Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Advances in Mathematics 190 (2005) 264–277
ADVANCES IN Mathematics
http://www.elsevier.com/locate/aim

Objects of categories as complex numbers

Marcelo Fiore[a,*,1] and Tom Leinster[b]

[a] *Computer Laboratory, University of Cambridge, 15 JJ Thomson Avenue, Cambridge CB3 0FD, UK*
[b] *Department of Mathematics, University of Glasgow, University Gardens, Glasgow G12 8QW, UK*

Received 10 September 2003; accepted 13 January 2004

Communicated by Ross Street

Abstract

In many everyday categories (sets, spaces, modules, etc.) objects can be both added and multiplied. The arithmetic of such objects is a challenge because there is usually no subtraction. We prove a family of cases of the following principle: if an arithmetic statement about the objects can be proved by pretending that they are complex numbers, then there also exists an honest proof.
© 2004 Elsevier Inc. All rights reserved.

Keywords: Category theory; Commutative algebra; Rings and algebras

1. Introduction

Consider the following absurd argument concerning planar, binary, rooted, unlabelled trees [1]. Every such tree is either the trivial tree or consists of a pair of trees joined together at the root, so the set T of trees is isomorphic to $1 + T^2$. Pretend that T is a complex number and solve the quadratic $T = 1 + T^2$ to find that T is a primitive sixth root of unity and so $T^6 = 1$. Deduce that T^6 is a one-element set; realise immediately that this is wrong. Notice that $T^7 \cong T$ is, however, not obviously wrong, and conclude that it is therefore right. In other words, conclude that there is a bijection $T^7 \cong T$ built up out of copies of the original bijection $T \cong 1 + T^2$: a tree is the same as seven trees.

*Corresponding author. Fax: +44-1223-334678.
E-mail address: Marcelo.Fiore@cl.cam.ac.uk (M. Fiore).
[1] Research supported by an EPSRC Advanced Research Fellowship.

0001-8708/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.aim.2004.01.002

The point of this paper is to show that 'nonsense proofs' of this kind are, actually, valid. Our main result is approximately this:

Let p, q_1 and q_2 be polynomials over $\mathbb{N}$. If

$$t = p(t) \Rightarrow q_1(t) = q_2(t)$$

for all complex numbers t, then

$$T \cong p(T) \Rightarrow q_1(T) \cong q_2(T)$$

for all objects T of any category in which it makes sense to add and multiply objects.

This is subject to some restrictions on the three polynomials. For instance (Theorem 17), it suffices to assume that $p(x) - x$ is primitive, has non-zero constant term and degree at least two, and has no repeated complex roots, and that neither q_1 nor q_2 is constant. The last condition is what forbids the conclusion $T^6 \cong 1$ in our example.

The story began with one sentence of Lawvere in 1990 [7, p. 11]:

> I was surprised to note that an isomorphism $x = 1 + x^2$ [...] always induces an isomorphism $x^7 = x$.

Provoked by this, Blass analysed the situation in detail, producing amongst other things an explicit bijection between the set T of trees and the set T^7 of 7-tuples of trees; he called the phenomenon 'Seven Trees in One' [1]. There are many such bijections, none of them particularly intuitive. Each corresponds to a way of building an isomorphism $T \xrightarrow{\sim} T^7$ from a given isomorphism $T \xrightarrow{\sim} 1 + T^2$ using only multiplication and addition. One such (not Blass's) runs as follows: first note that for each $n \geqslant 1$, we may multiply the given isomorphism by T^{n-1} (on the left, say) to obtain an isomorphism $T^n \xrightarrow{\sim} T^{n-1} + T^{n+1}$; use this repeatedly to build a chain of isomorphisms

$$\begin{aligned}
T &\xrightarrow{\sim} 1 + T^2 \xrightarrow{\sim} 1 + T + T^3 \xrightarrow{\sim} 1 + T + T^2 + T^4 \\
&\xrightarrow{\sim} 2T + T^4 \xrightarrow{\sim} 2T + T^3 + T^5 \xrightarrow{\sim} T + T^2 + T^5 \\
&\xrightarrow{\sim} T + T^2 + T^4 + T^6 \xrightarrow{\sim} T + T^3 + T^6 \xrightarrow{\sim} T^2 + T^6 \\
&\xrightarrow{\sim} T^2 + T^5 + T^7 \xrightarrow{\sim} T^2 + T^4 + T^6 + T^7 \xrightarrow{\sim} T^3 + T^6 + T^7 \\
&\xrightarrow{\sim} T^3 + T^5 + 2T^7 \xrightarrow{\sim} T^4 + 2T^7 \xrightarrow{\sim} T^4 + T^6 + T^7 + T^8 \\
&\xrightarrow{\sim} T^5 + T^7 + T^8 \xrightarrow{\sim} T^6 + T^8 \xrightarrow{\sim} T^7.
\end{aligned}$$

We began by trying this method on other polynomials. For example, we considered trees in which each vertex has either one or two branches coming up out of it, the set

T of which satisfies $T \cong 1 + T + T^2$. The complex solutions are $t = \pm i$, which of course satisfy $t^5 = t$, and indeed we were able to build an isomorphism $T^5 \cong T$ in a manner similar to the one for T^7 above. (In fact this example is of special interest: it leads to a 'categorified' or 'objectified' version of the Gaussian integers, as discussed in [3].) More generally, we were able to show that for $n \geqslant 2$,

$$T \cong 1 + T + T^n \Rightarrow T^{2n+1} \cong T$$

and with some effort, found a proof that

$$T \cong 1 + T + T^2 \Rightarrow (1 + T)^9 \cong 16(1 + T).$$

We hoped, of course, that there would turn out to be some general theorem of which all these isolated results were special cases. Our hope was fulfilled, and the subject of this paper is that theorem. We have therefore solved the problem posed by Blass in Section 2 of [1].

Here is the strategy. Our goal is to turn arguments using complex numbers into arguments using only addition and multiplication. Basic commutative algebra (Section 2) shows that subject to some conditions on the polynomials involved, if the implication

$$t = p(t) \Rightarrow q_1(t) = q_2(t)$$

holds for all complex numbers t then it holds for all elements t of all rings. We want to conclude that it holds for all rigs (ri*n*gs without *n*egatives, also known as semirings, Section 3). So the challenge is to discover how to turn a proof that uses subtraction into one that does not. In precise terms, this is a matter of cancellability in the underlying additive semigroup of the quotient rig

$$\mathbb{N}[x]/(x = p(x)).$$

We therefore develop (Section 4) a small amount of general theory of cancellability in semigroups, and using the assumptions on p, q_1 and q_2 we establish (Section 5) the necessary cancellability properties of this particular semigroup. This is the heart of the paper.

We finish (Section 6) by assembling the pieces to give a proof of the main theorem and looking at some further examples.

Our results can be seen as a first step in the study of *algebraic objects* in rig categories; i.e., objects A in categories where it makes sense to add and multiply such that $p_1(A) \cong p_2(A)$ for distinct polynomials p_1, p_2 over $\mathbb{N}$. We consider in particular the situation in which one of the polynomials is the identity (as did Gates in related work [4]). Our main motivation for this is that it is natural from the point of view of computer science. Indeed, it is common practice in programming language theory to specify recursively defined datatypes as fixed points of datatype constructors (see the literature on 'Domain Theory' and 'Denotational Semantics'). Thus, the recursive types $T \cong p(T)$ to which the results of the paper apply are tree-like structures (since p

is required to be a polynomial on $\mathbb{N}$ with degree at least two) that are inhabited (as p should also have non-zero constant term). For details on the computer science perspective, including also a treatment of list-like structures (when p has degree one), see [2].

2. Rings

Our rings will always be equipped with multiplicative identities, but need not be commutative.

Definition 1. Let $p_1, p_2, q_1, q_2 \in \mathbb{Z}[x]$. We say that

$$p_1(x) = p_2(x) \Rightarrow q_1(x) = q_2(x) \textbf{ ring-theoretically}$$

if the following equivalent conditions hold:

(a) for all rings A and all $a \in A$, if $p_1(a) = p_2(a)$ then $q_1(a) = q_2(a)$;
(b) as (a), but restricted to commutative rings;
(c) q_1 and q_2 represent the same element of the quotient ring $\mathbb{Z}[x]/(p_1 - p_2)$;
(d) $(p_1 - p_2)$ divides $(q_1 - q_2)$ in the ring $\mathbb{Z}[x]$.

(Condition (c) implies condition (a) by the universal property of the quotient, and the other implications are trivial.)

As suggested by Blass [1], the first step of 'rehabilitating' nonsense proofs is to rid them of complex numbers and turn them into ring theory. So, we start by assuming that

$$p_1(x) = p_2(x) \Rightarrow q_1(x) = q_2(x)$$

for complex numbers and try to deduce that the same implication holds ring-theoretically. This will not work in general: for instance, each of the implications

$$x = 2 + x + 2x^2 \Rightarrow x = 1 + x + x^2, \tag{1}$$

$$x = 1 + 3x + x^2 \Rightarrow x = 1 + 2x \tag{2}$$

holds for complex numbers but fails ring-theoretically. We therefore seek classes of polynomials for which the deduction is possible.

Proposition 2. *Let $p_1, p_2, q_1, q_2 \in \mathbb{Z}[x]$. Suppose that the polynomial $(p_1 - p_2) \in \mathbb{Z}[x]$ is primitive and has no repeated complex roots, and that each complex root t satisfies*

$q_1(t) = q_2(t)$. *Then*

$$p_1(x) = p_2(x) \Rightarrow q_1(x) = q_2(x) \text{ ring-theoretically}.$$

(Recall that a polynomial over $\mathbb{Z}$ is **primitive** if the greatest common divisor of its coefficients is 1. The following proof is little more than Gauss's Lemma: the product of primitive polynomials is primitive.)

Proof. By the division algorithm,

$$q_1 - q_2 = f \cdot (p_1 - p_2) + g$$

for some $f, g \in \mathbb{Q}[x]$ with $g = 0$ or $\deg(g) < \deg(p_1 - p_2)$. Each of the $\deg(p_1 - p_2)$ complex roots of $p_1 - p_2$ is a root of $q_1 - q_2$, so of g too; hence $g = 0$. If $f = 0$ then certainly $f \in \mathbb{Z}[x]$, and we are done. Otherwise we may write $f = \widetilde{f}/k$ with k a non-zero integer and $\widetilde{f} \in \mathbb{Z}[x]$ a primitive polynomial, and then

$$k \cdot (q_1 - q_2) = \widetilde{f} \cdot (p_1 - p_2),$$

so by Gauss's Lemma $k = \pm 1$ and $f \in \mathbb{Z}[x]$ again. □

Implications (1) and (2) show that the conditions on primitivity and distinctness of roots cannot be dropped.

3. Rigs

A **rig** is a set A equipped with elements 0 and 1 and binary operations $+$ and $\cdot$ such that $(A, 0, +)$ is a commutative monoid, $(A, 1, \cdot)$ is a monoid, and the distributive laws hold:

$$\begin{array}{rclcrcl} 0 & = & a0 & & 0 & = & 0a \\ ab + ac & = & a(b + c) & & ba + ca & = & (b + c)a \end{array}$$

for all $a, b, c \in A$.

Example 3. (a) Any ring is a rig.

(b) The initial rig is the set $\mathbb{N}$ of natural numbers, with its usual arithmetic.

(c) The free rig on one generator is the set $\mathbb{N}[x]$ of polynomials over $\mathbb{N}$.

(d) Any distributive lattice A is a rig: $+$ is least upper bound, 0 is the least element, $\cdot$ is greatest lower bound, and 1 is the greatest element.

One might be tempted to try to turn all rigs into rings by adjoining negatives formally. This can certainly be done (and defines a left adjoint to the inclusion functor **Rings** → **Rigs**), but destroys a lot of information. For example, if A is a

distributive lattice then $a + a = a$ for all $a \in A$, which in the presence of negatives implies $a = 0$, so A collapses to the trivial ring.

(e) The class of all cardinals, with their usual arithmetic, forms a large rig. A set-theoretically respectable version is $\{\text{cardinals} < \kappa\}$, for any infinite cardinal κ.

As with any kind of algebraic structure, it makes sense to talk about quotients of rigs. A **congruence** on a rig A is an equivalence relation $\sim$ on A such that if $\sim$ is regarded as a subset of the product rig $A \times A$ then it is a sub-rig. Explicitly, this means that if $a \sim a'$ and $b \sim b'$ then $a + b \sim a' + b'$ and $ab \sim a'b'$. This is precisely the condition needed on the equivalence relation $\sim$ in order that the set $A/\sim$ of equivalence classes inherits the structure of a rig.

Example 4. (a) The relation 'has the same degree as' defines a congruence $\sim$ on $\mathbb{N}[x]$. We call the quotient $\mathbb{N}[x]/\sim$ the **rig of degrees**. It has countably many elements, conveniently written as $L^{-\infty}, L^0, L^1, L^2, \ldots$, with operations

$$\begin{aligned} L^m + L^n &= L^{\max\{m,n\}}, & L^m \cdot L^n &= L^{m+n}, \\ 0 &= L^{-\infty}, & 1 &= L^0 \end{aligned}$$

$(m, n \in \{-\infty\} \cup \mathbb{N})$. These equations make sense if L is thought of as a large number.

(b) Dually, define the **codegree** of a polynomial $q(x)$ as the least $n \in \mathbb{N}$ for which $q(x)$ has a non-zero coefficient in x^n, or as ∞ if $q = 0$. The quotient of $\mathbb{N}[x]$ by the congruence 'has the same codegree as' is the **rig of codegrees**, which has elements $\varepsilon^{\infty}, \varepsilon^0, \varepsilon^1, \varepsilon^2, \ldots$ and operations appropriate for ε being small.

Any relation on a rig generates a congruence, which can be defined as the intersection of all congruences containing the given relation. We are particularly interested in the congruence $\sim$ on $\mathbb{N}[x]$ generated by declaring equivalent two polynomials p_1 and p_2. This can be described explicitly as follows: $q_1 \sim q_2$ if and only if there is a finite sequence of polynomials

$$q_1 = r_0, r_1, \ldots, r_{n-1}, r_n = q_2$$

(for some $n \in \mathbb{N}$) such that for each $i \in \{1, \ldots, n\}$, there exist $f \in \mathbb{N}[x]$ and $k \in \mathbb{N}$ satisfying

$$\{r_{i-1}(x), r_i(x)\} = \{f(x) + x^k p_1(x), f(x) + x^k p_2(x)\}. \tag{3}$$

We write the quotient rig $\mathbb{N}[x]/\sim$ as $\mathbb{N}[x]/(p_1 = p_2)$.

The situation for *rings* is much easier: congruences, defined analogously, correspond to ideals, and if we generate a congruence $\sim$ on the ring $\mathbb{Z}[x]$ by identifying polynomials p_1 and p_2 then $q_1 \sim q_2$ if and only if $(p_1 - p_2)$ divides $(q_1 - q_2)$ in $\mathbb{Z}[x]$. So the following definition is precisely analogous to Definition 1.

Definition 5. Let $p_1, p_2, q_1, q_2 \in \mathbb{N}[x]$. We say that

$$p_1(x) = p_2(x) \Rightarrow q_1(x) = q_2(x) \text{ rig-theoretically}$$

if the following equivalent conditions hold:

(a) for all rigs A and all $a \in A$, if $p_1(a) = p_2(a)$ then $q_1(a) = q_2(a)$;
(b) as (a), but restricted to commutative rigs;
(c) q_1 and q_2 represent the same element of the quotient rig $\mathbb{N}[x]/(p_1 = p_2)$;
(d) there is a finite sequence $r_0, \ldots, r_n$ of elements of $\mathbb{N}[x]$ (for some $n \in \mathbb{N}$) satisfying $r_0 = q_1$, $r_n = q_2$, and the condition described in (3).

We want to discuss categories in which objects can be added and multiplied. Such categories bear the same conceptual relation to rigs as monoidal categories do to monoids. So, a **rig category** is a category $\mathscr{A}$ equipped with a symmetric monoidal structure $(\oplus, 0)$ and a monoidal structure $(\otimes, I)$ with the latter distributing over the former up to coherent isomorphism—in other words, there are specified isomorphisms

$$0 \xrightarrow{\sim} T \otimes 0 \qquad 0 \xrightarrow{\sim} 0 \otimes T$$

$$(T \otimes U) \oplus (T \otimes V) \xrightarrow{\sim} T \otimes (U \oplus V) \qquad (U \otimes T) \oplus (V \otimes T) \xrightarrow{\sim} (U \oplus V) \otimes T$$

for each $T, U, V \in \mathscr{A}$. The distributivity, associativity and unit isomorphisms are required to satisfy various axioms; see [6] for details. Any polynomial $p \in \mathbb{N}[x]$ and object T of a rig category $\mathscr{A}$ give rise to a new object $p(T)$ of $\mathscr{A}$, which the axioms ensure is well-defined up to canonical isomorphism.

Example 6. (a) A **distributive category** is a category in which finite coproducts and products exist and the latter distribute over the former. Any such is naturally a rig category. Examples are the category of sets, the category of topological spaces, any bicartesian closed category, and any distributive lattice.

(b) The category of sets and partial functions, with disjoint union as $\oplus$ and cartesian product as $\otimes$, is a rig category.

(c) The category of modules over a fixed commutative ring, with the usual $\oplus$ and $\otimes$, is a rig category. The same goes for the category of representations of a group and the category of vector bundles over a topological space.

(d) A discrete rig category (one in which the only morphisms are the identities) is merely a rig.

The set (or class) of isomorphism classes of objects of a rig category forms a (possibly large) rig, called its **Burnside rig**. For instance, the Burnside rig of the distributive category of sets is the rig of cardinals (Example 3(e)). The Burnside rigs of the categories in Example 6(c) are basic to K-theory and representation theory. (In those subjects the Burnside rig is no sooner formed than turned into a ring, and, as pointed out in Example 3(d), this process potentially destroys information: the

'Eilenberg swindle' of K-theory. For this reason, among others, the categories of Example 6(c) are actually replaced by certain subcategories.)

By considering Burnside rigs and discrete rig categories we see that a further equivalent condition may be added to Definition 5:

(e) for all rig categories $\mathscr{A}$ and all $T \in \mathscr{A}$, if $p_1(T) \cong p_2(T)$ then $q_1(T) \cong q_2(T)$.

Moreover, suppose that the rig-theoretic implication holds and that we are given a specific isomorphism $p_1(T) \xrightarrow{\sim} p_2(T)$, for some T and $\mathscr{A}$. Then there exists a chain $r_0, \ldots, r_n$ of polynomials as in condition 5(d), and we can build from it a specific isomorphism $q_1(T) \xrightarrow{\sim} q_2(T)$. This is exactly how the 18-step isomorphism $T \xrightarrow{\sim} T^7$ in the Introduction was built.

If a proof can be done without using subtraction then it can certainly be done with subtraction available; in other words, if an implication holds rig-theoretically then it certainly holds ring-theoretically. This paper is about going the other way, and the next result shows that it is a question of cancellability.

Proposition 7. *Let $p_1, p_2, q_1, q_2 \in \mathbb{N}[x]$ and suppose that*

$$p_1(x) = p_2(x) \Rightarrow q_1(x) = q_2(x) \text{ ring-theoretically.}$$

Then there exists $s \in \mathbb{N}[x]$ such that

$$p_1(x) = p_2(x) \Rightarrow q_1(x) + s(x) = q_2(x) + s(x) \text{ rig-theoretically.}$$

Proof. We are given that there exists $r \in \mathbb{Z}[x]$ satisfying

$$q_1 - q_2 = r \cdot (p_1 - p_2)$$

in $\mathbb{Z}[x]$. We may write $r = r_1 - r_2$ for some $r_1, r_2 \in \mathbb{N}[x]$, and then

$$q_1 + r_1 p_2 + r_2 p_1 = q_2 + r_1 p_1 + r_2 p_2$$

in $\mathbb{N}[x]$. Put $s = r_1 p_1 + r_2 p_2$: then $q_1 + s$ and $q_2 + s$ represent the same element of the quotient rig $\mathbb{N}[x]/(p_1 = p_2)$, as required. □

4. Cancellation in commutative semigroups

A **commutative semigroup** $(A, *)$ is a set A equipped with a commutative associative binary operation $*$. In general, $a_1 * b = a_2 * b$ does not imply $a_1 = a_2$, but in this section we give a condition under which it does.

Later we will apply this to the underlying additive semigroup of a rig, but it seems to be easier to understand the following results if the semigroup operation $*$ is thought of as multiplication. Informally, take a commutative semigroup and call an element **'high'** if every element divides it. Then the set of high elements is closed under multiplication with an arbitrary element, and in it every element divides every other element. This says that the set of high elements is, if not empty, a group. So given an equation $a_1 * b = a_2 * b$ in which each a_i is high, we may post-multiply each side by a high element a then divide through by the high element $b * a$ to conclude that $a_1 = a_2$.

Formally, given a commutative semigroup $(A, *)$, define a relation $\leqslant_A$ on A by

$$b \leqslant_A a \iff \text{there exists } c \in A \text{ satisfying } b * c = a.$$

The notation is potentially misleading: $\leqslant_A$ is transitive but not necessarily reflexive (consider strictly positive numbers under addition) or antisymmetric (consider an abelian group). However, $\leqslant_A$ has the expected meaning when $(A, *) = (\mathbb{N}, +)$ or when $*$ is the least upper bound operation on a (semi-)lattice (cf. Example 3(d)).

An element a of A is called high if $b \leqslant_A a$ for all $b \in A$, and the set of high elements of A is written $H(A)$. This set may be empty (as for $(\mathbb{N}, +)$), or all of A (as for abelian groups), or somewhere in between (interesting examples of which occur later). We call A a **clique** if $b \leqslant_A a$ for all $a, b \in A$, or, equivalently, if $H(A) = A$.

Lemma 8. *Let $(A, *)$ be a commutative semigroup. Then $H(A)$ is a sub-semigroup of A, and $(H(A), *)$ is a clique.*

Proof. Let $a, b \in H(A)$. We have to show that $a * b \in H(A)$ and that there exists $c \in H(A)$ satisfying $a = b * c$. In fact, we can do both without knowing that b is high. First, we have $a * b \geqslant_A a \in H(A)$, hence $a * b \in H(A)$. Second, we have $a \geqslant_A b * a$, so there exists $d \in A$ satisfying $a = b * a * d$; take $c = a * d \in H(A)$. □

We observed earlier that any abelian group is a clique. The next result is very nearly the converse.

Lemma 9. *A commutative semigroup is a clique if and only if it is empty or an abelian group.*

Proof. Let $(B, *)$ be a non-empty clique, and pick an element d. Since $d \leqslant_B d$, there exists $z \in B$ satisfying $d * z = d$. For any $b \in B$ we have $b \geqslant_B d$, so there exists $c \in B$ satisfying $b = d * c$, and then using commutativity, $b * z = b$. Hence z is a unit for $*$. Also, B has inverses: if $b \in B$ then $b \leqslant_B z$. □

Corollary 10. *If A is a commutative semigroup then $H(A)$ is either empty or, when equipped with the inherited binary operation, an abelian group.*

This result can also be obtained as a corollary of a fundamental result in semigroup theory known as Green's theorem [5]; see [2] for details.

Beware that even if the semigroup A has a unit, this is very likely *not* the unit of $H(A)$. (Indeed, the units can only be the same if A is an abelian group to start with.) For a simple example, take A to be a lattice and $*$ to be least upper bound: then the unit of A is the least element, but $H(A)$ is the singleton consisting of the greatest element.

Corollary 11. *Let a_1 and a_2 be high elements of a commutative semigroup $(A, *)$. If there exists $b \in A$ such that $a_1 * b = a_2 * b$ then $a_1 = a_2$.*

As explained, we will apply these results when A is the underlying additive semigroup of a rig. In that case, although we will not need to know it, we have:

Corollary 12. *If A is a rig then $H(A, +)$ is either empty or, when equipped with the inherited binary operations, a ring.*

Proof. In the notation of Lemma 9, the multiplicative unit is $1 + z$. The only non-trivial check is that $a \cdot z = z$ for all high a. □

5. High polynomials

Our final task is to find the high elements of the quotient rig $\mathbb{N}[x]/(p_1 = p_2)$ (or rather, its underlying additive semigroup). We do this under several assumptions. These are: p_1 is the identity polynomial, and p_2 has non-zero constant term and degree at least two. As briefly mentioned at the end of the Introduction, these assumptions have natural interpretations in applications to computer science.

Fix a polynomial $p(x) \in \mathbb{N}[x]$. To lighten the notation, we say that a polynomial f is high to mean that its image $[f]$ in the quotient $\mathbb{N}[x]/(x = p(x))$ is high, and write $g \leqslant f$ to mean $[g] \leqslant_{\mathbb{N}[x]/(x=p(x))} [f]$. Observe that $\leqslant$ is compatible with multiplication: if $g_1 \leqslant f_1$ and $g_2 \leqslant f_2$ then $g_1 g_2 \leqslant f_1 f_2$.

Lemma 13. *If p has non-zero constant term then $1 \leqslant x \leqslant x^2 \leqslant \cdots$.*

Proof. We have $p(x) \geqslant 1$ by hypothesis, so $x \geqslant 1$, and multiplying through by x^n gives $x^{n+1} \geqslant x^n$ for all $n \in \mathbb{N}$. □

Lemma 14. *If p has non-zero constant term and degree at least two then*

(a) *$x \geqslant nx$ for all $n \in \mathbb{N}$, and*
(b) *$x \geqslant x^n$ for all $n \in \mathbb{N}$.*

Proof. The hypotheses imply that $x \geqslant 1 + x^d$ for some $d \geqslant 2$. By induction, it is enough to prove each of (a) and (b) when $n = 2$; and using Lemma 13,

$$x \geqslant x^d \geqslant x^{d-1}(1 + x^d) = x^{d-1} + x^{2d-1} \geqslant 2x$$

and

$$x \geqslant x^d \geqslant x^2. \qquad \square$$

Proposition 15. *If p has non-zero constant term and degree at least two then every non-constant polynomial is high.*

Proof. By Lemma 14,

$$x \geqslant kx \geqslant x^{n_1} + x^{n_2} + \cdots + x^{n_k}$$

for all $k, n_1, \ldots, n_k \in \mathbb{N}$; in other words, x is high. The result then follows from Lemma 13. $\square$

We will not need to know it, but under the hypotheses of the proposition, the high polynomials are precisely the non-constants. This can be proved by considering the 3-element quotient rig of $\mathbb{N}[x]$ in which the equivalence classes are $\{0\}$, the set of non-zero constants, and the set of non-constants. So by Corollary 12, the non-constant polynomials form a ring. The quotient rig $\mathbb{N}[x]/(x = p(x))$ is the disjoint union of this ring with the set of natural numbers.

We arrived at the results of this section and the last after a conversation between one of us (M.F.) and Bill Lawvere, in which Lawvere mentioned a result of Steve Schanuel's that 'the infinite-dimensional elements of the rig form a ring'. Recently we learned that the result that the non-constant polynomials form a ring was also known to Schanuel. But we do not know any details of this work beyond what is in [8].

6. The main theorem

We have now done all the work and can read off the main theorem, of which we give two slightly different versions.

Theorem 16. *Let $p, q_1, q_2 \in \mathbb{N}[x]$ be polynomials such that p has non-zero constant term and degree at least two, and q_1 and q_2 have degree at least one. If*

$$x = p(x) \Rightarrow q_1(x) = q_2(x) \text{ ring-theoretically}$$

then the same is true rig-theoretically.

Proof. Assemble Proposition 7, Corollary 11, and Proposition 15. □

Theorem 17. *Let p, q_1, q_2 be polynomials as in the first sentence of Theorem* 16. *Suppose that the polynomial $p(x) - x \in \mathbb{Z}[x]$ is primitive and has no repeated complex roots, and that each complex root t satisfies $q_1(t) = q_2(t)$. Then*

$$x = p(x) \Rightarrow q_1(x) = q_2(x) \text{ rig-theoretically.}$$

Proof. Assemble Proposition 2 and Theorem 16. □

The first version is more general, and is the one to use when the complex solutions of $x = p(x)$ are hard to find. To apply it we verify by the division algorithm that the ring-theoretic implication holds. The second version is more useful when the solutions of $x = p(x)$ are known or easily calculated, and is applied by simply checking that $q_1(t) = q_2(t)$ for each solution t.

The proofs are constructive. So, if we are given polynomials p, q_1, q_2 satisfying the hypotheses of either Theorem 16 or 17, then unwinding the proof gives an explicit sequence of polynomials demonstrating the rig-theoretic implication (as in Definition 5(d)). Combining this with the observations of Section 3, if we are also given an isomorphism $T \xrightarrow{\sim} p(T)$ for some object T of a rig category then we obtain an explicit isomorphism $q_1(T) \xrightarrow{\sim} q_2(T)$.

Example 18. We write $+$ and $\times$ for the monoidal structures of a rig category. (a) Returning to our original example, $1 - x + x^2$ has distinct complex roots $e^{\pm i\pi/3}$, both of which satisfy $x^6 = 1$, so Theorem 17 tells us that

$$T \cong 1 + T^2 \Rightarrow T^7 \cong T$$

for any object T of a rig category. In particular, $T^7 \cong T$ when T is the set of binary trees.

(b) If $n \geqslant 2$ then the solutions of the equation $x = 1 + x + x^n$ are the complex nth roots of -1, each of which satisfies $x^{2n} = 1$. Hence

$$T \cong 1 + T + T^n \Rightarrow T^{2n+1} \cong T \quad \text{and} \quad T + T^{2n} = 1 + T$$

for any object T of a rig category.

(c) The case $n = 2$ of the previous example is particularly easy to work with since the solutions are $\pm i$. For instance, if T is an object of a rig category satisfying $T \cong 1 + T + T^2$ then there are also isomorphisms

$$T^4 \cong 2 + T^2, \quad T + T^3 \cong 1 + T^2, \quad (1 + T)^9 \cong 16(1 + T)$$

(observing for the last one that $1 \pm i = \sqrt{2}e^{\pm i\pi/4}$). Our paper [3] explores this example in depth.

(d) Let m and n be coprime positive integers, one of which is even. Then the complex roots of $(1+x^m)(1+x^n)$ are distinct and each satisfy $x^{2mn}=1$, so

$$T \cong 1 + T + T^m + T^n + T^{m+n}$$

implies

$$T^{2mn+1} \cong T \quad \text{and} \quad T + T^{2mn} \cong 1 + T$$

for any object T of a rig category.

(e) The following randomly-chosen example illustrates the power and generality of Theorem 16. Let

$$p(x) = 3 + 2x^3 + 4x^5,$$

$$q_1(x) = 6x + 10x^2 + x^3 + 3x^4 + 2x^5 + 7x^6 + 12x^7,$$

$$q_2(x) = 3 + 2x + 2x^2 + 9x^3 + 5x^6 + 4x^8.$$

A routine application of the division algorithm shows that $p(x) - x$ divides $q_1(x) - q_2(x)$ in $\mathbb{Z}[x]$, so by Theorem 16,

$$x = p(x) \Rightarrow q_1(x) = q_2(x)$$

rig-theoretically. Certainly this implication would be tiresome to prove by hand. Indeed, without the results of this paper it would not be at all clear that there was any systematic way of finding such a proof.

Observe finally that Theorems 16 and 17 are sharp: none of the hypotheses can be dropped. For the condition that p has non-zero constant term, consider the implication

$$x = x + x^2 \Rightarrow x^2 = x^3.$$

This holds ring-theoretically, but fails when x is the element ε^1 of the rig of codegrees (Example 4(b)). For the condition that p has degree at least two, consider

$$x = 1 + x \Rightarrow x = x^2,$$

which holds ring-theoretically but fails when x is the element L^1 of the rig of degrees (Example 4(a)). For the condition that q_1 and q_2 are non-constant, consider the original example of

$$x = 1 + x^2 \Rightarrow x^6 = 1,$$

which holds ring-theoretically but fails when x is the element $\aleph_0$ of the rig of countable cardinals. And we saw in Section 2 that the extra hypotheses in

Theorem 17 (primitivity and distinctness of roots) cannot be dropped, otherwise the implication might not even hold ring-theoretically.

Acknowledgments

We are grateful to Bill Lawvere and Steve Schanuel for a remark that gave us the last piece of the jigsaw: see Section 5. We also thank Susan Howson, Peter Johnstone, Maxim Kontsevich, and Mark Lawson for useful remarks. The second author gratefully acknowledges the support of the Institut des Hautes Études Scientifiques, where he held a William Hodge Fellowship.

References

[1] A. Blass, Seven trees in one, J. Pure Appl. Algebra 103 (1995) 1–21.
[2] M. Fiore, Isomorphisms of generic recursive polynomial types, Proceedings of the 31st Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, ACM Press, 2004, pp. 77–88.
[3] M. Fiore, T. Leinster, An objective representation of the Gaussian integers, J. Symbolic Comput., in press.
[4] R. Gates, On the generic solution to $P(X) \cong X$ in distributive categories, J. Pure Appl. Algebra 125 (1998) 191–212.
[5] J.A. Green, On the structure of semigroups, Ann. Math. 54 (1951) 163–172.
[6] M. Laplaza, Coherence for distributivity, in: Coherence in Categories, Lecture Notes in Mathematics, Vol. 281, Springer, Berlin, 1972, pp. 29–65.
[7] F.W. Lawvere, Some thoughts on the future of category theory, in: Proceedings of Como 1990, Lecture Notes in Mathematics, Vol. 1488, Springer, Berlin, 1991, pp. 1–13.
[8] S.H. Schanuel, Negative sets have Euler characteristic and dimension, in: Proceedings of Como 1990, Lecture Notes in Mathematics, Vol. 1488, Springer, Berlin, 1991, pp. 379–385.