The problem with Cartesian Fermat Categories and how to fix them (hopefully!)

JS PL (he/him), joint work[1] with:



Carlos Pascual Miralles

[1] Ben MacAdam previously did lots of work on this with Robin Cockett and Jonathan Gallagher.

Slogan

- **Goal:** Give a multisorted version of **Fermat theories**.
- **Want:** So that they are still **Cartesian differential categories**.

Story Time!

- **Cartesian differential categories (CDC)** is a categorical framework for the foundations of the multivariable differential calculus (over Euclidean spaces).

   Blute, R., Cockett, R., Seely, R.A.G. **Cartesian Differential Categories**

- CDC have been able to recapture lots of concept from differential calculus: linear maps, partial derivatives, differential equations, de Rham complex, etc. Also used in computer science for the foundations of differential lambda calculus, differentiable programming, automatic differentiation, machine learning, etc.

- Another approach for categorical foundations of differential calculus is via **Fermat theories**.

   E. Dubuc, A. Kock, **On 1-form classifiers**

   Carchedi, D. & Roytenberg, D. **On Theories of Superalgebras of Differentiable Functions**.

  which briefly are Lawvere theories which add Hadamard's lemma as an axiom.

- Ben MacAdam (an ex-grad student of Robin Cockett at the same time that I was) showed that every Fermat theory is a CDC (FMCS2016 and in unpublished notes).

- However in general, CDC are not Lawvere theories. So we would like to generalize Fermat theories such that they still give CDC.

Story Time!

- In 2025, Carlos was my undergrad student at Macquarie who just wanted to learn about doing math research. He chose to learn about Fermat theories and how they were CDC (because I had never really sat down and looked at Fermat theories and Ben's proof in detail). We made quick progress, and early on we came up with a definition that generalized Fermat theories, which we called Cartesian Fermat categories (CFC). We got the result we wanted: every CFC is a CDC.

- We presented our work in the Macquarie Seminar, the Topos Institute Seminar, and at CMS2025. And people seemed to really like this story (and I find it cool!).

- When I was visiting Geoff Cruttwell at Mount Allison before CT2025, I discovered a big issue... Unfortunately our definition was too strong: Fermat theories (in particular smooth functions!) were not CFC. So while our CDC result was true, we weren't recapturing the thing we were trying to generalize! (the problem is the multivariable case)

- So we went back to the drawing board... And I realized one fix was to use something called a **system of linear maps**.

- Later Robin Cockett informed me that Ben MacAdam had done this approach (unpublished notes), which Ben called **secant theories**. So seems we recaptured Ben's previous work.

- **Today's Story:** This "new" definition of a Cartesian Fermat category (CFC) (which I am still not satisfied with...)

Cartesian Left k-Linear Categories – Definition

The underlying category of a Cartesian differential/Fermat category is a Cartesian left k-linear category, which is a category where we can take sums and scalar multiplication of maps, but where not every map is k-linear.

Definition

For a commutative semiring k, a **left k-linear category** is a category $\mathbb{X}$ such that each homset $\mathbb{X}(A, B)$ is a k-module with:

$$+ : \mathbb{X}(A, B) \times \mathbb{X}(A, B) \to \mathbb{X}(A, B) \qquad 0 \in \mathbb{X}(A, B) \qquad \cdot : k \times \mathbb{X}(A, B) \to \mathbb{X}(A, B)$$

such that pre-composition preserves the k-linear structure:

$$(s \cdot f + t \cdot g) \circ x = s \cdot (f \circ x) + t \cdot (g \circ x)$$

A map $f : A \to B$ is said to be k-**linear** if post-composition by f preserves the k-linear structure:

$$f \circ (s \cdot x + t \cdot y) = s \cdot (f \circ x) + t \cdot (f \circ y)$$

A **Cartesian left k-linear category** (CLLC_k) is a left k-linear with finite products such that the projection maps $\pi_j : A_1 \times \ldots \times A_n \to A_j$ are k-linear.

Cartesian Differential Categories – Definition

Definition

A **Cartesian k-differential category** (CDC_k) is a Cartesian left k-linear category $\mathbb{X}$ equipped with a **differential combinator** D, which is a family of operators:

$$\mathsf{D} : \mathbb{X}(A, B) \to \mathbb{X}(A \times A, B) \qquad \frac{f : A \to B}{\mathsf{D}[f] : A \times A \to B}$$

where $\mathsf{D}[f]$ is called the **derivative** of f, and which satisfies seven axioms which capture the basics of the derivative such as the chain rule, symmetry of partial derivatives, etc.

For readability, let's use element notation: $\mathsf{D}[f](a, b)$

- **[CD.1]** $\mathsf{D}[s \cdot f + t \cdot g](a, b) = s \cdot \mathsf{D}[f](a, b) + t \cdot \mathsf{D}[g](a, b)$
- **[CD.2]** $\mathsf{D}[f](a, s \cdot b + t \cdot c) = s \cdot \mathsf{D}[f](a, b) + t \cdot \mathsf{D}[f](a, c)$ **k-LINEAR IN 2ND ARG.**
- **[CD.3]** $\mathsf{D}[\pi_j](\vec{a}, \vec{b}) = b_j$
- **[CD.4]** $\mathsf{D}[\langle f_1, \ldots, f_n \rangle](a, b) = \langle \mathsf{D}[f_1](a, b), \ldots, \mathsf{D}[f_n](a, b) \rangle$
- **[CD.5]** $\mathsf{D}[g \circ f](a, b) = \mathsf{D}[g]\left(f(a), \mathsf{D}[f](a, b)\right)$ **CHAIN RULE**
- **[CD.6]** $\mathsf{D}\left[\mathsf{D}[f]\right](a, , 0, 0, b) = \mathsf{D}[f](a, b)$
- **[CD.7]** $\mathsf{D}\left[\mathsf{D}[f]\right](a, b, c, d) = \mathsf{D}\left[\mathsf{D}[f]\right](a, c, b, d)$ **SYMMETRY OF MIXED PARTIALS**

Example

Lawvere theory of poylonimials, Lawvere theory of smooth functions, categories with biproducts, differential lambda calculus, free and cofree constructions, abelian functor calculus, etc.

Fermat Theories – Definition

Definition

Let k be a commutative ring. A **Fermat theory** is a Lawvere theory[a] that is an extension of the theory of commutative k-algebras (so we can add and *multiply* maps) such that for every map $f : n+1 \to 1$, there is a unique map $\mathfrak{F}_1(f) : n+2 \to 1$ such that:

$$f(x+y, \vec{z}) = f(x, \vec{z}) + \mathfrak{F}_1(f)(x, \vec{z}, y)y$$

where here I'm using function notation for simplicity.

Using permutations, for each $1 \leq i \leq n+1$, there is a unique map $\mathfrak{F}_i(f) : n+2 \to 1$ such that:

$$f(\vec{z}_1, x+y, \vec{z}_2) = f(\vec{z}_1, x, \vec{z}_2) + \mathfrak{F}_i(f)(\vec{z}_1, x, \vec{z}_2, y)y$$

[a] If it helps, think of n as k^n or $\mathbb{R}^n$

Example

Lawvere theory of polynomials, Lawvere theory of smooth functions, etc.

Fermat Theories are CDCs

For any map $f : n + 1 \to 1$, we can define its partial derivatives $\partial_i(f) : n + 1 \to 1$ as follows:

$$\partial_i(f)(\vec{x}) = \mathfrak{F}_i(f)(\vec{x}, 0)$$

Theorem (B. MacAdam)

A Fermat theory is a CDC_k where for a map $f : n + 1 \to 1$, its derivative $\mathsf{D}[f] : 2n + 2 \to 1$ is defined as follows:

$$\mathsf{D}[f](\vec{x}, \vec{y}) = \sum_{i=1}^{n+1} \mathfrak{F}_i(f)(\vec{x}, 0)y_i = \sum_{i=1}^{n+1} \partial_i(f)(\vec{x})y_i$$

which we extend to tuples for maps $n + 1 \to m$. For the case $0 \to 1$, the derivatives are just 0.

Some mismatches

- So Fermat theories are CDC. Great!
- In general though, CDC are not: 1) Lawvere theories and 2) don't have multiplication...
- So let's address these first.

Let's look at the differential combinator again

If we look at the differential combinator[2], which takes a $f : n \to 1$ and then spits out a $\mathsf{D}[f] : n + n \to 1$

$$\mathsf{D}[f](\vec{x}, \vec{y}) = \sum_{i=1}^{n} \partial_i(f)(\vec{x}) y_i$$

The multiplication is hidden in the definition of the differential combinator.

This is actually **[CD.2]** which says that the differential is k-linear in its 2nd argument.

In a similar fashion we can group together the $\mathfrak{F}_i[f]$ into a $\mathfrak{F}[f] : n + n + n \to 1$

$$\mathfrak{F}[f](\vec{x}, \vec{y}, \vec{z}) = \sum_{i=1}^{n} \mathfrak{F}_i[f](\vec{x}, \vec{y}) z_i$$

So here we see that $\mathfrak{F}[f]$ is k-linear in its 3rd argument.

And we can extend this to $n \to m$ coordinate wise to get $n + n + n \to m$:

$$\mathfrak{F}[\langle f_1, \ldots, f_n \rangle] = \langle \mathfrak{F}[f_1], \ldots, \mathfrak{F}[f_n] \rangle$$

[2] Remember that in a Lawvere theory, product is given by sum $n \times m = n + m$

Extending to the multisorted case

So what we should be asking is that in a Cartesian left k-linear category is that:

$$\mathfrak{F} : \mathbb{X}(A, B) \to \mathbb{X}(A \times A \times A, B) \qquad \frac{f : A \to B}{\mathfrak{F}[f] : A \times A \times A \to B}$$

such that $\mathfrak{F}[f]$ is k-linear in its third argument:

$$\mathfrak{F}[f](a, b, s \cdot c + t \cdot d) = s \cdot \mathfrak{F}[f](a, b, c) + t \cdot \mathfrak{F}[f](a, b, d)$$

This is what we do in CDC to avoid multiplication. And it's what our original definition... BUT, this actually excludes many Lawvere theories of polynomials over certain k.

The issue comes form the fact that we don't just k-linearity, we want to say that $\mathfrak{F}[f]$ is a monomial in its third argument:

$$\mathfrak{F}[f](\vec{x}, \vec{y}, \vec{z}) = \sum_{i=1}^{n} \mathfrak{F}_i[f](\vec{x}, \vec{y}) z_i$$

For example if $k = \mathbb{Z}_2$, then x^2 is $\mathbb{Z}_2$-linear and this would break uniqueness...

The solution is to look at **system of linear maps**!

System of Linear Maps

For a Cartesian left k-linear category, a **system of linear maps** $\mathcal{L}$ is a collection of maps:

$$\mathcal{L} = \{f : C \times A \to B\}$$

where $f : C \times A \to B \in \mathcal{L}$ is said to be $\mathcal{L}$-linear in A or $\mathcal{L}$-linear in context C, such that this collection of maps is closed under composition (in context), projections, products, addition, sections, and context change. Essentially it's a way of declaring when we can say that a map has a linear argument. By definition being $\mathcal{L}$-linear implies being k-linear, but not the converse!



Blute, R., Cockett, R., Seely, R.A.G. **Cartesian Differential Storage Categories**

Example

- For polynomials, being monomial/degree one in argument forms a system of linear maps.

$$P(\vec{z}, \vec{x}) = \sum_{i=1}^{n} P_i(\vec{z}) x_i$$

- For smooth functions, similar to polynomial (but here this is the same as being $\mathbb{R}$-linear!)

$$F(\vec{z}, \vec{x}) = \sum_{i=1}^{n} F_i(\vec{z}) x_i$$

- In any CDC there is a canonical choice of linearity called D-linear:

$$\mathrm{D}[f](c, 0, 0, a) = f(c, a)$$

This is precisely the system of linear maps for polynomials and smooth functions above.

Extending to the multisorted case

Ok so now we have a Cartesian left k-linear category with a system of linear maps $\mathcal{L}$ such that:

$$\mathfrak{F} : \mathbb{X}(A, B) \to \mathbb{X}(A \times A \times A, B) \qquad \frac{f : A \to B}{\mathfrak{F}[f] : A \times A \times A \to B}$$

such that $\mathfrak{F}[f]$ is $\mathcal{L}$-linear in its third argument.

Now the Hadamard axiom should be that $\mathfrak{F}[f]$ is the **unique** map such that:

$$f(a + b) = f(a) + \mathfrak{F}[f](a, b, b)$$

PROBLEM: In a Fermat theory, while the $\mathfrak{F}_i[f]$ have this uniqueness property, their sum $\mathfrak{F}[f]$ do not! See comment under Prop 1.1 (which we originally overlooked!):



E. Dubuc, A. Kock, **On 1-form classifiers**

Essentially the issue is we can't split A into smaller parts, like splitting R^n into $R \times \ldots \times R$.

So original definition was too strong (but it was so nice! and remember it did give us a CDC)

But there is some uniqueness floating around!

Let's look forward to getting a CDC. We define the differential combinator as follows:

$$\mathrm{D}[f](a, b) = \mathcal{F}[f](a, 0, b)$$

In a Fermat theory, follows from from Prop/Def 1.7:



E. Dubuc, A. Kock, **On 1-form classifiers**

that for any $F : n \times n \times n \to 1$ which is monomial linear in its third argument such that:

$$f(\vec{x} + \vec{y}) = f(\vec{x}) + F(\vec{x}, \vec{y}, \vec{y})$$

Then we have that:

$$F(\vec{x}, \vec{0}, \vec{y}) = \mathrm{D}[f](\vec{x}, \vec{y}) = \mathcal{F}[f](\vec{x}, 0, \vec{y})$$

Cartesian Fermat Categories

Definition

A **Cartesian k-Fermat category** (CFC_k) is a Cartesian left k-linear category $\mathbb{X}$ with a system of linear maps $\mathcal{L}$ such that for any map $f : A \to B$:

- There exists a map $F : A \times A \times A \to B$ which is $\mathcal{L}$-linear in its third argument and the following equality holds:

  $$f(a + b) = f(a) + F(a, b, b)$$

  We call F a $\mathcal{L}$-**Fermat extension** of f (which is not necessarily unique!)

- For all $\mathcal{L}$-Fermat extensions F_1 and F_2 of f, the following equality holds:

  $$F_1(a, 0, b) = F_2(a, 0, b)$$

Theorem

A CFC_k is a CDC_k where for a map $f : A \to B$, its derivative $\mathrm{D}[f] : A \times A \to B$ is defined as:

$$\mathrm{D}[f](a, b) = F(a, 0, b)$$

for some $\mathcal{L}$-Fermat extension of f (exists by 1st axiom and is well-defined by 2nd axiom).

Unfortunately I can only prove **[CD.1]**-**[CD.6]**... missing **[CD.7]** which was the symmetry of mixed partial derivative axiom!

Cartesian Fermat Categories

Definition

A **Cartesian k-Fermat category** (CFC_k) is a Cartesian left k-linear category $\mathbb{X}$ with a system of linear maps $\mathcal{L}$ such that for any map $f : A \to B$:

- There exists a map $F : A \times A \times A \to B$ which is $\mathcal{L}$-linear in its third argument and the following equality holds:
$$f(a + b) = f(a) + F(a, b, b)$$
We call F a $\mathcal{L}$-**Fermat extension** of f (which is not necessarily unique!)
- For all $\mathcal{L}$-Fermat extensions F_1 and F_2 of f, the following equality holds:
$$F_1(a, 0, b) = F_2(a, 0, b)$$
- If $\widetilde{F} : A^9 \to B$ is an $\mathcal{L}$-Fermat extension of an $\mathcal{L}$-Fermat extension then[a]:
$$\widetilde{F}(a, 0, b, 0, 0, 0, c, 0, d) = \widetilde{F}(a, 0, c, 0, 0, 0, b, 0, d)$$

[a] This is literally **[CD.7]**

Theorem

A CFC_k is a CDC_k where for a map $f : A \to B$, its derivative $\mathrm{D}[f] : A \times A \to B$ is defined as:

$$\mathrm{D}[f](a, b) = F(a, 0, b)$$

for some $\mathcal{L}$-Fermat extension of f (exists by 1st axiom and is well-defined by 2nd axiom).

We do recapture Fermat theories

Theorem

A Fermat theory is a CFC_k.

Concluding Remarks

- Personally I am still unsatisfied with the definition of a CFC_k, I really want to get rid of **[CD.7]** or at least give a better version of it. But as far as I can tell, even in a Fermat theory, there's no symmetry of the mixed Fermat extensions. (Help!)

- The statement of Prop/Def 1.7 in Dubuc-Kock actually can be written as saying for every $f : A \to B$ there exists a unique map $\mathrm{D}[f] : A \times A \to B$ which is $\mathcal{L}$-linear in its second argument for which there exists a (not necessarily unique) map $H : A \times A \times A \times A \to B$ which is $\mathcal{L}$-bilinear and symmetric in its last two arguments, such that:

$$f(a+b) = f(a) + \mathrm{D}[f](a,b) + H(a,b,b,b)$$

  I can show **[CD.1]**-**[CD.6]** again but can't seem to get **[CD.7]** (ugh!). But seems to be better fit for CDC perspective, since this might be how we define a **Fermat CDC**.

- There is another axiomatization of Fermat theories using *difference quotients*, that is, for every map $f(x)$, there exists a unique map $\Delta(f)(x,y)$ such that:

$$f(x) - f(y) = \Delta(f)(x,y)(x-y)$$

  Note $\Delta(f)$ is not the same as $\mathfrak{F}[f]$ but we can build one from the other. In this case the derivative is $f'(x) = \Delta(f)(x,x)$. Carlos and I have also worked out the $\mathrm{CD}/\mathrm{FC}_k$ version of this.

  

  Carchedi, D. & Roytenberg, D. **On Theories of Superalgebras of Differentiable Functions**.

- Ideas welcomed!

Big thank you to Rick!

Please join me in saying a big thank you to Rick for organizing[3] these virtual CT Octoberfest yearly and another succesful one this year!

[3] Who strategically scheduled to finish the conference before the Patriots' game later today.