(x) \le x$, so $f(x) = x$. □

In this proof we did not use surjectivity of f, only that it is a field homomorphism $\mathbf{R} \to \mathbf{R}$. (We did need $y \neq 0 \implies f(y) \neq 0$, but this is true for all field homomorphisms: from $f(1) = 1$ we get for $y \neq 0$ that $1 = y(1/y)$, so $1 = f(1) = f(y)f(1/y)$ and thus $f(y) \neq 0$.) Therefore we have proved a slightly stronger result.

Corollary 3.2. *The only field homomorphism $\mathbf{R} \to \mathbf{R}$ is the identity.*

4. Field automorphisms of $\mathbf{Q}_p$

Theorem 4.1. *The only field automorphism of $\mathbf{Q}_p$ is the identity.*

Proof. Let $f\colon \mathbf{Q}_p \to \mathbf{Q}_p$ be a field automorphism.

Step 1: $f(r) = r$ for all $r \in \mathbf{Q}$.

The argument is the same one used in the proof of Theorem 3.1.

Step 2: If $|x|_p = 1$ then $|f(x)|_p = 1$.

We use the multiplicative decomposition

$$\mathbf{Z}_p^\times = \mu_{p-1} \times (1 + p\mathbf{Z}_p).$$

For $x \in \mathbf{Z}_p^\times$ write $x = \omega v$ where $\omega^{p-1} = 1$ and $v \in 1 + p\mathbf{Z}_p$. Then $f(x) = f(\omega)f(v)$. Since $\omega^{p-1} = 1$ we get $f(\omega)^{p-1} = f(1) = 1$, so $|f(\omega)|_p = 1$. For every $n \in \mathbf{Z}^+$ that is not divisible by p, v is an nth power in $\mathbf{Z}_p$ by Hensel's lemma for the polynomial $X^n - v$ with approximate root 1. When v is an nth power, $f(v)$ is also an nth power ($v = a^n \implies f(v) = f(a)^n$), so $f(v)$ is an nth power for infinitely many positive integers n. Thus $n \mid \operatorname{ord}_p f(v)$ for infinitely many n, so $\operatorname{ord}_p(f(v)) = 0$ and therefore $|f(v)|_p = 1$. Finally, $|f(x)|_p = |f(\omega)|_p |f(v)|_p = 1$.

Step 3: For x and y in $\mathbf{Q}_p$, $|f(x) - f(y)|_p = |x - y|_p$.

This is clear if $x = y$, so assume $x \neq y$. Write $x - y = p^n u$ where $u \in \mathbf{Z}_p^\times$. Then $f(x) - f(y) = f(x-y) = f(p^n u) = f(p)^n f(u)$. By Step 1 we have $f(p) = p$, so $f(x) - f(y) = p^n f(u)$. By Step 2 we have $|f(u)|_p = 1$, so $|f(x) - f(y)|_p = |p^n|_p = 1/p^n = |x - y|_p$.

Step 4: $f(x) = x$ for all $x \in \mathbf{Q}_p$.

For each $x \in \mathbf{Q}_p$ let r_n be a sequence of rational numbers tending p-adically to x, so $|x - r_n|_p \to 0$ as $n \to \infty$. By Step 3 $|x - r_n|_p = |f(x) - f(r_n)|_p$, which equals $|f(x) - r_n|_p$ since $f(r_n) = r_n$ (Step 1). Thus $|f(x) - r_n|_p \to 0$ as $n \to \infty$, so $f(x) = \lim_{n\to\infty} r_n = x$. □

Our proof did not use surjectivity of f, so just as in the real case we really proved a stronger result.

Corollary 4.2. *The only field homomorphism $\mathbf{Q}_p \to \mathbf{Q}_p$ is the identity.*

Remark 4.3. The real numbers lie in the larger field $\mathbf{C}$, which is 2-dimensional over $\mathbf{R}$, but it turns out we can't dig inside $\mathbf{R}$ in a finite-dimensional way: if a field K is contained in $\mathbf{R}$ and $\mathbf{R}$ is finite-dimensional over K then $K = \mathbf{R}$. The proof uses the complex numbers and a piece of algebra called the Artin–Schreier theorem. There is a p-adic analogue: if a field K is contained in $\mathbf{Q}_p$ and $\mathbf{Q}_p$ is finite-dimensional over K then $K = \mathbf{Q}_p$. The proof uses a generalization of the automatic continuity in the proof of Theorem 4.1, together with Galois theory. See `https://math.stackexchange.com/questions/2893911`.

5. Automorphisms of a rational function field

For a field K, we write $K(t)$ for the field of rational functions in one indeterminate with coefficients in K. A field automorphism of $K(t)$ can be created using an invertible linear fractional transformation: $f(t) \mapsto f((at+b)/(ct+d))$ for all $f \in K(t)$, where $a, b, c, d \in K$ and $ad - bc \neq 0$. Such an automorphism of $K(t)$ fixes all the constants (the elements of K) and it can be shown that every field automorphism of $K(t)$ that fixes the elements of K arises in this way.[2]

It is not true in general that a field automorphism of $K(t)$ has to fix all of K, or even map K to K. For example, if $K = F(u)$ for a field F and an indeterminate u, then $K(t) = F(u)(t) = F(t)(u)$ and we get a field automorphism of $K(t)$ by swapping the elements t and u and fixing the elements of F. However, if K is $\mathbf{R}$ or $\mathbf{Q}_p$, then we'll use our earlier work to show the field automorphisms of $K(t)$ must fix all of K, so they are described using linear fractional transformations as indicated above.

Theorem 5.1. *Each field automorphism of $\mathbf{R}(t)$ fixes every element of $\mathbf{R}$.*

Proof. Each real number is an nth power of some real number for infinitely many positive integers n. For example, this is true for odd n. We will use this property to show for each field automorphism $\varphi\colon \mathbf{R}(t) \to \mathbf{R}(t)$ that $\varphi(\mathbf{R}) \subset \mathbf{R}$. Then φ is a field homomorphism $\mathbf{R} \to \mathbf{R}$, so φ fixes all of $\mathbf{R}$ by Corollary 3.2.

Since φ is multiplicative, if $c \in \mathbf{R}$ then $\varphi(c)$ is an nth power for infinitely many $n \geq 1$ (*e.g.*, for odd n). To show $\varphi(c) \in \mathbf{R}$, it suffices to show nonconstant elements of $\mathbf{R}(t)$ are *not* nth powers for infinitely many n.

Let $f(t) \in \mathbf{R}(t)$ with $f(t) \notin \mathbf{R}$. Write f in reduced form as g/h, where g and h are in $\mathbf{R}[t] - \{0\}$ and are relatively prime. They are not both constant, since f is not constant. We'll show that if f is an nth power then $n \leq \max(\deg g, \deg h)$.

Suppose $f(t)$ is an nth power of a rational function written in reduced form as $a(t)/b(t)$, so $a(t)$ and $b(t)$ are relatively prime polynomials and they are not both constant (otherwise $f = (a/b)^n$ would be constant, but f is nonconstant). Since $g/h = (a/b)^n$, clearing denominators gives us $b^n g = a^n h$ in $\mathbf{R}[t]$. From $a^n \mid b^n g$ and a and b being relatively prime, $a^n \mid g$. In a similar way, $b^n \mid h$. If a is nonconstant then g is nonconstant (it's a multiple of a^n), and $a^n \mid g \Rightarrow \deg(a^n) \leq \deg g$, so $n \deg a \leq \deg g$. Thus $n \leq \deg g$ (since $\deg a > 0$). If b is nonconstant, then we get $n \leq \deg h$ in a similar way. At least one of these bounds on n holds, so $n \leq \max(\deg g, \deg h)$. $\square$

The argument in the last paragraph of the previous proof did not depend on the coefficients being in $\mathbf{R}$. For all fields K, no element of $K(t) - K$ can be an nth power for infinitely many n. That will let us carry over part of the previous proof from $\mathbf{R}(t)$ to $\mathbf{Q}_p(t)$.

Theorem 5.2. *Each field automorphism of $\mathbf{Q}_p(t)$ fixes every element of $\mathbf{Q}_p$.*

Proof. As in the previous theorem, it suffices (now by Corollary 4.2) to show a field automorphism φ of $\mathbf{Q}_p(t)$ maps $\mathbf{Q}_p$ to $\mathbf{Q}_p$.

Unlike $\mathbf{R}$, elements of $\mathbf{Q}_p$ might not be nth powers in $\mathbf{Q}_p$ infinitely often: p is not an nth power in $\mathbf{Q}_p$ for $n > 1$. But elements of $1 + p\mathbf{Z}_p$ are nth powers infinitely often, and that will turn out to be enough.

[2] See https://math.stackexchange.com/questions/13129.

Step 1: $\varphi(1+p\mathbf{Z}_p) \subset \mathbf{Q}_p$. Each $v \in 1+p\mathbf{Z}_p$ is an nth power in $1+p\mathbf{Z}_p$ for infinitely many n, so $\varphi(v)$ is an nth power in $\mathbf{Q}_p(t)$ for infinitely many n. Just as in the previous proof, we conclude that $\varphi(v) \in \mathbf{Q}_p$ (in fact, $\varphi(v) \in \mathbf{Z}_p^\times$, but that's not strictly needed here).

Step 2: $\varphi(\mathbf{Q}_p) \subset \mathbf{Q}_p$. This is obvious at 0. For $x \in \mathbf{Q}_p^\times$, write $x = p^n u$ for $u \in \mathbf{Z}_p^\times$ and $u = \omega v$, where $\omega^{p-1} = 1$ and $v \in 1 + p\mathbf{Z}_p$. Then $\varphi(x) = \varphi(p^n \omega v) = \varphi(p)^n \varphi(\omega)\varphi(v)$. By Step 1, $\varphi(v) \in \mathbf{Q}_p$. Why are $\varphi(p)$ and $\varphi(\omega)$ in $\mathbf{Q}_p$?

Since φ fixes every rational number (from fixing 1 and being a field homomorphism), $\varphi(p) = p$. Since $\varphi(\omega)^{p-1} = \varphi(\omega^{p-1}) = \varphi(1) = 1$, $\varphi(\omega)$ is a $(p-1)$-th root of unity. The polynomial $T^{p-1} - 1$ has $p - 1$ roots in $\mathbf{Q}_p$, which matches the degree of the polynomial, so all of the roots of $T^{p-1} - 1$ in a larger field like $\mathbf{Q}_p(t)$ must be its roots in $\mathbf{Q}_p$, and thus $\varphi(\omega) \in \mathbf{Q}_p$. $\square$