# Formally real local rings, and infinitesimal stability.[*]

## Anders Kock

We propose here a topos-theoretic substitute for the theory of formally-real field, and real-closed field. By 'substitute' we mean that the notion is *not* just a *lifting* of the corrsponding classical notion, but at the same time a generalisation which takes into account the mathematical applications of the specific topos-theoretic features of the notion. Thus in [1], it was argued that the good topos theoretic substitute for the notion of *field* is the notion of *local ring* object. Rousseau, in [6], has argued that topos-theoretic results often mathematically are identical to classical results which depend smoothly on a parameter.

We study here properties which depend smoothly on parameters in the sense that they are *infinitesimally stable*: they are not changed by infinitesimal changes of the parameters. More precisely, we study ring-theoretic properties $\phi$ so that if $\phi$ holds for a given object $A$, then $\phi$ also holds for the ring object $A[\varepsilon]$ of dual numbers over $A$. It was precisely the ring-of-dual-numbers that motivated [1]. Clearly, the notion of field is not infinitesimally stable, whereas the notion of local ring is.

## 1    Two basic ring constructions

If $A$ is a commutative ring object in a category $\underline{E}$ with finite products, then there are several ways of making $A \times A$ into a commutative ring object. We are interested in the following two classical ways (in both cases, the additive structure, or even the $A$-module structure, is coordinatewise):

**Ring of dual numbers**: $A[\varepsilon] = A \times A$, with multiplication

$$(a,b) \cdot (c,d) = (a \cdot c, a \cdot d + b \cdot c).$$

Multiplicative unit 1 is $(1,0)$. The element $(0,1)$ is denoted $\varepsilon$; $\varepsilon^2 = 0$.

**Gauss-numbers**: $A[i] = A \times A$, with multiplication

$$(a,b) \cdot (c,d) = (a \cdot c - b \cdot d, a \cdot d + b \cdot c).$$

Multiplicative unit 1 is $(1,0)$. The element $(0,1)$ is denoted $i$; $i^2 = -1$.

Assume now that $\underline{E}$ is a category where coherent logic has a good semantics (say, $\underline{E}$ a topos or a pretopos). Consider any finite coherent formula $\phi(z_1, \ldots, z_n)$

(in the sense of [3], see e.g. [5] §5) about $n$-tuples of elements from rings. Then clearly there is a simple way of constructing a coherent formula $\phi_\varepsilon$ with $2n$ free variables such that for any ring object $A$,

$$A \models \phi_\varepsilon(x_1, y_1, \ldots, x_n, y_n)$$

iff $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ (1.1)

$$A[\varepsilon] \models \phi((x_1, y_1), \ldots, (x_n, y_n))$$

(there are several examples below). Similarly, an $n$-ary formula $\phi$ gives rise to a $2n$-ary formula $\phi_i$ such that (1.1) holds when $\phi_\varepsilon$ and $A[\varepsilon]$ are replaced by $\phi_i$ and $A[i]$, respectively.

Therefore also, if $T$ is a coherent theory of commutative rings, there is a coherent theory $T_\varepsilon$ such that $A \models T_\varepsilon$ iff $A[\varepsilon] \models T$. Similarly with $T_i$: $A \models T_i$ iff $A[i] \models T$.

We say that a theory $T$ is $\varepsilon$-stable or infinitesimally stable if $A \models T$ implies $A[\varepsilon] \models T$, or equivalently, if $T_\varepsilon \subseteq T$. By the well known metatheorem for coherent logic [3] we have in particular:

**Proposition 1.1** *A coherent theory $T$ of commutative rings is $\varepsilon$-stable if and only if, for every $T$-model $A$ in <u>Set</u>, $A[\varepsilon]$ is also a $T$-model.*

An immediate application is

**Proposition 1.2** *The coherent theory $T_L$ of local rings is $\varepsilon$-stable.*

For, if $A$ is a local ring in <u>Set</u>, then so is $A[\varepsilon]$. Note also that no coherent field notion is stable; for $A[\varepsilon]$ is not always (in fact never) a field.

**Remark 1.3** One could similarly talk about $i$-stable properties and theories, but we do not know of any significant example. The notion of local ring is not $i$-stable: First, we note that in $A[i] = A \times A$, an element $z = (x, y)$ is invertible if and only if $x^2 + y^2$ is invertible, and then $z^{-1} = (x^2 + y^2)^{-1} \bar{z}$ where $\bar{z} = (x, -y)$. Next, let $F$ be a field in <u>Set</u> of characteristic $\neq 2$, and suppose there is an element $j \in F$ with $j^2 = -1$ (for instance $F = \mathbb{C}$). Then certainly $F$ is a local ring, but $F[i] = F \times F$ is not. For, $(1, j)$ and $(1, -j)$ are non-invertible since $1^2 + j^2 = 0$, but their sum is $(2, 0)$ which is invertible. So $F[i]$ is not local.

## 2 Some $\varepsilon$-stable theories

Consider for each natural number $n$ the coherent sequent $s_n$:

$$\forall x_1, \ldots, x_n : \bigvee_{i=1}^{n} (x_i \text{ is invertible }) \Rightarrow \sum_{i=1}^{n} x_i^2 \text{ is invertible.}$$

We let $T_{FR}$ denote the coherent theory of local rings whose axioms are the sequents $s_n$. We call $T_{FR}$ the theory of *formally-real* rings.

2

**Proposition 2.1** *If $K$ is a field in <u>Set</u>, then $K \models T_{FR}$ if and only if $K$ is formally real in the classical sense '$-1$ is not a square sum' (cf. e.g. [2] XI.2).*

The proof is straightforward.

**Proposition 2.2** *The theory $T_{FR}$ of formally-real rings is $\varepsilon$-stable.*

**Proof.** By Prop. 1.1, it suffices to consider a formally real ring $A$ in <u>Set</u> and prove that $A[\varepsilon]$ is formally real. Let $(x_i, y_i) \in A[\varepsilon] = A \times A$ for $i = 1, \ldots, n$. Now $(x, y) \in A[\varepsilon]$ is invertible iff $x$ is invertible in $A$. So one of the $(x_i, y_i)$'s is invertible iff one of the $x_i$'s is invertible, which implies that $\sum x_i^2$ is invertible (by formal-realness of $A$). But then also

$$\left( \sum x_i^2, \sum 2x_i y_i \right)$$

is invertible in $A[\varepsilon]$. The displayed element is the square sum of the $(x_i, y_i)$'s.

**Proposition 2.3** *If $A$ is formally real and local, then $A[i]$ is local.*

**Proof** easy, using part of remark 1.3.

Clearly, no coherent theory of algebraically closed field is $\varepsilon$-stable. But the notion of algebraically closed local ring is not $\varepsilon$-stable either (algebraically closed ring means: monic polynomials have roots). For, if it were, $\mathbb{C}[\varepsilon]$ in <u>Set</u> would be algebraically closed local, which it is not, since $\varepsilon$ has no square root.

However, Wraith [7] has displayed a coherent theory $T_{SC}$ of 'separably closed local rings'. It has the property that for a ring $A$ in <u>Set</u>, $A \models T_{SC}$ if and only if $A$ is a Henselian local ring with separably closed residue field (or: $A$ is strictly Henselian, in the terminology of [4], chapter VIII). The *existence* of such a theory has been known for some time, using theorems of Makkai-Reyes, Deligne, and Hakim; cf. [3].

**Proposition 2.4** *The theory $T_{SC}$ of separably closed local rings is $\varepsilon$-stable.*

**Proof.** Again, by Prop. 1.1, it suffices to consider rings in <u>Set</u>. Let $A$ be a Henselian local ring with separably closed residue field $k$. Then $A[\varepsilon]$ is local, and its residue field is also $k$. So we just have to prove that $A[\varepsilon]$ is Henselian. We use the description of this notion given in [4] VII prop. 3 no. 2, so we must prove that for monic polynomials $P(X)$ over $A[\varepsilon]$, simple roots in $k$ lift to $A[\varepsilon]$. Now we have canonical ring maps

$$A[\varepsilon] \xrightarrow{\quad q_2 \quad} A \xrightarrow{\quad q_1 \quad} k.$$

If $P(X)$ is a monic polynomial over $A[\varepsilon]$, we denote its image under $q_2$ and $q_1 \circ q_2$ by $\overline{P}(X)$ and $\overline{\overline{P}}(X)$, respectively. Assume $\overline{\overline{P}}$ has a simple root $\in k$. Since $a$ is Henselian, this root may be lifted to a root $b \in A$ of $\overline{P}(X)$, and $b$ is necessarily a simple root (meaning $\overline{P}'(b)$ is invertible). Now

$$P(X) = \overline{P}(X) + \varepsilon \cdot Q(X).$$

To lift $b$ means to find a $c \in A$ such that $P(b + \varepsilon c) = 0$. Now

$$P(b + \varepsilon c) = \overline{P}(b + \varepsilon c) + \varepsilon Q(b + \varepsilon c)$$
$$= \overline{P}(b) + \varepsilon c \overline{P}'(b) + \varepsilon \cdot Q(b) + \varepsilon(\varepsilon c Q'(b)).$$

The first term vanishes since $\overline{P}(b) = 0$. The last term vanishes since $\varepsilon^2 = 0$. Thus to find $c$ means to solve

$$0 = \varepsilon c \overline{P}'(b) + \varepsilon Q(b)$$

which can be done since $\overline{P}'(b)$ is invertible in $A$. This proves the proposition.

## 3 A substitute for the notion of real-closed field

We shall say that a ring object $A$ is a *separably-real-closed local* ring if $A$ is formally real local, and $A[i]$ is separably closed ($A[i]$ is local by prop. 2.3).

Let, as above, $T_L$, $T_{FR}$, and $T_{SC}$ be the (coherent) theories of local, formally real, and separably-closed local, rings, respectively. Then the theory of separably-real-closed local ring is

$$T_{SRCL} = T_L \cup T_{FR} \cup (T_{SC})_i,$$

and as such, it is a coherent theory.

**Proposition 3.1** *The theory $T_{SRCL}$ is $\varepsilon$-stable.*

**Proof.** The theories $T_L$, $T_{FR}$, and $T_{SC}$ are $\varepsilon$-stable by propositions 1.2, 2.2, and 2.4. The result will now follow from the followin general

**Lemma**  *If $T$ is an $\varepsilon$-stable theory, then so is $T_i$.*

**Proof.** We have

$$(A \models T_i) \Rightarrow (A[i] \models T) \Rightarrow (A[i][\varepsilon] \models T)$$

(by $\varepsilon$-stability of $T$)

$$\Rightarrow (A[\varepsilon][i] \models T) \Rightarrow (A[\varepsilon] \models T_i),$$

since obviously $A[\varepsilon][i] = A[i][\varepsilon]$.

Besides (or related to) the $\varepsilon$-stability of $T_{SRCL}$, a justification of this theory lies in the following conjecture[1]: The Dedekind reals in an elementary topos with NNO satisfy $T_{SRCL}$. A support for this conjecture is

---

[1]Remark (2008): The conjecture was verified by Peter Johnstone in 1978, see the letter at the end of the present file.

4

**Proposition 3.2** *The sheaf R of germs of continuous real-valued functions on a topological space X is a separably-real-closed local ring object.*

**Proof.** We have $R[i] = C =$ sheaf of germs of continuous complex-valued functions. To see that $C \models T_{SC}$, it suffices, since $T_{SC}$ is a coherent theory, to see that for each $x \in X$, $C_x \models T_{SC}$. But $C_x$ is well known to be Henselian (and have $\mathbb{C}$ as residue field), see e.g. [4] VII §4.

We note that $R$ is not a real-closed local ring in the sense of $R[i] = C$ being an algebraically closed local ring. For, if it were, then one could solve $x^2 = \text{id}$ around the origin of $X = \mathbb{C}$, which cannot be done continuously (there is homotopy obstruction).

# 4  Strict order structure

In this section, $A$ will denote a fixed separably real closed local ring object. Any formally real ring, and in particular $A$, is an algebra over the rationals $\mathbb{Q}$; for, $n = 1^2 + \ldots + 1^2$ ($n$ times) and is thus invertible.

We equip $A$ with a binary "Strict order relation" $<$ by posing for arbitrary $a : X \to A$

$$a > 0 \quad \text{iff} \quad \vdash_X \exists y (y^2 = a \text{ and } y \text{ invertible}).$$

We put $a > b$ if $a - b > 0$.

**Proposition 4.1** *The following coherent sentences hold:*

1. $\forall a : a > 0 \Rightarrow a$ *invertible.*

2. $\forall a_1, a_2 : a_1 > 0$ *and* $a_2 > 0$ *implies* $a_1 \cdot a_2 > 0$.

3. $\forall a : a$ *invertible implies* $a > 0 \vee (-a) > 0$

4. $\forall e, f : e > 0$ *and* $f > 0$ *implies* $e + f > 0$ *(and hence* $a_1 > a_2$ *and* $b_1 > b_2$ *implies* $a_1 + b_1 > a_2 + b_2$).

5. $\forall e, f : e + f > 0$ *implies* $e > 0 \vee f > 0$. *(I am indebted to Peter Johnstone for this observation.)*

**Proof.** Again, by coherence, it suffices to prove these in *Set*. The first and second are immediate. To prove the third, consider the monic polynomial $X^2 = a$. Since $A[i]$ is separably closed and 2 is invertible, this polynomial has a root, $x + iy$, say, with $x$ and $y \in A$. So

$$(x + iy)^2 = a$$

that is

$$x^2 - y^2 = a \tag{4.1}$$

and

$$2xy = 0. \tag{4.2}$$

Since $a$ is invertible, we conclude from (4.1) and localness:

$$x \text{ invertible} \quad \text{or} \quad y \text{ invertible},$$

whence from (4.2)

$$y = 0 \quad \text{or} \quad x = 0.$$

If $y = 0$, $x^2 = a$. If $x = 0$, $y^2 = -a$, whence $a > 0$ or $(-a) > 0$, respectively.

To prove 4), assume $x^2 = e$ and $y^2 = f$, with $x$ and $y$ invertible. By $A$ being formal-real, we conclude $e + f$ invertible, so by 3)

$$e + f > 0 \quad \text{or} \quad -(e+f) > 0.$$

We just have to exclude the latter possibility. But $-(e+f) > 0$ implies

$$-(x^2 + y^2) = -(e+f) = z^2$$

for some invertible $z$, whence $x^2 + y^2 + z^2 = 0$, contradicting formal real-ness.

To prove 5): if $e + f > 0$, then $e + f$ is invertible. Since $A$ is a local ring, either $e$ or $f$ is invertible, say $e$ is. Then by 3) either $e > 0$ (in which case we are done), or $(-e) > 0$, whence $f = (e+f) + (-e) > 0$ by 4).

**Corollary 4.2** *The relation $>$ is transitive.*

**Proof.** $((a > b) \text{ and } (b > c))$ implies $((a - b) > 0 \text{ and } (b - c) > 0)$, which in turn implies $((a - b) + (b - c) > 0)$, thus $a - c > 0$, thus $a > c$.

**Corollary 4.3** *If $n$ is a positive natural number, then $n > 0$ and $n^{-1} > 0$ in $A$.*

**Proof.** By prop. 4.1 (3), it suffices to exclude $-n > 0$ and $-n^{-1} > 0$, which is easy.

We now leave the world of coherent logic by introducing the predicate '$\leq$'. We put $b \leq 0$ iff $\forall a : a > 0$ implies $a > b$. Also, put $b \leq c$ if $b - c \leq 0$.

**Proposition 4.3** $\forall a, b : a \leq 0$ *and* $b \leq 0$ *implies* $a + b \leq 0$.

**Proof.** Let $c > 0$. We must prove $a + b < c$. Now $c = \frac{1}{2}c + \frac{1}{2}c$, and $\frac{1}{2}c > 0$ by $c > 0$ and Coroll. 4.3. Thus $(a < \frac{1}{2}c)$ and $(b < \frac{1}{2}c)$, whence $a + b < \frac{1}{2}c + \frac{1}{2}c$ (using Prop. 4.1 (4)). This proof is intuitionistically valid, hence valid in $\underline{E}$ .

Again, it is clear that this Proposition implies the transitive law for $\leq$. Also, $\forall a : a \leq a$, so that $\leq$ is a preorder. We cannot conclude that it is a partial order.

The next Propositions have evident corollaries obtained by adding elements to both sides of the various (strict or nonstrict) inequality signs. We omit these corollaries.

6

**Proposition 4.4** $\forall a, b : a \geq 0$ *and* $b > 0$ *implies* $a + b > 0$.

**Proof.** Since $\forall d : d < 0$ implies $d < a$, we also have

$$\forall d : d < b \quad \text{implies} \quad d < a + b.$$

In particular, this holds for $d = \frac{1}{2}b$, thus $0 < \frac{1}{2}b < a + b$, whence $a + b > 0$ by transitivity of $>$.

**Proposition 4.5** *For all a and b, we have*

1. $a > 0$ *implies* $a \geq 0$

2. $a \geq 0$ *and* $b \geq 0$ *implies* $a \cdot b \geq 0$

3. $a \geq 0$ *and* $b > 0$ *implies* $a \cdot b \geq 0$.

**Proof.** 1) If $a > 0$ and $c < 0$, then by transitivity of $<$, $c < a$. Since this holds for any $c < 0$, $a \geq 0$.

2) The following corrects my erroneous proof in the originally circulated version (June 1977) of the present paper. It depends on the following Proposition, due to Peter Johnstone; this Proposition at the same time refutes a remark to the contrary effect in the June 1977 version.

**Proposition 4.6** *We have*
$$\neg(z > 0) \Leftrightarrow z \leq 0.$$

**Proof.** By Prop. 4.1 (5) we have

$$\forall z, y : z + y > 0 \quad \Rightarrow \quad z > 0 \vee y > 0$$

hence

$$\forall z, y : z + y > 0 \wedge \neg(z > 0) \quad \Rightarrow \quad y > 0;$$

hence

$$\forall z : \neg(z > 0) \quad \Rightarrow \quad [\forall y : z + y > 0 \rightarrow y > 0].$$

But it is easy to see that the formula in the square bracket is equivalent to $z \leq 0$:

$$\forall y : z + y > 0 \Rightarrow y > 0$$
$$\Leftrightarrow$$
$$\forall a : z + (a - z) > 0 \Rightarrow (a - z) > 0$$
$$\Leftrightarrow$$
$$\forall a : a > 0 \Rightarrow a > z.$$

This proves the implication $\Rightarrow$. For the other one, observe that

$$z \leq 0 \wedge (z > 0) \Rightarrow z > z,$$

using the definition of $\leq$. But $z > z$ is false.

**Proof** of Prop. 4.5 (2) and (3). Assume $a \geq 0$ and $b \geq 0$. To prove $a \cdot b \geq 0$, it suffices to prove $\neg(a \cdot b < 0)$. But if $a \cdot b < 0$, $a \cdot b$ is invertible, so in particular, $b$ is; by Proposition 4.6, $b < 0$ is incompatible with $b \geq 0$, so that $b > 0$. Simlarly $a > 0$. Thus $a \cdot b > 0$, by Proposition 4.1 (2), contradicting the assumption $a \cdot b < 0$.

Now (3) in the Proposition follows by combining (1) and (2). Again, the present proofs are intutionistically valid, hence valid in $\underline{E}$ .

Let us finally remark that we cannot conclude $a \leq b$ and $b \leq a$ implies $a = b$. ($\mathbb{R}[\varepsilon]$ in $\underline{Set}$ furnishes a counterexample.)

REFERENCES :
1.   A. Kock, Universal projective geometry via topos theory, Journ. Pure Appl. Algebra 9 (1976), 1-24.
2.   S. Lang, Algebra, Addison-Wesley 1965.
3.   M. Makkai and G.E. Reyes, First order categorical logic, Springer Lecture Notes Vol. 611 (1977).
4.   M. Raynaud, Anneaux locaux Henseliennes, Springer Lecture Notes Vol. 169 (1970).
5.   G.E. Reyes, Sheaves and concepts, Aarhus Preprint Series 75/76 No. 8.
6.   C. Rousseau, Thesis, Université de Montréal 1977.
7.   G.C. Wraith, Intuitionistic Galois Theory, manuscript, University of Sussex 1977. To appear in Proceedings of the Durham Conference on Sheaves and Logic 1977.

AARHUS UNIVERSITY
June 1977/ Corrections February 1979.

# REMARKS ON THE PREVIOUS PAPER

Extracts from two letters from Peter Johnstone to Anders Kock, March 1978

Concerning your conjecture that the Dedekind reals are always a separably real-closed local ring: As you remark (Proposition 3.2) this is true in any spatial topos, from the "classical" fact that the stalks of the sheaf of continuous real- (or complex-) valued functions on a space are always Henselian. In fact this observation is sufficient to prove your conjecture in any Grothendieck topos, for the simple reason that "the generic unramifiable polynomial over $\mathbb{C}$" lives in a spatial topos. Explicitly, let $f$ be a monic polynomial of degree $n$ over $\mathbb{C}$ in a topos $\underline{E}$ . Then the coefficients of $f$ define a geometric morphism $\overline{f} : \underline{E} \to Shv(\mathbb{C}^n)$, and $f$ is unramifiable if the image of $\overline{f}$ is contained in the open subtopos of points in $\mathbb{C}^n$ where at least one of the hyperdiscriminants is nonzero. But over this space, the sheaf of continuous $\mathbb{C}$-valued functions is separably closed, and so we can cover the space with open subsets on which the generic polynomial has a simple root. Pulling back this cover along $\overline{f}$, we get a localization of $\underline{E}$ over which $f$ has a simple root.

Given that this result is true, there probably ought to be a better proof of ot than this. I suspect that in order to get a direct proof we are going to need a formulation of $T_{SRCL}$ which does not menion the ring $A[i]$.

I feel that there ought to be something analogous to the hyperdiscriminants which would tell you (in the classical case) when a given polynomial over a formally real field has at least one simple root in a formally real extension; but so far, I have not found a way of distinguishing between real and complex roots that can be expressed coherently.

Re-reading what I wrote earlier, it occurs to me that perhaps the theory of separably real-closed local ring needs to be written in the language of ordered rings rather than in the language of rings. Define the theory *OLR* of ordered local rings to consists of the theory of rings plus a unary predicate $P$ satisfying
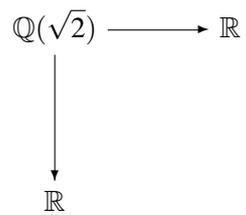
$$P(a) \vdash \exists b(ab = 1)$$
$$\exists b \vdash P(a) \vee P(-a)$$
$$P(a) \wedge P(b) \vdash P(ab) \wedge P(a+b)$$
$$P(a+b) \vdash P(a) \vee P(b)$$
$$P(0) \vdash \text{ false}$$

Then the underlying ring of an OLR is formally real local; conversely, in *Set* every formally real local ring admits an ordering (since we can order its residue field and then pull back). You showed that every SRCL ring admits a unique ordering (the positive elements being the invertible squares). But in a topos it is not true even locally that a formally real local ring can be ordered: consider, in the topos of

diagrams of the form

$$
\begin{array}{ccc}
\bullet & \longrightarrow & \bullet \\
\downarrow & & \\
\bullet & &
\end{array}
$$
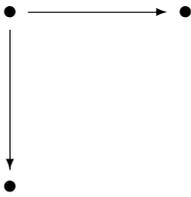
the formally real field

$$
\begin{array}{ccc}
\mathbb{Q}(\sqrt{2}) & \longrightarrow & \mathbb{R} \\
\downarrow & & \\
\mathbb{R} & &
\end{array}
$$

where the two embeddings are different. Now it seems to me that one has rather better chances, in the theory of ordered fields, of saying that a polynomial has a simple root in an *ordered* extension field (not, of course, the same thing as a formally real extension field); for example, with a quadratic one can make the assertion that the discriminant is positive.