

A complete equational axiomatisation of partial differentiation

Gordon D. Plotkin ¹

*Google Research
Mountain View, United States*

Abstract

We formalise the well-known rules of partial differentiation in a version of equational logic with function variables and binding constructs. We prove the resulting theory is complete with respect to polynomial interpretations. The proof makes use of Severi's interpolation theorem that all multivariate Hermite problems are solvable. We also present a number of related results, such as decidability and equational completeness.

Keywords: Partial differentiation, second-order equational logic, completeness, Hermite interpolation

1 Introduction

There has been recent increasing interest in categorical axiomatisations of differential structure. For example, for forward differentiation, see [2]; for reverse differentiation see [4]; and for tangent structures, see [3]. A natural question is whether the axioms are well chosen. The authors generally show they hold in natural structures. For example, in the case of cartesian differential categories they hold for the category of finite powers of the reals and smooth functions. So the axioms are, in that sense, correct. But one can additionally ask, if, or in what sense, they are complete, that is whether there are missing axioms.

Here we interest ourselves in an allied basic logical question which we hope will help with the categorical one: are the standard rules for manipulating partial derivatives complete? The rules are indeed well known: derivatives of products and sums are given by simple formulas involving their immediate subexpressions and their partial derivatives; derivatives of real constants and variables are 0 or 1; the chain rule takes care of function applications; and partial derivatives with respect to different variables commute. They are surely complete.

To prove this, we proceed by first formalising the standard rules of partial differentiation using a suitable kind of equational logic. Our axiom system consists of the ring axioms, addition and multiplication tables for the reals, and four axioms for partial differentiation. As partial differentiation involves variable binding, we are outside the scope of standard equational logic. So we instead use an expanded version which allows binding constructs and makes use of function variables. An equational logic of this kind, called *second-order equational logic*, has been presented by Fiore, Hur, and Mahmoud [6,7]. We employ a minor, although entirely equivalent, variant of their logic.

The resulting theory of partial differentiation is naturally interpreted using smooth functions, but, for completeness, it turns out that it suffices to use only polynomial functions. Indeed, the theory is complete even if the interpretation of function variables is restricted to natural number polynomial functions, and the interpretation of variables to natural numbers. This remains the case if the constants are restricted to the rationals or to 0 and 1 (equivalently the integers).

To establish completeness we employ an interpolation theorem, a well-known theorem of Severi [12] on the solvability of Hermite interpolation problems. Hermite interpolation generalises Lagrange interpolation. In

¹ Email: plotkin@google.com

Lagrange interpolation, one seeks a real polynomial taking prescribed values at prescribed points, and this has an evident generalization to multivariate Lagrange interpolation. In multivariate Hermite interpolation, one additionally prescribes values of (possibly higher-order) partial derivatives.

Theorem 5.4 establishes completeness relative to natural number polynomial interpretations of function variables and natural number interpretations of variables. (The polynomials are adaptations of real polynomials solving suitable Hermite interpolation problems.) We also give a number of other results. Theorem 5.5 characterises the theorems of our equational theory in terms of an equivalence relation between canonical forms. Theorem 5.8 shows that the equivalence relation holds if, and only if, it can be established using just one axiom for partial differentiation, that the order of partial differentiation by two variables does not matter (with a slightly more refined notion of canonical form this axiom can be eliminated, when no axioms for partial differentiation are needed). Theorem 5.9 shows that the theory is not only complete with respect to standard interpretations over the reals, it is also equationally complete (also known as Hilbert-Post complete), that is, it has no equationally consistent proper extensions, see [14]. This remains the case if the constants are restricted to the rationals but not if they are restricted to the integers. Theorem 5.10 shows that the subtheory in which the constants are restricted to the rationals is decidable, and that, in case an equation does not hold, a counterexample to it can be found. An important remaining question is the complexity of the decision problem.

2 Axiomatisation of partial differentiation

We assume disjoint sets of variables, ranged over by x, y, z, \dots , and of function variables, ranged over by f, g, h, \dots ; each function variable has a given arity $n \geq 0$, written $f : n$. The set of variables is assumed denumerable, as is the set of function variables of each arity. Expressions have the following forms:

$$e ::= r \quad (r \in \mathbb{R}) \mid x \mid e_0 + e_1 \mid e_0 e_1 \mid f(e_0, \dots, e_{n-1}) \quad (f : n) \mid \text{PDiff}(x.e_0, e_1)$$

The expression $\text{PDiff}(x.e_0, e_1)$ is read as the partial derivative of e_0 with respect to x , evaluated at e_1 . In this expression, the variable x has binding power over e_0 , but not over e_1 . Below, we use the more familiar, and more natural, expression $\frac{\partial e_0}{\partial x} \Big|_{x=e_1}$ (while more natural, this expression obscures the role of the variable x). We write $\frac{\partial}{\partial x} e$ for $\frac{\partial e}{\partial x} \Big|_{x=x}$, and note that the last occurrence of x is free; we write $\frac{\partial^n e}{\partial x_{n-1} \dots \partial x_0}$ for $\frac{\partial}{\partial x_{n-1}} \dots \frac{\partial}{\partial x_0} e$.

Free variables and α -equivalence, are defined as usual, and, as is also usual, we identify α -equivalent expressions. We write $\text{FV}(e)$ for the set of free variables of an expression e and $\text{FnV}(e)$ for the set of its function variables. Every expression e has a size $|e|$, defined in an evident way.

The simultaneous substitution

$$e[e'_0/x_0, \dots, e'_{n-1}/x_{n-1}]$$

of expressions for variables is defined as may be expected, with the clause for $\text{PDiff}(x.e_0, e_1)$, with its bound variable, being:

$$\text{PDiff}(x.e_0, e_1)[e'_0/x_0, \dots, e'_{n-1}/x_{n-1}] = \text{PDiff}(x.e_0[e'_0/x_0, \dots, e'_{n-1}/x_{n-1}], e_1[e'_0/x_0, \dots, e'_{n-1}/x_{n-1}])$$

where $x \notin \text{FV}(e'_0) \cup \dots \cup \text{FV}(e'_{n-1})$. There is also a notion of substitution of *abstracts* $(x_0, \dots, x_{n-1}).e'$ for function variables. For $f : n$, we define

$$e[(x_0, \dots, x_{n-1}).e'/f]$$

by structural induction on e as follows:

$$r[(x_0, \dots, x_{n-1}).e'/f] = r$$

$$x[(x_0, \dots, x_{n-1}).e'/f] = x$$

$$(e_0 + e_1)[(x_0, \dots, x_{n-1}).e'/f] = e_0[(x_0, \dots, x_{n-1}).e'/f] + e_1[(x_0, \dots, x_{n-1}).e'/f]$$

$$(e_0 e_1)[(x_0, \dots, x_{n-1}).e'/f] = e_0[(x_0, \dots, x_{n-1}).e'/f] e_1[(x_0, \dots, x_{n-1}).e'/f]$$

$$g(e_0, \dots, \textcircled{e_{n-1}})[(x_0, \dots, x_{n-1}).e'/f] = \begin{cases} e'[e_0[(x_0, \dots, x_{n-1}).e'/f]/x_0, \dots, e_{n-1}[(x_0, \dots, x_{n-1}).e'/f]/x_{n-1}] & (g = f) \\ g(e_0[(x_0, \dots, x_{n-1}).e'/f], \dots, e_{n-1}[(x_0, \dots, x_{n-1}).e'/f]) & (g \neq f) \end{cases}$$

can be $\neq n-1$ if $g \neq f$

$$\text{PDiff}(x. e_0, e_1)[(x_0, \dots, x_{n-1}).e'/f] = \text{PDiff}(x. e_0[(x_0, \dots, x_{n-1}).e'/f], e_1[(x_0, \dots, x_{n-1}).e'/f])$$

$(x \notin \text{FV}(e') \setminus \{x_0, \dots, x_{n-1}\})$

There is a more general, similarly defined, simultaneous substitution of several abstracts:

$$e[(x_{11}, \dots, x_{1n_1}).e'_1/f_1, \dots, (x_{kn_k}, \dots, x_{kn_k}).e'_k/f_k]$$

where $f_i : n_i$, for $i = 1, k$.

The axioms of the theory of partial differentiation consist of those for commutative rings, the addition and multiplication tables for the real constants, and four axioms for partial differentiation. The four axioms are the usual rules for the differentiation of addition and multiplication, the chain rule for binary functions, and the commutativity of partial differentiation with respect to different variables.

$$\frac{\partial x + y}{\partial x} = 1 \qquad \frac{\partial yx}{\partial x} = y$$

$$\frac{\partial f(g_0(x), g_1(x))}{\partial x} = \frac{\partial f(x_0, g_1(x))}{\partial x_0} \Big|_{x_0=g_0(x)} \frac{\partial g_0(x)}{\partial x} + \frac{\partial f(g_0(x), x_1)}{\partial x_1} \Big|_{x_1=g_1(x)} \frac{\partial g_1(x)}{\partial x}$$

$$\frac{\partial}{\partial y} \frac{\partial}{\partial x} f(x, y) = \frac{\partial}{\partial x} \frac{\partial}{\partial y} f(x, y)$$

The theorems of the theory of partial differentiation are obtained by closing the axioms under the evident rules for equality, congruence, and both kinds of substitution, following the usual pattern for equational systems. In particular, the congruence rule for partial differentiation is as follows:

$$\frac{e_0 = e_1 \quad e'_0 = e'_1}{\text{PDiff}(x. e_0, e'_0) = \text{PDiff}(x. e_1, e'_1)}$$

If an equation $e_0 = e_1$ holds in this theory we write

$$\vdash e_0 = e_1$$

We next establish a number of expected consequences of the axioms, including a general form of the chain rule. We employ an evident abbreviation $\sum_{i=m}^n e_i$ for finite sum expressions.

Lemma 2.1 *The following equations are provable in the theory of partial differentiation:*

(i)

$$\frac{\partial x}{\partial x} = 1$$

(ii)

$$\frac{\partial f(x) + g(x)}{\partial x} = \frac{\partial f(x)}{\partial x} + \frac{\partial g(x)}{\partial x}$$

(iii)

$$\frac{\partial f(x)g(x)}{\partial x} = g(x)\frac{\partial f(x)}{\partial x} + f(x)\frac{\partial g(x)}{\partial x}$$

(iv)

$$\frac{\partial f(g_0(x), \dots, g_{n-1}(x))}{\partial x} = \sum_{i=0}^{n-1} \frac{\partial f(g_0(x), \dots, g_{i-1}(x), x_i, g_{i+1}(x), \dots, g_{n-1}(x))}{\partial x_i} \Big|_{x_i=g_i(x)} \frac{\partial g_i(x)}{\partial x}$$

(v)

$$\frac{\partial e}{\partial x} = 0 \quad (x \notin \text{FV}(e))$$

Proof.

(i) This follows from the axiom for addition, substituting 0 for y .

(ii) This follows from the axiom for addition and the binary chain rule.

(iii) This follows from the axiom for multiplication and the binary chain rule.

(iv) First, note that, substituting 0 for y in the axiom for multiplication, we obtain $\frac{\partial 0}{\partial x} = 0$. Next, for the case $n = 0$ of the chain rule, substitute $(x_0, x_1).h()$ for f and $(z).0$ for both g_0 and g_1 in the binary chain rule to get the following provable equations:

$$\begin{aligned} \frac{\partial h()}{\partial x} &= \frac{\partial h()}{\partial x_0} \Big|_{x_0=0} \frac{\partial 0}{\partial x} + \frac{\partial h()}{\partial x_1} \Big|_{x_1=0} \frac{\partial 0}{\partial x} \\ &= \frac{\partial h()}{\partial x_0} \Big|_{x_0=0} 0 + \frac{\partial h()}{\partial x_1} \Big|_{x_1=0} 0 \\ &= 0 \end{aligned}$$

Next, for the unary chain rule, substitute $(x_0, x_1).h(x_0)$ for f and $(z).0$ for g_1 in the binary chain rule to get the following provable equations:

$$\begin{aligned} \frac{\partial h(g_0(x_0))}{\partial x} &= \frac{\partial h(x_0)}{\partial x_0} \Big|_{x_0=g_0(x)} \frac{\partial g_0(x)}{\partial x} + \frac{\partial h(x_0)}{\partial x_1} \Big|_{x_1=0} \frac{\partial 0}{\partial x} \\ &= \frac{\partial h(x_0)}{\partial x_0} \Big|_{x_0=g_0(x)} \frac{\partial g_0(x)}{\partial x} \end{aligned}$$

Finally, for $n \geq 2$, we proceed by induction. The axiom provides the base case, so assuming the induction hypothesis for n , we prove it for $n + 1$. Substituting $(z).z$ for g_0 and g_1 in the binary chain rule, and using $\frac{\partial x}{\partial x} = 1$, we obtain the following equation:

$$\frac{\partial f(x, x)}{\partial x} = \frac{\partial f(x'_0, x)}{\partial x'_0} \Big|_{x'_0=x} + \frac{\partial f(x, x'_1)}{\partial x'_1} \Big|_{x'_1=x}$$

Then, substituting $(x, y).h(g_0(x), g_1(y), \dots, g_n(y))$ for f , we obtain:

$$\frac{\partial h(g_0(x), g_1(x), \dots, g_n(x))}{\partial x} = \frac{\partial h(g_0(x'_0), g_1(x), \dots, g_n(x))}{\partial x'_0} \Big|_{x'_0=x} + \frac{\partial h(g_0(x), g_1(x'_1), \dots, g_n(x'_1))}{\partial x'_1} \Big|_{x'_1=x}$$

For the first of these two summands we have:

$$\begin{aligned}
 \left. \frac{\partial h(g_0(x'_0), g_1(x), \dots, g_n(x))}{\partial x'_0} \right|_{x'_0=x} &= \frac{\partial h(g_0(x'_0), g_1(x), \dots, g_n(x))}{\partial x'_0} [x/x'_0] \\
 &= \left(\left. \frac{\partial h(x_0, g_1(x), \dots, g_n(x))}{\partial x_0} \right|_{x_0=g_0(x'_0)} \frac{\partial g_0(x'_0)}{\partial x'_0} \right) [x/x'_0] \\
 &= \left. \frac{\partial h(x_0, g_1(x), \dots, g_n(x))}{\partial x_0} \right|_{x_0=g_0(x)} \frac{\partial g_0(x)}{\partial x}
 \end{aligned}$$

with the second equality being an application of the unary chain rule and the other two equalities being syntactic identities (the second up to α -equivalence).

For the second of these, making use of the induction hypothesis (the n -ary chain rule) we have:

$$\begin{aligned}
 \left. \frac{\partial h(g_0(x), g_1(x'_1), \dots, g_n(x'_1))}{\partial x'_1} \right|_{x'_1=x} &= \frac{\partial h(g_0(x), g_1(x'_1), \dots, g_n(x'_1))}{\partial x'_1} [x/x'_1] \\
 &= \left(\sum_{i=1}^n \left. \frac{\partial h(g_0(x), g_1(x'_1), \dots, g_{i-1}(x'_1), x_i, g_{i+1}(x'_1), \dots, g_n(x'_1))}{\partial x_i} \right|_{x_i=g_i(x'_1)} \frac{\partial g_i(x'_1)}{\partial x'_1} \right) [x/x'_1] \\
 &= \sum_{i=1}^n \left. \frac{\partial h(g_0(x), g_1(x), \dots, g_{i-1}(x), x_i, g_{i+1}(x), \dots, g_n(x))}{\partial x_i} \right|_{x_i=g_i(x)} \frac{\partial g_i(x)}{\partial x}
 \end{aligned}$$

Combining the results, we reach the desired conclusion.

(v) For $n = 0$, the chain rule is $\frac{\partial f()}{\partial x} = 0$, where $f:0$. Substituting $() \cdot e$ for f we obtain $\frac{\partial e}{\partial x} = 0$ as $x \notin \text{FV}(e)$. \square

It will prove useful to know the interaction between differentiation and substitution:

Lemma 2.2 *The following are equivalent:*

- (i) *The general chain rule.*
- (ii) *The following substitution principle, that for all expressions e and e_0, \dots, e_{n-1} we have:*

$$\vdash \frac{\partial e[e_0/x_0, \dots, e_{n-1}/x_{n-1}]}{\partial x} = \sum_{i=0}^{n-1} \frac{\partial e}{\partial x_i} [e_0/x_0, \dots, e_{n-1}/x_{n-1}] \frac{\partial e_i}{\partial x}$$

where x_0, \dots, x_{n-1} are distinct variables such that $x \notin \text{FV}(e) \setminus \{x_0, \dots, x_{n-1}\}$.

Proof. In one direction assume the general chain rule. Choose $f:n$ not in any of the $\text{FnV}(e_i)$. Using the general chain rule we have:

$$\vdash \frac{\partial f(e_0, \dots, e_{n-1})}{\partial x} = \sum_{i=0}^{n-1} \left. \frac{\partial f(e_0, \dots, e_{i-1}, x_i, e_{i+1}, \dots, e_{n-1})}{\partial x_i} \right|_{x_i=e_i} \frac{\partial e_i}{\partial x}$$

Then, substituting $(x_0, \dots, x_{n-1}) \cdot e$ for f on the left we have the following identities (recall that we identify α -equivalent expressions):

$$\begin{aligned}
 \frac{\partial f(e_0, \dots, e_{n-1})}{\partial x} [(x_0, \dots, x_{n-1}) \cdot e / f] &= \frac{\partial f(e_0, \dots, e_{n-1})[(x_0, \dots, x_{n-1}) \cdot e / f]}{\partial x} \quad (\text{as } x \notin \text{FV}(e) \setminus \{x_0, \dots, x_{n-1}\}) \\
 &= \frac{\partial e[e_0/x_0, \dots, e_{n-1}/x_{n-1}]}{\partial x} \quad (\text{as } f \notin \text{FnV}(e_0, \dots, e_{n-1}))
 \end{aligned}$$

Next, substituting on the right, we have the identities:

$$\begin{aligned}
 & \left(\sum_{i=0}^{n-1} \frac{\partial f(e_0, \dots, e_{i-1}, x_i, e_{i+1}, \dots, e_{n-1})}{\partial x_i} \right) \Big|_{x_i=e_i} \frac{\partial e_i}{\partial x} \left[(x_0, \dots, x_{n-1}) \cdot e/f \right] \\
 &= \sum_{i=0}^{n-1} \left(\frac{\partial f(e_0, \dots, e_{i-1}, x_i, e_{i+1}, \dots, e_{n-1})}{\partial x_i} \Big|_{x_i=e_i} \left[(x_0, \dots, x_{n-1}) \cdot e/f \right] \right) \frac{\partial e_i}{\partial x} \\
 & \quad \text{(as } f \notin \text{FnV}(e_i), \text{ for } i = 0, \dots, n-1) \\
 &= \sum_{i=0}^{n-1} \left(\frac{\partial f(e_0, \dots, e_{i-1}, x_i, e_{i+1}, \dots, e_{n-1})}{\partial x_i} \left[(x_0, \dots, x_{n-1}) \cdot e/f \right] \Big|_{x_i=e_i} \right) \frac{\partial e_i}{\partial x} \\
 & \quad \text{(as } x_i \notin \text{FV}(e) \setminus \{x_0, \dots, x_{n-1}\}) \\
 &= \sum_{i=0}^{n-1} \frac{\partial e[e_0/x_0, \dots, e_{i-1}/x_{i-1}, x_i/x_i, e_{i+1}/x_{i+1}, \dots, e_{n-1}/x_{n-1}]}{\partial x_i} \Big|_{x_i=e_i} \frac{\partial e_i}{\partial x} \\
 & \quad \text{(as } f \notin \text{FnV}(e_0, \dots, e_{i-1}, e_{i+1}, \dots, e_{n-1})) \\
 &= \sum_{i=0}^{n-1} \frac{\partial e[e_0/x_0, \dots, e_{i-1}/x_{i-1}, e_{i+1}/x_{i+1}, \dots, e_{n-1}/x_{n-1}]}{\partial x_i} \Big|_{x_i=e_i} \frac{\partial e_i}{\partial x} \\
 &= \sum_{i=0}^{n-1} \frac{\partial e}{\partial x_i} [e_0/x_0, \dots, e_{n-1}/x_{n-1}] \frac{\partial e_i}{\partial x}
 \end{aligned}$$

and the conclusion follows.

In the other direction, assuming the substitution principle, take e to be $f(x_0, \dots, x_{n-1})$ (when $\text{FV}(e) \setminus \{x_0, \dots, x_{n-1}\} = \emptyset$) and e_i to be $g_i(x_i)$ for $i = 0, n-1$, to obtain:

$$\vdash \frac{\partial f(g_0(x), \dots, g_{n-1}(x))}{\partial x} = \sum_{i=0}^{n-1} \frac{\partial f(x_0, \dots, x_{n-1})}{\partial x_i} [g_0(x)/x_0, \dots, g_{n-1}(x)/x_{n-1}] \frac{\partial g_i(x)}{\partial x}$$

As

$$\frac{\partial f(x_0, \dots, x_{n-1})}{\partial x_i} [g_0(x)/x_0, \dots, g_{n-1}(x)/x_{n-1}]$$

and

$$\frac{\partial f(g_0(x), \dots, g_{i-1}(x), x_i, g_{i+1}(x), \dots, g_{n-1}(x))}{\partial x_i} \Big|_{x_i=g_i(x)}$$

are α -equivalent, we are done. \square

A (*real*) *polynomial* is (as usual) an expression containing no function variables or partial differentiations; it is a *natural number polynomial* if all its constants are natural numbers. We write $P(x_0, \dots, x_{n-1})$ to indicate that P is a real polynomial whose (necessarily free) variables are included in $\{x_0, \dots, x_{n-1}\}$; for expressions e_0, \dots, e_{n-1} we may then write $P(e_0, \dots, e_{n-1})$ for $P[e_0/x_0, \dots, e_{n-1}/x_{n-1}]$. We write $P \sim P'$ to mean that the polynomials P and P' are equal modulo the axioms for commutative rings and the addition and multiplication tables for the real constants.

We remark that our equational system is an instance of a second-order equational logic for binding. This logic has *operator constants* op of arities of the form $(b_0, \dots, b_{n-1}; m)$ ($b_i, m \geq 0$). These are used to form a compound expression from n abstracts, of respective arities b_0, \dots, b_{n-1} , and m expressions. With this notation, $+$ has arity $(; 2)$ and PDiff has arity $(1; 1)$. This system is single-sorted; there is a natural generalisation to a multisorted version.

In [6] Fiore and Hur gave a multisorted second-order equational logic, and in [7] Fiore and Mahmoud considered the single-sorted case. The differences between their systems and ours are inessential: what we term ‘function variables’ they term ‘metavariables’; whereas our function variables come with preassigned arities, theirs do not, and they rather utilise suitable environments; their arities are of the simpler form (b_0, \dots, b_{n-1}) but are no less general, as the extra arguments become 0-ary abstracts; and whereas we have both a substitution operation for function variables and a context rule, they employ a single equivalent rule. A single substitution version of their equivalent rule is:

$$\frac{e_0 = e_1}{e_0[(x_0, \dots, x_{n-1}) \cdot e'/f] = e_1[(x_0, \dots, x_{n-1}) \cdot e'/f]}$$

They employ a simultaneous substitution version of this rule, which is, in any case, equivalent to the single substitution rule.

Returning to our single-sorted version of second-order equational logic, we say that an equational theory is equationally inconsistent if $x = y$ is a theorem (with x, y different), and we say that an equational theory is equationally complete if any extension of it by a single non-theorem is equationally inconsistent. Note that this is a syntactic criterion, independent of any particular model.

Polynomials provide an example of equational completeness, though one without a binding operator. The theory is that of commutative rings with constants for all reals and addition and multiplication tables (i.e., our theory less partial differentiation). The idea of the proof is that distinct polynomials have distinct values for suitable choices of values for their variables. This can be formalised within the logic, and so, given an equation between distinct polynomials, two distinct reals can be proved equal, and that, in turn, enables one to prove $x = y$.

The $\lambda\beta\eta$ -calculus provides a partial example with a binding operator. It can be formalised as a second-order theory with a binary application operator $\text{ap}:(;2)$ and a unary lambda abstraction operator $\lambda:(1;0)$ and the two equations

$$\text{ap}(\lambda((x).f(x)), y) = f(y) \quad \lambda((y).\text{ap}(x, y)) = x$$

While not equationally complete, it is partially so in the sense that, by Böhm's theorem [1], no two distinct $\beta\eta$ -normal forms can be consistently equated. The theory of the Beta-Bernoulli process [13] also has binding operators; it is formulated using an equational logic that is a bit different than ours in order to accommodate algebraic effects. The theory is almost equationally complete, in the sense that it has just one consistent extension. In this extension all quantitative information about probabilities is lost, and so can be ruled out for the purposes at hand.

3 Semantics

We give a semantics of our partial differentiation expressions using smooth functions $h:\mathbb{R}^n \rightarrow \mathbb{R}$. An *environment* ρ is a function from variables to reals; a *function environment* φ is a function from function variables to smooth functions of arity n to smooth functions on \mathbb{R}^n . We write $\rho[r_0/x_0, \dots, r_{n-1}/x_{n-1}]$ for the environment with value r_i on x_i (for $i = 0, n-1$) and value $\rho(x)$ on any other variable x ; function environments $\varphi[h_0/f_0, \dots, h_{n-1}/f_{n-1}]$, where h_i is m -ary if $f_i:m$, are defined similarly. We write 0 for both the constantly 0 environment and the function environment yielding constantly 0 functions.

We define the denotation $\mathcal{S}[e]\varphi\rho$ of an expression relative to a function environment and an environment by the following clauses:

$$\begin{aligned} \mathcal{S}[r]\varphi\rho &= r \\ \mathcal{S}[x]\varphi\rho &= \rho(x) \\ \mathcal{S}[e + e']\varphi\rho &= \mathcal{S}[e]\varphi\rho + \mathcal{S}[e']\varphi\rho \\ \mathcal{S}[ee']\varphi\rho &= \mathcal{S}[e]\varphi\rho \times \mathcal{S}[e']\varphi\rho \\ \mathcal{S}[f(e_0, \dots, e_{n-1})]\varphi\rho &= \varphi(f)(\mathcal{S}[e_0]\varphi\rho, \dots, \mathcal{S}[e_{n-1}]\varphi\rho) \\ \mathcal{S}[\text{PDiff}(x.e, e')]\varphi\rho &= D(r \in \mathbb{R} \mapsto \mathcal{S}[e]\varphi\rho[r/x])(\mathcal{S}[e']\varphi\rho) \end{aligned}$$

Prima facie this definition may not be proper as the last clause depends on the function $r \in \mathbb{R} \mapsto \mathcal{S}[e]\varphi\rho[r/x]$ being differentiable. However one proves by structural induction that the definition is proper and, in addition, that for any variables x_0, \dots, x_{n-1} the function sending $r_0, \dots, r_{n-1} \in \mathbb{R}$ to $\mathcal{S}[e]\varphi\rho[r_0/x_0, \dots, r_{n-1}/x_{n-1}]$ is smooth.

The denotation $\mathcal{S}[e]\varphi\rho$ of an expression depends only on the values ascribed to its free function variables by φ , and its free variables by ρ . When e has no function variables, we just write $\mathcal{S}[e]\rho$ for its denotation, omitting φ , and when it is closed (has no free variables of either kind) we just write $\mathcal{S}[e]$. As usual, we write $\models e_0 = e_1$ to mean that e_0 and e_1 have the same denotation, given any function environment and any environment.

We omit the proof of the following standard substitution lemma:

Lemma 3.1

- (i) Variable substitution and denotation commute, that is, we have:

$$\mathcal{S}[e[e_0/x_0, \dots, e_{n-1}/x_{n-1}]]\varphi\rho = \mathcal{S}[e]\varphi\rho[\mathcal{S}[e_0]\varphi\rho/x_0, \dots, \mathcal{S}[e_{n-1}]\varphi\rho/x_{n-1}]$$

- (ii) Function variable substitution and denotation commute, that is, we have:

$$\mathcal{S}[e[(x_0, \dots, x_{n-1}).e'/f]]\varphi\rho = \mathcal{S}[e]\varphi[u_0, \dots, u_{n-1} \mapsto \mathcal{S}[e']\varphi\rho[u_0/x_0, \dots, u_{n-1}/x_{n-1}]/f]\rho$$

Using this lemma, it is then straightforward to prove consistency:

Theorem 3.2 (Consistency) *For any expressions e_0 and e_1 we have:*

$$\vdash e_0 = e_1 \implies \models e_0 = e_1$$

Any polynomial $P(x_0, \dots, x_{n-1})$ defines an n -ary polynomial function on the reals, also written $P(x_0, \dots, x_{n-1})$. Note that

$$P(r_0, \dots, r_{n-1}) = \mathcal{S}[P]0[r_0/x_0, \dots, r_{n-1}/x_{n-1}]$$

and we recall that $P(x_0, \dots, x_{n-1}) \sim P'(x_0, \dots, x_{n-1})$ if, and only if, the two functions $P(x_0, \dots, x_{n-1})$ and $P'(x_0, \dots, x_{n-1})$ are equal.

We say that a function environment is a *(natural number) polynomial* environment if all its values are (natural number) polynomial functions and that an environment is a natural number environment if all its values are natural numbers. We write

$$\models_{\mathbb{N}} e_0 = e_1$$

to mean that e_0 and e_1 have the same denotation, given any natural number polynomial function environment, and any natural number environment.

Our semantics of the theory of partial differentiation employs a concrete notion of function, viz. smooth functions over the reals. For the general theory of second-order logic, a more abstract, and thereby more general, notion of function is needed. This can be formulated variously in terms of abstract clones [5], Lawvere theories, or, more conceptually, in terms of monoids in a presheaf category of contexts, see [8, 6, 7].

4 Canonical forms

To show completeness we need suitable canonical forms c ; up to a suitable equivalence relation $c \approx c'$, they provide normal forms for our theory.

Let x_0, x_1, \dots be a countably infinite sequence of distinct variables. For any function variable $f : n$, argument sequence $m = i_1, \dots, i_k \in [n]^*$ (where $[n] = \{0, \dots, n-1\}$), and expressions e_0, \dots, e_{n-1} we set:

$$f_m(e_0, \dots, e_{n-1}) = \frac{\partial^k f(x_0, \dots, x_{n-1})}{\partial x_{i_1} \dots \partial x_{i_k}}[e_0/x_0, \dots, e_{n-1}/x_{n-1}]$$

We remark that if two such expressions $f_m(e_0, \dots, e_{n-1})$ and $f_{m'}(e'_0, \dots, e'_{n'-1})$ are equal then so are: f and f' ; m and m' ; n and n' ; and e_i and e'_i , for $i = 1, n$. This is used implicitly below to ensure the uniqueness of expression case analyses. We also note that, as may be expected, for any φ and ρ we have:

$$\mathcal{S}[f_m(e_0, \dots, e_{n-1})]\varphi\rho = D_m(\varphi(f))(\mathcal{S}[e_0]\varphi\rho, \dots, \mathcal{S}[e_{n-1}]\varphi\rho)$$

The chain rule takes the following form for such applications of partial derivatives of functions:

Lemma 4.1 *We have:*

$$\vdash \frac{\partial f_m(e_0, \dots, e_{n-1})}{\partial x} = \sum_{i=0}^{n-1} f_{im}(e_0, \dots, e_{n-1}) \frac{\partial e_i}{\partial x}$$

Proof. Supposing $m = i_1, \dots, i_k \in [n]^*$, then we have

$$\begin{aligned} \frac{\partial f_m(e_0, \dots, e_{n-1})}{\partial x} &= \frac{\partial}{\partial x} \left(\frac{\partial^k f(x_0, \dots, x_{n-1})}{\partial x_{i_1} \dots \partial x_{i_k}}[e_0/x_0, \dots, e_{n-1}/x_{n-1}] \right) \\ &= \sum_{i=0}^{n-1} \left(\frac{\partial}{\partial x} \frac{\partial^k f(x_0, \dots, x_{n-1})}{\partial x_{i_1} \dots \partial x_{i_k}} \right)[e_0/x_0, \dots, e_{n-1}/x_{n-1}] \frac{\partial e_i}{\partial x} \quad (\text{by Lemma 2.2}) \\ &= \sum_{i=0}^{n-1} f_{im}(e_0, \dots, e_{n-1}) \frac{\partial e_i}{\partial x} \end{aligned}$$

where the first two identities are syntactic and the second is a provable equality. \square

We define a set of *canonical forms* c and a set of *atomic expressions* a (both sets of expressions) by simultaneous induction:

- (i) (a) Any atomic expression is a canonical expression.
- (b) Any $r \in \mathbb{R}$ is a canonical expression, as are $c + c'$ and cc' if c and c' are.
- (ii) (a) Any variable x is an atomic expression.
- (b) For any $f : n$, $m \in [n]^*$, and canonical forms c_0, \dots, c_{n-1} , $f_m(c_0, \dots, c_{n-1})$ is an atomic expression.

The set of *immediate* atomic subexpressions of a canonical form c is defined by structural recursion:

$$\text{Im}(a) = \{a\} \quad \text{Im}(r) = \emptyset \quad \text{Im}(c + c') = \text{Im}(cc') = \text{Im}(c) \cup \text{Im}(c')$$

Lemma 4.2 *For any canonical forms c and \bar{c} there is a canonical form c' such that:*

$$\vdash \left. \frac{\partial c}{\partial x} \right|_{x=\bar{c}} = c'$$

and with $\text{FnV}(c') \subseteq \text{FnV}(c) \cup \text{FnV}(\bar{c})$ and $\text{FV}(c') \subseteq \text{FV}(c) \cup \text{FV}(\bar{c})$,

Proof. The proof is by induction on the size of c . The cases where c is a constant, variable, sum, or product are handled using Lemma 2.1. This leaves the case of an atomic expression $f_m(c_0, \dots, c_{n-1})$, for which we use the fact that

$$\vdash \left. \frac{\partial f_m(c_0, \dots, c_{n-1})}{\partial x} \right|_{x=\bar{c}} = \sum_{i=0}^{n-1} f_{im}(c_0, \dots, c_{n-1}) \left. \frac{\partial c_i}{\partial x} \right|_{x=\bar{c}}$$

holds, by Lemma 4.1. □

Lemma 4.3 (Canonicalisation) *For any expression e there is a canonical form $\text{CF}(e)$ such that $\vdash e = \text{CF}(e)$ and with $\text{FnV}(\text{CF}(e)) \subseteq \text{FnV}(e)$ and $\text{FV}(\text{CF}(e)) \subseteq \text{FV}(e)$.*

Proof. We prove this by structural induction on e . If e is a variable or a constant it is already a canonical form. The cases where e is a sum or a product follow from the fact that the canonical forms are closed under sums and products.

The case where e has the form $f(e_0, \dots, e_{n-1})$ is immediate from the induction hypothesis.

Finally, the case where e has the form $\left. \frac{\partial e_0}{\partial x} \right|_{x=e_1}$ is handled using Lemma 4.2. □

We have the following corollary of the canonicalisation lemma:

Corollary 4.4 *For any closed expression e , we have:*

$$\vdash e = \mathcal{S}[e]$$

Proof. As e is closed, the same, by Lemma 4.3, is true of $\text{CF}(e)$, which latter must therefore have no immediate atomic subexpressions or free variables, and so is a closed polynomial. There is therefore a real r such that $\vdash \text{CF}(e) = r$ and so, by the lemma, $\vdash e = r$. By consistency we have $r = \mathcal{S}[e]$. □

We next define the equivalence relations that will allow us to characterise provable equality between canonical forms. We also define polynomials that will be used to formulate our Hermite interpolation problems.

For any $m, m' \in [n]^*$ we write $m \sim m'$ to mean that m is a permutation of m' . For any atomic expressions a, a' we write $a \approx a'$ to mean that

- a and a' are identical variables, or else
- for some function variable $f : n$, and for some $m \sim m'$, they have the forms $f_m(c_0, \dots, c_{n-1})$ and $f_{m'}(c'_0, \dots, c'_{n-1})$ with $\vdash c_i = c'_i$ for $i = 0, n-1$.

We evidently have:

$$a \approx a' \implies \vdash a = a' \tag{1}$$

We now fix an assignment v_a of variables (termed *separation* variables) to atomic expressions such that:

$$v_a = v_{a'} \iff a \approx a'$$

For any canonical form c , we define its *node* polynomial P_c by structural induction:

$$P_r = r \ (r \in \mathbb{R}) \quad P_a = v_a \quad P_{c+c'} = P_c + P_{c'} \quad P_{cc'} = P_c P_{c'}$$

The separation variables occurring in P_c are the v_a with a an immediate atomic subexpression of c .

We then define our equivalence relation between canonical forms by:

$$c \approx c' \iff P_c \sim P_{c'}$$

Note that this relation only depends on the choice of the v_a for the immediate atomic expressions of c and c' . Note too that the relation extends that between atomic expressions, i.e., $a \approx a'$ holds with a, a' taken as atomic expressions iff it does when they are taken as canonical forms.

Lemma 4.5

- (i) For any canonical expression c with node polynomial $P_c(v_{a_0}, \dots, v_{a_{m-1}})$, for atomic expressions a_0, \dots, a_{m-1} , we have:

$$\vdash c = P_c(a_0, \dots, a_{m-1})$$

- (ii) For any canonical expressions c and c' we have:

$$c \approx c' \implies \vdash c = c'$$

Proof. The first part is established by a straightforward structural induction. In the case where c is an atomic expression a , we have $v_a = P_c = v_{a_0}$ and so $a \approx a_0$ and so $\vdash a = a_0$, i.e., $\vdash c = P_c(a_0)$.

For the second part, suppose that we have canonical expressions c and c' such that $c \approx c'$ (and so $P_c \sim P_{c'}$). Let $v_{a_0}, \dots, v_{a_{m-1}}$ ($j = 0, m$) include all the variables of P_c and $P_{c'}$. As $P_c \sim P_{c'}$, we have:

$$\vdash P_c[a_0/v_{a_0}, \dots, a_{m-1}/v_{a_{m-1}}] = P_{c'}[a_0/v_{a_0}, \dots, a_{m-1}/v_{a_{m-1}}]$$

Applying the first part to c and c' , it follows that $\vdash c = c'$, as required. \square

5 Completeness

To establish completeness we use environments distinguishing non-equivalent canonical forms, i.e., canonical forms c and c' such that $c \not\approx c'$. We first need to be able to distinguish finite sets of non-equivalent polynomials (these will be the node polynomials of sub-canonical forms of c or c'):

Lemma 5.1 Let $P_i(x_0, \dots, x_{m-1})$ be n mutually inequivalent polynomials. Then, for some natural number choices of x_0, \dots, x_{m-1} , they take on different values.

Proof. None of the polynomials $Q_{ij} = P_i - P_j$ are identically 0 ($i \neq j$), and so neither is $Q = \prod_{ij} Q_{ij}$. So Q is non-zero for some natural number choices of x_0, \dots, x_{m-1} . That choice differentiates distinct P_i . \square

We next need to be able to solve multivariate Hermite interpolation problems in order to give prescribed values to atomic expressions. A multivariate Hermite interpolation problem of dimension $d \geq 0$ is given by:

- A finite set of nodes $x_i \in \mathbb{R}^d$ ($i = 1..k$), and
- For each node x_i , finitely many conditions $D_{m_{ij}}(h)(x_i) = r_{ij}$ ($m_{ij} \in [d]^*$, $r_{ij} \in \mathbb{R}$) on the partial derivatives of a function $h: \mathbb{R}^d \rightarrow \mathbb{R}$. The set of conditions must be *consistent*, in the sense that each r_{ij} is determined by the node x_i and the permutation equivalence class of m_{ij} .

A *solution* to this problem is a d -ary polynomial function satisfying all the conditions. According to a theorem of Severi [12] (and see [9,10]) every such problem has a solution by a polynomial of degree $k(\max(|m_{ij}|) + 1) - 1$.

A finite set C of canonical forms is *saturated* if the following two conditions hold:

- (i) For any canonical form c and immediate atomic subexpression a of c

$$c \in C \implies a \in C$$

- (ii)

$$f_m(c_0, \dots, c_{n-1}) \in C \implies c_0, \dots, c_{n-1} \in C$$

Every finite set of canonical forms can evidently be extended to a finite saturated set of canonical forms.

Theorem 5.2 (Polynomial Separation) *Let C be a finite set of canonical forms. Then there is a polynomial function environment φ and an environment ρ that distinguish any two inequivalent elements of C .*

Proof. We can assume w.l.o.g. that C is saturated. By Lemma 5.1 we can find an assignment $\bar{\rho}$ of reals to the separation variables of the node polynomials P_c , ($c \in C$) which separates them, i.e., makes inequivalent ones take on different values, r_c . As $P_c \sim P_{c'}$ iff $c \approx c'$, we have $c \approx c'$ iff $r_c = r_{c'}$. In the case that c is an atomic expression a we have $P_a = v_a$ and so $r_a = \bar{\rho}(v_a)$. We define an environment ρ by:

$$\rho(x) = \begin{cases} r_x & (x \in C) \\ 0 & (\text{otherwise}) \end{cases}$$

We next set up an n -dimensional Hermite interpolation problem for each function variable $f:n$ appearing in the atomic expressions in C . For each atomic expression $f_m(c_0, \dots, c_{n-1}) \in C$ we add a node

$$(r_{c_0}, \dots, r_{c_{n-1}})$$

and a condition:

$$D_m(h)(r_{c_0}, \dots, r_{c_{n-1}}) = r_{f_m(c_0, \dots, c_{n-1})}$$

Note that the nodes are determined by the values of the node polynomials.

To show these conditions consistent, let $a = f_m(c_0, \dots, c_{n-1})$ and $a' = f_{m'}(c'_0, \dots, c'_{n-1})$ be two atomic expressions in C such that $m \sim m'$ and $r_{c_i} = r_{c'_i}$ for $i = 0, n-1$. Then, by the above, we have $c_i \approx c'_i$, and so, by part (2) of Lemma 4.5, $\vdash c_i = c'_i$, for $i = 0, n-1$. So $a \approx a'$ as atomic expressions and so too, therefore, as canonical forms. It follows that $r_a = r_{a'}$, as required for consistency.

We can therefore obtain a polynomial function environment by taking $\varphi(f)$ to be an m -ary interpolating polynomial for the corresponding Hermite interpolation problem, if f appears in some $c \in C$, and constantly 0, otherwise.

We claim that:

$$\mathcal{S}[c]\varphi\rho = r_c$$

for every $c \in C$. We establish the claim by structural induction on the expressions in C .

First, consider a variable $x \in C$. We have $\mathcal{S}[x]\varphi\rho = \rho(x) = r_x$, as required.

Next, consider an atomic expression $f_m(c_0, \dots, c_{n-1}) \in C$. By saturation we have $c_0, \dots, c_{n-1} \in C$. So, by the induction hypothesis, we have $\mathcal{S}[c_i]\varphi\rho = r_{c_i}$ for $i = 0, n-1$. We then have:

$$\begin{aligned} \mathcal{S}[f_m(c_0, \dots, c_{n-1})]\varphi\rho &= D_m(\varphi(f))(\mathcal{S}[c_0]\varphi\rho, \dots, \mathcal{S}[c_{n-1}]\varphi\rho) \\ &= D_m(\varphi(f))(r_{c_0}, \dots, r_{c_{n-1}}) \\ &= r_{f_m(c_0, \dots, c_{n-1})} \end{aligned}$$

with the last line holding as $\varphi(f)$ solves the Hermite interpolation problem for f set up above.

Lastly, consider a canonical form $c \in C$ which is not an atomic expression. Let $v_{a'_0}, \dots, v_{a'_{n-1}}$ be the variables of P_c . For every a'_i there is an immediate atomic subexpression a_i of c such that $a_i \approx a'_i$ (and so $v_{a'_i} = v_{a_i}$). By part (1) of Lemma 4.5 we then have

$$\vdash c = P_c[a_0/v_{a_0}, \dots, a_{n-1}/v_{a_{n-1}}]$$

As C is saturated, it contains the a_i . Since c is not itself an atomic expression, we can apply the induction hypothesis to the a_i and so we have $\mathcal{S}[a_i]\varphi\rho = r_{a_i}$.

We may then calculate that:

$$\begin{aligned} \mathcal{S}[c]\varphi\rho &= \mathcal{S}[P_c[a_0/v_{a_0}, \dots, a_{n-1}/v_{a_{n-1}}]]\varphi\rho \\ &= \mathcal{S}[P_c(v_{a_0}, \dots, v_{a_{n-1}})]\varphi\rho[\mathcal{S}[a_0]\varphi\rho/v_{a_0}, \dots, \mathcal{S}[a_{n-1}]\varphi\rho/v_{a_{n-1}}] \\ &= \mathcal{S}[P_c(v_{a_0}, \dots, v_{a_{n-1}})]\varphi\rho[r_{a_0}/v_{a_0}, \dots, r_{a_{n-1}}/v_{a_{n-1}}] \\ &= \mathcal{S}[P_c(v_{a_0}, \dots, v_{a_{n-1}})]\varphi\rho[\bar{\rho}(v_{a_0})/v_{a_0}, \dots, \bar{\rho}(v_{a_{n-1}})/v_{a_{n-1}}] \\ &= r_c \end{aligned}$$

with the last line holding by the definition of r_c . This concludes the inductive proof.

Finally, taking $c \not\approx c' \in C$, we have $r_c \neq r_{c'}$, and so, by the above, $\mathcal{S}[c]\varphi\rho \neq \mathcal{S}[c']\varphi\rho$, and so, as required, φ and ρ distinguish any two inequivalent elements of C , concluding the proof. \square

We next strengthen the polynomial separation theorem to natural number separation:

Theorem 5.3 (Natural Number Separation) *Let C be a finite set of canonical forms. Then there is a natural number polynomial function environment $\varphi_{\mathbb{N}}$ and a natural number environment $\rho_{\mathbb{N}}$ that distinguish any two inequivalent elements of C .*

Proof. By polynomial separation (Theorem 5.2), there is a polynomial function environment φ and an environment ρ such that distinguish any two inequivalent elements of C .

Let $f_i : m_i$ (for $i = 0, n-1$) be the function variables occurring in the $c \in C$, and, for $i = 0, n-1$, let $P_i(x_0, \dots, x_{m_i-1})$ be polynomials defining $\varphi(f_i)$. Let r_0, \dots, r_{q-1} be the constants occurring in these P_i and, for $i = 0, n-1$, let $Q_i(x_0, \dots, x_{m_i-1}, y_0, \dots, y_{q-1})$ be natural number polynomials such that

$$Q_i[r_0/y_0, \dots, r_{q-1}/y_{q-1}] = P_i$$

and where no y_k occurs freely in any $c \in C$. (The Q_i can be obtained from the P_i by replacing r_k by y_k , for $k = 0, q-1$.)

For $c \in C$, set $d_c = \text{CF}(c[\dots, (x_0, \dots, x_{m_i}). Q_i/f_i, \dots])$. As f_i is a list of function variables including all those of e , and as $\text{FnV}(c) \subseteq \text{FnV}(e)$ by Lemma 4.3, we see that $c[\dots, (x_0, \dots, x_{m_i}). Q_i/f_i, \dots]$ has no function variables and so, again by Lemma 4.3, neither does its canonical form d_c . As any canonical form with no function variables is a polynomial, it follows that d_c is a polynomial.

Setting $\bar{\rho} = \rho[\dots, r_k/y_k, \dots]$, we calculate:

$$\begin{aligned} \mathcal{S}[c]\varphi\rho &= \mathcal{S}[c]\varphi[\dots, (u_0, \dots, u_{m_i} \in \mathbb{R} \mapsto \mathcal{S}[P_i]\rho[\dots, u_j/x_j, \dots])/f_i, \dots]\rho \\ &= \mathcal{S}[c]\varphi[\dots, (u_0, \dots, u_{m_i} \in \mathbb{R} \mapsto \mathcal{S}[Q_i[\dots, r_k/y_k, \dots]]\rho[\dots, u_j/x_j, \dots])/f_i, \dots]\rho \\ &= \mathcal{S}[c]\varphi[\dots, (u_0, \dots, u_{m_i} \in \mathbb{R} \mapsto \mathcal{S}[Q_i]\rho[\dots, u_j/x_j, \dots, r_k/y_k, \dots])/f_i, \dots]\rho \\ &= \mathcal{S}[c]\varphi[\dots, (u_0, \dots, u_{m_i} \in \mathbb{R} \mapsto \mathcal{S}[Q_i]\bar{\rho}[\dots, u_j/x_j, \dots])/f_i, \dots]\bar{\rho} \\ &= \mathcal{S}[c[\dots, (x_0, \dots, x_{m_i}). Q_i/f_i, \dots]]\varphi\bar{\rho} \\ &= \mathcal{S}[d_c]\bar{\rho} \end{aligned}$$

where we have used Lemma 3.1, canonicalisation, and consistency (Theorem 3.2). As φ, ρ separate inequivalent $c \in C$ we therefore see that the corresponding d_c are inequivalent as functions and so as polynomials. By Lemma 5.1 there is a natural number environment $\rho_{\mathbb{N}}$ separating any two inequivalent d_c , and so any d_c and $d_{c'}$ with c and c' inequivalent.

Next, define a natural number polynomial function environment $\varphi_{\mathbb{N}}$ by setting:

$$\varphi_{\mathbb{N}}(f) = \begin{cases} u_0, \dots, u_{n_i} \in \mathbb{R} \mapsto \mathcal{S}[Q_i]\rho_{\mathbb{N}}[u_0/x_0, \dots, u_{n_i}/x_{n_i}] & (f = f_i) \\ 0 & (\text{otherwise}) \end{cases}$$

For any $c \in C$ we have:

$$\begin{aligned} \mathcal{S}[c]\varphi_{\mathbb{N}}\rho_{\mathbb{N}} &= \mathcal{S}[c]0[\dots, (u_0, \dots, u_{n_i} \in \mathbb{R} \mapsto \mathcal{S}[Q_i]\rho_{\mathbb{N}}[\dots, u_j/x_j, \dots])/f_i, \dots]\rho_{\mathbb{N}} \\ &= \mathcal{S}[c[\dots, (x_0, \dots, x_{m_i}). Q_i/f_i, \dots]]0\rho_{\mathbb{N}} \\ &= \mathcal{S}[d_c]\rho_{\mathbb{N}} \end{aligned}$$

again making use of Lemma 3.1, canonicalisation, and consistency.

So as $\rho_{\mathbb{N}}$ separates d_c and $d_{c'}$ whenever c and c' are inequivalent, we see that the natural number polynomial function environment $\varphi_{\mathbb{N}}$ and the natural number environment $\rho_{\mathbb{N}}$ separate any two inequivalent elements of C , concluding the proof. \square

We can now prove completeness relative to natural number polynomial function environments and natural number environments:

Theorem 5.4 (Natural number completeness) *For any expressions e and e' we have:*

$$\models_{\mathbb{N}} e = e' \implies \vdash e = e'$$

Proof. Suppose that $\models_{\mathbb{N}} e = e'$ but, for the sake of contradiction, that $\not\vdash e = e'$. By canonicalisation (Lemma 4.3) we then have $\not\vdash c = c'$, where $c = \text{CF}(e)$ and $c' = \text{CF}(e')$; so, by part (2) of Lemma 4.5, we have $c \not\approx c'$. Therefore, by Theorem 5.3, there is a natural number polynomial function environment $\varphi_{\mathbb{N}}$ and a natural number environment $\rho_{\mathbb{N}}$ such that $\mathcal{S}[[c]]\varphi_{\mathbb{N}}\rho_{\mathbb{N}} \neq \mathcal{S}[[c']]\varphi_{\mathbb{N}}\rho_{\mathbb{N}}$, obtaining the required contradiction. \square

We next upgrade part (2) of Lemma 4.5 to an equivalence and use that to analyse the theorems of our theory in terms of canonical forms and their \approx relation; we also obtain a useful characterisation of the equivalence of atomic expressions in terms of the equivalence of their canonical subexpressions:

Theorem 5.5

(i) *For any expressions e and e' we have:*

$$\vdash e = e' \iff \text{CF}(e) \approx \text{CF}(e')$$

(ii) *For any canonical forms c and c' we have:*

$$\vdash c = c' \iff c \approx c'$$

(iii) *For any atomic forms a and a' , $a \approx a'$ holds iff:*

- (a) *a and a' are identical variables, or else*
- (b) *for some function variable $f : n$, and for some $m \sim m'$, they have the forms $f_m(c_0, \dots, c_{n-1})$ and $f_{m'}(c'_0, \dots, c'_{n-1})$ with $c_i \approx c'_i$ for $i = 0, n-1$.*

Proof. For the second part, we already have the implication from right to left, by part (2) of Lemma 4.5. The other direction follows from consistency and separation. For the first part, by canonicalisation we have that $\vdash e = e'$ holds iff $\vdash \text{CF}(e) = \text{CF}(e')$ does, and conclusion follows from the second part. The third part follows immediately from the second part and the definition of the equivalence relation between atomic expressions. \square

We next seek a local criterion for canonical expression equivalence. Let A be a set of atomic forms, and let w_a ($a \in A$) be an A -indexed set of variables. We define polynomials $P_{w,c}$ for canonical forms whose immediate atomic subexpressions are included in A by structural induction on c :

$$P_{w,r} = r \ (r \in \mathbb{R}) \quad P_{w,a} = w_a \quad P_{w,c+c'} = P_{w,c} + P_{w,c'} \quad P_{w,cc'} = P_{w,c}P_{w,c'}$$

and we say that w is *equivalence-characterising* if, for all $a, a' \in A$, we have:

$$w_a = w_{a'} \iff a \approx a'$$

Lemma 5.6 *Let c, c' be canonical forms, and let A be a set of atomic forms including the immediate atomic subexpressions of c and c' and let w_a ($a \in A$) be an equivalence-characterising A -indexed set of variables. Then:*

$$c \approx c' \iff P_{w,c} \sim P_{w,c'}$$

Proof. It suffices to show that $P_c \sim P_{c'}$ holds iff $P_{w,c} \sim P_{w,c'}$ does. Choose $a_1, \dots, a_k \in A$ such that w_{a_1}, \dots, w_{a_k} enumerates $\{w_a \mid a \in A\}$. Then, as both w and v are equivalence-characterising, v_{a_1}, \dots, v_{a_k} enumerates $\{v_a \mid a \in A\}$. A straightforward induction shows that $P_c[w_{a_1}/v_{a_1}, \dots, w_{a_k}/v_{a_k}] = P_{w,c}$, and similarly for c' . Assuming $P_c \sim P_{c'}$, we then find:

$$P_{w,c} = P_c[w_{a_1}/v_{a_1}, \dots, w_{a_k}/v_{a_k}] \sim P_{c'}[w_{a_1}/v_{a_1}, \dots, w_{a_k}/v_{a_k}] = P_{w,c'}$$

that is, $P_c \sim P_{c'}$ implies $P_{w,c} \sim P_{w,c'}$. The converse is proved similarly. \square

As we now show, the equivalence relation on canonical forms can be viewed as a combination of polynomial equivalence and the commutativity of partial differentiation with respect to different variables. We write

$$\vdash_{\text{RTC}} e = e'$$

to mean that e and e' can be proved equal using only the ring axioms, the addition and multiplication tables, and the commutativity axiom for partial differentiation (i.e., without using any partial differentiation axioms other than commutativity).

Lemma 5.7 *For any canonical form c , and set of atomic expressions A with $\text{Im}(c) \subseteq A$ and such that for all $a, a' \in A$ if $a \approx a'$ then $\vdash_{\text{RTC}} a = a'$, and for any \approx -characterising A -indexed variable assignment w we have*

$$\vdash_{\text{RTC}} c = P_{w,c}[a_1/v_{a_1}, \dots, a_n/v_{a_n}]$$

where v_{a_1}, \dots, v_{a_n} are the variables of $P_{w,c}$ for $a_1, \dots, a_n \in A$.

Proof. The proof is a straightforward structural induction on c . In the case of an atomic form a , we have $P_{w,c} = v_a = v_{a_1}$ (the second equality by assumption). So as w is \approx -characterising, we have $a \approx a_1$ and so $\vdash_{\text{RTC}} a = a_1$, by assumption. \square

Theorem 5.8 *For any canonical forms c and c' we have:*

$$\vdash c = c' \iff c \approx c' \iff \vdash_{\text{RTC}} c = c'$$

Proof. We prove that for all c, c'

$$c \approx c' \implies \vdash_{\text{RTC}} c = c'$$

by induction on $\max(|c|, |c'|)$, when the conclusion follows immediately from Theorem 5.5.

In the case where c (say) is an atomic expression a , then $P_{c'} \sim P_c = v_a$, and so $\vdash_{\text{RTC}} c' = a'$ for some $a' \in \text{Im}(c')$ with $v_{a'} = v_a$. As $v_{a'} = v_a$, we have $a \approx a'$, and we then use the induction hypothesis to prove $\vdash_{\text{RTC}} a = a'$ using the characterisation of equivalence of atomic forms given by part (3) of Theorem 5.5.

Otherwise neither c nor c' is an atomic form and so, setting $A = \text{Im}(c) \cup \text{Im}(c')$, using the induction hypothesis, we have $\vdash_{\text{RTC}} a = a'$ whenever $a \approx a'$, for $a, a' \in A$. Let w be any \approx -characterising A -indexed variable assignment. Applying Lemma 5.7 we find $\vdash_{\text{RTC}} c = P_{w,c}[a_1/v_{a_1}, \dots, a_n/v_{a_n}]$ and $\vdash_{\text{RTC}} c' = P_{w,c'}[a_1/v_{a_1}, \dots, a_n/v_{a_n}]$. Further, by Lemma 5.6, we have $P_{w,c} \sim P_{w,c'}$. It follows that $\vdash_{\text{RTC}} P_{w,c}[a_1/v_{a_1}, \dots, a_n/v_{a_n}] = P_{w,c'}[a_1/v_{a_1}, \dots, a_n/v_{a_n}]$. Putting these three things together we obtain $\vdash_{\text{RTC}} c = c'$, as desired. \square

As may be evident, one can further strengthen the definition of canonical forms so that fewer axioms are needed to prove equivalence. For example, the partial differentiations in atomic expressions can be put in a standard order using the commutativity axiom. One can show that two such canonical forms are provably equal if, and only, if they can be proved so without using any axioms for partial differentiation; the proof of this fact parallels that of Theorem 5.8. Should one wish, one can go further and rewrite polynomials as sums of distinct multinomials; two such canonical forms are provably equal if, and only if, they can be proved so using only the associativity and commutativity of $+$.

We turn next to equational completeness. Let $f_i : m_i$ ($i = 1, k$) and x_j ($j = 1, l$) be the function variables and the free variables of two expressions e and e' . Then a *natural number counterexample* to the equation $e = e'$ consists of natural number polynomials $P_i(y_0, \dots, y_{m_i})$ and natural numbers k_1, \dots, k_l such that, the closed terms $\underline{e} =_{\text{def}} e[\dots, (y_0, \dots, y_{m_i}). P_i/f_i, \dots][\dots k_j/x_j, \dots]$ and $\underline{e}' =_{\text{def}} e'[\dots, (y_0, \dots, y_{m_i}). P_i/f_i, \dots][\dots k_j/x_j, \dots]$ are provably equal to different constants. As we now see, the natural number separation theorem yields such counterexamples, and thereby enables us to establish equational completeness:

Theorem 5.9 *The theory of partial differentiation is equationally complete.*

Proof. Suppose that $\not\vdash e = e'$. We show that adding the equation $e = e'$ makes the theory equationally inconsistent. As $\not\vdash e_0 = e_1$, by natural number polynomial completeness (Theorem 5.4), there is a natural number polynomial function environment φ , and a natural number environment ρ such that $\mathcal{S}[\underline{e}]\varphi\rho \neq \mathcal{S}[\underline{e}']\varphi\rho$. We can then define a natural number counterexample to the equation. Let $f_i : m_i$ ($i = 1, k$) and x_j ($j = 1, l$) be the function variables and the free variables of e and e' . For $i = 1, k$, take $P_i(y_0, \dots, y_{m_i})$ to be a natural number polynomial defining $\varphi(f_i)$ and take k_j to be $\rho(x_j)$, and define \underline{e} and \underline{e}' as above. By the substitution rule we have $\vdash \underline{e} = \underline{e}'$. By Lemma 3.1, we have $\mathcal{S}[\underline{e}] = \mathcal{S}[e]\varphi\rho$ and $\mathcal{S}[\underline{e}'] = \mathcal{S}[e']\varphi\rho$. So $\mathcal{S}[\underline{e}] \neq \mathcal{S}[\underline{e}']$, and therefore by Corollary 4 and consistency, \underline{e} and \underline{e}' are provably equal to different constants.

Thus, assuming $\vdash e = e'$, we can prove two distinct real constants $\mathcal{S}[\llbracket e \rrbracket]$ and $\mathcal{S}[\llbracket e' \rrbracket]$ equal. Using addition and multiplication tables and the ring axioms, we then successively find that $\vdash 1 = 0$ and $\vdash x = 0$ and so, as required, $\vdash x = y$. \square

Instead of allowing all reals as constants, one could restrict to \mathbb{Q} or to none (except for the ring constants). The above development goes through straightforwardly in both cases, except for equational completeness. The proof of equational completeness for \mathbb{Q} goes through as the required polynomials, being over the natural numbers, and the required reals, being natural numbers, are definable, and the argument from the equality of two distinct rationals to that of two distinct variables also goes through.

However, the last part of that argument does not go through when the only constants are the ring constants. For example it is then consistent to add the equation $1 + 1 = 0$. As a model, one can employ the *boolean differential calculus* [11]. One works over the boolean ring and defines the partial differentiation of boolean functions $h: \mathbb{B}^n \rightarrow \mathbb{B}$ by:

$$\frac{\partial h(x_0, \dots, x_{n-1})}{\partial x_i} = h(x_0, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_{n-1}) + h(x_0, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_{n-1})$$

This can be equivalently written, somewhat more transparently, as:

$$\frac{\partial h(x_0, \dots, x_{n-1})}{\partial x_i} = h(x_0, \dots, x_{i-1}, x_i + dx_i, x_{i+1}, \dots, x_{n-1}) - h(x_0, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_{n-1})$$

taking $dx_i = 1$. Another model of $1 + 1 = 0$ can be constructed from the clone of polynomials in several variables over the boolean ring, with partial differentiation defined as usual on polynomials over a ring.

With the constants restricted to the rationals, the question of decidability of the equational theory of partial differentiation makes sense, and we have:

Theorem 5.10 *With the constants restricted to the rationals, the equational theory of partial differentiation is decidable, and, further, natural number counterexamples to unprovable equations can be effectively obtained.*

Proof. We interleave two search procedures. One searches for a proof, the other searches through all possible natural number counterexamples. Since we know from the proof of equational completeness (Theorem 5.9) that a natural number counterexample exists if an equation is not provable, this procedure will terminate, yielding either a proof or a counterexample. \square

There is another, more direct, way to prove decidability and to find counterexamples. For decidability, one uses the effectiveness of the proof of canonicalisation (Lemma 4.3) to find canonical forms, and the characterisations of the equivalence of atomic expressions and canonical forms given by Theorem 5.5 and Lemma 5.6, to obtain a recursive algorithm to decide the equivalence of canonical forms.

With this decision procedure in hand, and with the observation that the proof in [9] of the Severi theorem is effective, one observes that the proof of polynomial separation (Theorem 5.2) is effective. So too, therefore is the proof of natural number separation (Theorem 5.3), and, following the proof of equational completeness (Theorem 5.9), we finally see that, if they exist, counterexamples can be found effectively.

Acknowledgements

I thank Martín Abadi, Jonathan Gallagher, and Tarmo Uustalu for useful discussions, and Sam Staton for helpful remarks including indicating how the axioms for partial differentiation could be made finitary.

References

- [1] Barendregt, Hendrik P., “The Lambda Calculus: Its Syntax and Semantics”, Studies in Logic and the Foundations of Mathematics, **103**, 2nd edition, North-Holland, 1985.
- [2] Blute, Richard F., Robin Cockett, and Robert A.G. Seely, *Cartesian Differential Categories*, Theory and Applications of Categories, **22** (2009), 622–672.
- [3] Cockett, Robin and Geoff Cruttwell, *Differential Structure, Tangent Structure, and SDG*, Applied Categorical Structures, **22**(2) (2014), 331–417.
- [4] Cockett, Robin, Geoff Cruttwell, Jonathan Gallagher, Jean-Simon Pacaud Lemay, Benjamin MacAdam, Gordon Plotkin, and Dorette Pronk, *Reverse Derivative Categories*, Proc. 28th. Annual Conference on Computer Science Logic (eds. Maribel Fernández and Anca Muscholl), LIPIcs, **152**, 18:1–18:16, Schloss Dagstuhl, 2020.

- [5] Cohn, Paul Moritz, “Universal algebra”, Mathematics and Its Applications, **6**, 2nd edn., Dordrecht: Reidel, 1981.
- [6] Fiore, Marcelo and Chung-Kil Hur, *Second-order equational logic*, Proc. 24th. International Workshop on Computer Science Logic (eds. Anuj Dawar and Helmut Veith), LNCS **6247**, 320–335, Springer, 2010.
- [7] Fiore, Marcelo and Ola Mahmoud, *Second-order algebraic theories*, Proc. 35th. International Symposium on Mathematical Foundations of Computer Science (eds. Hliněný P., Kučera A.), LNCS **6281**, 368–380, Springer, 2010.
- [8] Fiore, Marcelo, Gordon Plotkin, and Daniele Turi, *Abstract syntax and variable binding*, Proc. 14th. Annual IEEE Symposium on Logic in Computer Science, 193–202, 1999.
- [9] Lorentz, Rudolph A., “Multivariate Birkhoff Interpolation”, Lecture Notes in Mathematics, **1516**, Springer, 1992.
- [10] Lorentz, Rudolph A., *Multivariate Hermite interpolation by algebraic polynomials: A survey*, Journal of Computational and Applied Mathematics, **122** (2000), 167–201.
- [11] Posthoff, Christian, and Bernd Steinbach, “Logic Functions and Equations”, Springer, 2004.
- [12] Severi, Francesco, “Vorlesungen über Algebraische Geometrie”, Teubner, Berlin, 1921.
- [13] Staton, Sam, Dario Stein, Hongseok Yang, Nathanael L. Ackerman, Cameron E. Freer, and Daniel M. Roy, *The Beta-Bernoulli process and algebraic effects*, Proc. 45th. ICALP (eds. Ioannis Chatzigiannakis, Christos Kaklamanis, Dániel Marx, and Donald Sannella), LIPIcs, **107**, 141:1–141:15, Schloss Dagstuhl, 2018.
- [14] Taylor, Walter, *Survey 79: Equational logic*, Houston Journal of Mathematics, **5**(S) (1979), 1–83.