

# The Theory of Differential Categories

## Lecture 2: Cartesian Differential Categories

JS Pacaud Lemay



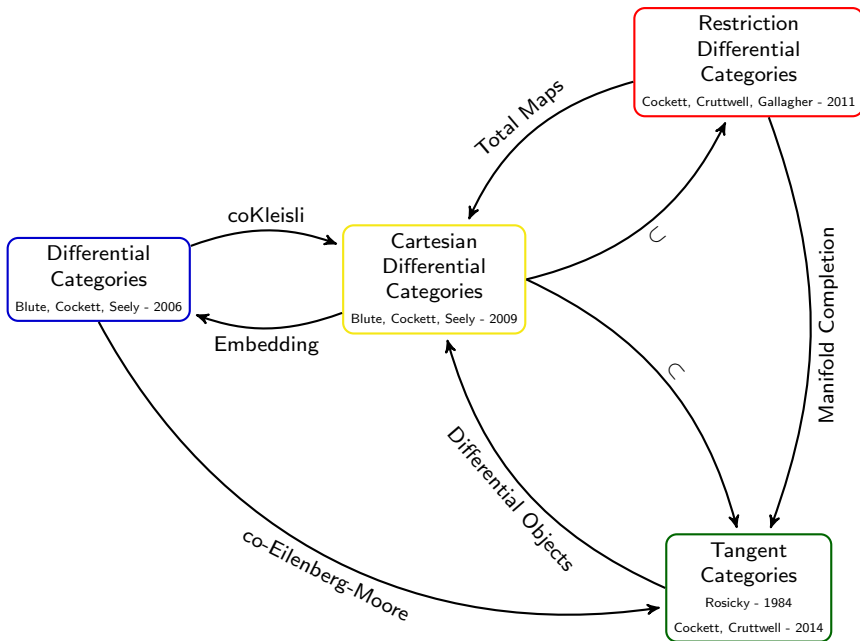
**Kellogg College**  
University of Oxford



DEPARTMENT OF  
**COMPUTER  
SCIENCE**



# The Differential Category World: It's all connected!



# Today's Story: Differential Categories

## Cartesian Differential Categories:

- Formalize differentiation in multivariable calculus of Euclidean spaces.
- Provide the categorical semantics of the differential  $\lambda$ -calculus.

## Main Reference:

 R. Blute, R. Cockett, R.A.G. Seely, [Cartesian Differential Categories](#)

A **Cartesian differential category** is:

- (i) A Cartesian left additive category;
- (ii) With a differential combinator.

A **Cartesian differential category** is:

- (i) A **Cartesian left additive category**;
- (ii) With a differential combinator.

## Cartesian Left Additive Category - Definition

A **left additive category** is a category  $\mathbb{X}$  which is *skew-enriched* over commutative monoids:



Campbell, A., 2018. [Skew-enriched categories](#).

Explicitly, every homset is a commutative monoid, so we can add maps and have zero maps:

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

such that composition preserves the addition in the following sense:

$$(f + g) \circ x = f \circ x + g \circ x \qquad 0 \circ x = 0$$

A map  $f$  is **additive** if  $f \circ (x + y) = f \circ x + f \circ y$  and  $f \circ 0 = 0$ .

A **Cartesian left additive category** (CLAC) is a left additive category with finite products such that the projection maps  $\pi_0 : A \times B \rightarrow A$  and  $\pi_1 : A \times B \rightarrow B$  are additive.

## Example

- Every category with finite biproducts is a CLAC where every map is additive. For example,  $\text{VEC}_k$  the category of  $k$ -vector spaces and  $k$ -linear maps is a CLAC.
- $\text{VEC}_k^\omega$  the category of  $k$ -vector spaces and arbitrary set functions is a CLAC, where the sum of set functions is defined point-wise  $(f + g)(x) = f(x) + g(x)$ .
- Let  $\text{Poly}_k$  be the Lawvere theory of polynomials, that is, the category whose objects are  $n \in \mathbb{N}$  and where a map  $P : n \rightarrow m$  is a tuple of polynomials:

$$P = \langle p_1, \dots, p_m \rangle \quad p_i \in R[x_1, \dots, x_n]$$

Then  $\text{Poly}_k$  is a CLAC (where  $n \times m = n + m$ ).

- Let  $\text{SMOOTH}$  be the category of smooth real functions, that is, the category whose objects are the Euclidean vector spaces  $\mathbb{R}^n$  and whose maps are smooth function  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , which is actually an  $m$ -tuple of smooth functions:

$$F = \langle f_1, \dots, f_m \rangle \quad f_i : \mathbb{R}^n \rightarrow \mathbb{R}$$

Then  $\text{SMOOTH}$  is a CLAC. Note that  $\text{Poly}_{\mathbb{R}}$  is a sub-CLAC of  $\text{SMOOTH}$ .

A **Cartesian differential category** is:

- (i) A Cartesian left additive category;
- (ii) With a **differential combinator**.



A **differential combinator** on a Cartesian left additive category  $\mathbb{X}$  is a combinator  $D$ , which is a family of functions  $\mathbb{X}(A, B) \rightarrow \mathbb{X}(A \times A, B)$ , which written as an inference rule:

$$\frac{f : A \rightarrow B}{D[f] : A \times A \rightarrow B}$$

Before giving the axioms, let's look at some examples!

## Example

SMOOTH is a Cartesian differential category where the differential combinator is defined as the directional derivative of a smooth function. A smooth function  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is in fact a tuple:

$$F = \langle f_1, \dots, f_m \rangle$$

of smooth functions  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ . Then the Jacobian matrix of  $F$  at vector  $\vec{x} \in \mathbb{R}^n$  is the matrix  $\nabla(F)(\vec{x})$  of size  $m \times n$  whose coordinates are the partial derivatives of the  $f_i$ :

$$\nabla(F)(\vec{x}) := \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\vec{x}) & \frac{\partial f_1}{\partial x_2}(\vec{x}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{x}) \\ \frac{\partial f_2}{\partial x_1}(\vec{x}) & \frac{\partial f_2}{\partial x_2}(\vec{x}) & \dots & \frac{\partial f_2}{\partial x_n}(\vec{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\vec{x}) & \frac{\partial f_m}{\partial x_2}(\vec{x}) & \dots & \frac{\partial f_m}{\partial x_n}(\vec{x}) \end{bmatrix}$$

So for a smooth function  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , its derivative  $D[F] : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^m$  is then defined as:

$$D[F](\vec{x}, \vec{y}) := \nabla(F)(\vec{x}) \cdot \vec{y} = \left\langle \sum_{i=1}^n \frac{\partial f_1}{\partial x_i}(\vec{x}) y_i, \dots, \sum_{i=1}^n \frac{\partial f_m}{\partial x_i}(\vec{x}) y_i \right\rangle$$

where  $\cdot$  is matrix multiplication and  $\vec{y}$  is seen as a  $n \times 1$  matrix. For example, Let  $f(x, y) = x^2 y$ .

$$D[f]((x, y), (a, b)) = 2xya + x^2 b$$

## Example

Any category with finite biproduct  $\oplus$  is a CDC, where for a map  $f : A \rightarrow B$ :

$$D[f] := A \oplus A \xrightarrow{\pi_1} A \xrightarrow{f} B$$

For example,  $\text{VEC}_k$  is a CDC where  $D[f](x, y) = f(y)$ .

## Example

$\text{POLY}_k$  is a CDC where for a map  $P : n \rightarrow m$  with  $P = \langle p_1, \dots, p_n \rangle$ ,  $D[P] : n \times n \rightarrow m$  is:

$$D[P] := \left\langle \sum_{i=1}^n \frac{\partial p_1}{\partial x_i} y_i, \dots, \sum_{i=1}^n \frac{\partial p_n}{\partial x_i} y_i \right\rangle$$

where  $\sum_{i=1}^n \frac{\partial p_1}{\partial x_i} y_i \in R[x_1, \dots, x_n, y_1, \dots, y_n]$ . Note that  $\text{POLY}_{\mathbb{R}}$  is a sub-CDC of  $\text{SMOOTH}$ .

## Differential Combinator - Definition

A **differential combinator** on a Cartesian left additive category  $\mathbb{X}$  is a combinator  $D$ , which is a family of functions  $\mathbb{X}(A, B) \rightarrow \mathbb{X}(A \times A, B)$ , which written as an inference rule:

$$\frac{f : A \rightarrow B}{D[f] : A \times A \rightarrow B}$$

To help us with the axioms, we will use the following notation/proto-term logic:

$$D[f](a, b) := \frac{df(x)}{dx}(a) \cdot b$$

### Example

The notation comes from SMOOTH:  $D[F](\vec{x}, \vec{y}) := \nabla(F)(\vec{x}) \cdot \vec{y}$ .

### Remark

There is a sound and complete term logic for Cartesian differential categories. In short: anything we can prove using the term logic, holds in any Cartesian differential category. So doing proofs in the term logic is super useful!

- Additivity of Combinator:

$$D[f + g] = D[f] + D[g]$$

$$D[0] = 0$$

$$\frac{df(x) + g(x)}{dx}(a) \cdot b = \frac{df(x)}{dx}(a) \cdot b + \frac{dg(x)}{dx}(a) \cdot b$$

$$\frac{d0}{dx}(a) \cdot b = 0$$

- Additivity in Second Argument

$$D[f] \circ \langle a, b + c \rangle = D[f] \circ \langle a, b \rangle + D[f] \circ \langle a, c \rangle$$

$$D[f] \circ \langle x, 0 \rangle = 0$$

$$\frac{df(x)}{dx}(a) \cdot (b + c) = \frac{df(x)}{dx}(a) \cdot b + \frac{df(x)}{dx}(a) \cdot c$$

$$\frac{df(x)}{dx}(a) \cdot 0 = 0$$

- Identities + Projections

$$D[1] = \pi_1$$

$$D[\pi_i] = \pi_i \circ \pi_1$$

$$\frac{dx}{dx}(a) \cdot b = b$$

$$\frac{dx_i}{d(x_0, x_1)}(a_0, a_1) \cdot (b_0, b_1) = b_i$$

- Pairings

$$D[\langle f, g \rangle] = \langle D[f], D[g] \rangle$$

$$\frac{d\langle f(x), g(x) \rangle}{dx}(a) \cdot b = \left\langle \frac{df(x)}{dx}(a) \cdot b, \frac{dg(x)}{dx}(a) \cdot b \right\rangle$$

### Example

In SMOOTH, if  $F = \langle f_1, \dots, f_n \rangle$ , then  $D[F](\vec{x}, \vec{y}) := \langle D[f_1](\vec{x}, \vec{y}), \dots, D[f_n](\vec{x}, \vec{y}) \rangle$ .

Chain Rule:

$$D[g \circ f] = D[g] \circ \langle f \circ \pi_0, D[f] \rangle$$

$$\frac{dg(f(x))}{dx}(a) \cdot b = \frac{dg(x)}{dx}(f(a)) \cdot \left( \frac{df(x)}{dx}(a) \cdot b \right)$$

$$\frac{f : A \rightarrow B}{D[f] : A \times A \rightarrow B}$$


---


$$D[D[f]] : (A \times A) \times (A \times A) \rightarrow B$$

- Linearity in Second Argument

$$D[D[f]] \circ \langle a, 0, 0, b \rangle = D[f] \circ \langle a, b \rangle$$

$$\frac{d \frac{df(x)}{dx}(y) \cdot z}{d(y, z)}(a, 0) \cdot (0, b) = \frac{df(x)}{dx}(a) \cdot b$$

- Symmetry

$$D[D[f]] \circ \langle \langle a, b \rangle, \langle c, d \rangle \rangle = D[D[f]] \circ \langle \langle a, c \rangle, \langle b, d \rangle \rangle$$

$$\frac{d \frac{df(x)}{dx}(y) \cdot z}{d(y, z)}(a, b) \cdot (c, d) = \frac{d \frac{df(x)}{dx}(y) \cdot z}{d(y, z)}(a, c) \cdot (b, d)$$



## Cartesian Differential Categories - Definition

A **Cartesian differential category** is:

- (i) A Cartesian left additive category;
- (ii) With a differential combinator.

$$\frac{f : A \rightarrow B}{D[f] : A \times A \rightarrow B}$$

Before we give some more examples: let's see what we can do within a CDC!

## Partial Derivatives I

Suppose we have a map  $f : A \times B \rightarrow C$  and we only want to differentiate with respect to  $A$ .

We can zero out in  $D[f] : (A \times B) \times (A \times B) \rightarrow C$  to obtain a partial derivative!

Define the partial derivative  $D_0[f] : (A \times B) \times A \rightarrow C$  as follows:

$$D_0[f] := (A \times B) \times A \xrightarrow{(1_A \times 1_B) \times \langle 1_A, 0 \rangle} (A \times B) \times (A \times B) \xrightarrow{D[f]} C$$

$$D_0[f](a, b, c) := \frac{df(x, b)}{dx}(a) \cdot c := \frac{df(x, y)}{d(x, y)}(a, b) \cdot (c, 0)$$

Similarly, define the partial derivative  $D_1[f] : (A \times B) \times B \rightarrow C$  as follows:

$$D_1[f] := (A \times B) \times B \xrightarrow{(1_A \times 1_B) \times \langle 0, 1_B \rangle} (A \times B) \times (A \times B) \xrightarrow{D[f]} C$$

$$D_1[f](a, b, d) := \frac{df(a, y)}{dy}(b) \cdot d := \frac{df(x, y)}{d(x, y)}(a, b) \cdot (0, d)$$

You can also do this with maps  $f : A_0 \times \dots \times A_n \rightarrow B$ .

## Partial Derivatives II

A consequence of symmetry rule, CD.7, is that for  $f : A \times B \rightarrow C$ , doing the partial derivative with respect to  $A$  then  $B$  is the same as doing the partial derivative with respect to  $B$  then  $A$ .

$$\frac{d \frac{df(x,y)}{dy}(b) \cdot d}{dx}(a) \cdot c = \frac{d \frac{df(x,y)}{dx}(a) \cdot c}{dy}(b) \cdot d$$

Additivity in the second argument, CD.2, tells us that for  $f : A \times B \rightarrow C$ ,  $D[f]$  is the sum of the partial derivatives!

$$\begin{aligned} \frac{df(x,y)}{d(x,y)}(a,b) \cdot (c,d) &= \frac{df(x,y)}{d(x,y)}(a,b) \cdot ((c,0) + (0,d)) \\ &= \frac{df(x,y)}{d(x,y)}(a,b) \cdot (c,0) + \frac{df(x,y)}{d(x,y)}(a,b) \cdot (0,d) \\ &= \frac{df(x,b)}{dx}(a) \cdot c + \frac{df(a,y)}{dy}(b) \cdot d \end{aligned}$$

### Example

For a smooth map  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $D[f]$  is the sum of its partial derivatives:

$$D[f] : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \qquad D[f](\vec{v}, \vec{w}) := \nabla(f)(\vec{v}) \cdot \vec{w} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{v}) w_i$$

# Linear Maps I

In a Cartesian differential category, there is a natural notion of **linear maps**. A map  $f : A \rightarrow B$  is said to be linear if:

$$D[f] := A \times A \xrightarrow{\pi_1} A \xrightarrow{f} B$$
$$\frac{df(x)}{dx}(a) \cdot b = f(b)$$

## Example

- In a category with finite biproducts, every map is linear (by definition!).
- In  $\text{POLY}_k$ ,  $P = \langle p_1, \dots, p_m \rangle$  is linear if each  $p_i \in k[x_1, \dots, x_n]$  is a polynomial of degree 1, that is, a sum of the form  $p_i = \sum_{j=1}^n a_j x_j$ .
- In  $\text{SMOOTH}_{\mathbb{R}}$ , a smooth function  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is linear in the Cartesian differential sense precisely when it is  $\mathbb{R}$ -linear in the classical sense:

$$F(s\vec{x} + t\vec{y}) = sF(\vec{x}) + tF(\vec{y})$$

for all  $s, t \in \mathbb{R}$  and  $\vec{x}, \vec{y} \in \mathbb{R}^n$ .

- Linear  $\Rightarrow$  Additive, but not necessarily the converse!  
(But in the above examples: Additive  $\Rightarrow$  Linear)
- Identity maps and projection maps are linear by CD.3

## Linear Maps II

A map  $f : A \times B \rightarrow C$  can also be linear in its second argument if it is linear with respect to its partial derivative:

$$D_1[f] := (A \times B) \times B \xrightarrow{\pi_0 \times 1} A \times B \xrightarrow{f} C$$
$$\frac{df(a, y)}{dy}(b) \cdot c = f(a, c)$$

The linearity in the second argument rule, CD.6, says that for any  $f : A \rightarrow B$ ,  $D[f]$  is linear in its second argument:

$$\frac{d \frac{df(x)}{dx}(a) \cdot y}{dy}(b) \cdot c = \frac{df(x)}{dx}(a) \cdot c$$

### Example

For a smooth map  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $D[f]$  is linear in its second argument:

$$D[f] : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R} \qquad D[f](\vec{v}, \vec{w}) := \nabla(f)(\vec{v}) \cdot \vec{w} = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{v}) w_i$$

## Example

Every model of the differential  $\lambda$ -calculus induces a Cartesian differential category. Conversely, every Cartesian differential category which is Cartesian closed such that the evaluation maps are linear in their second argument gives rise to a model of the differential  $\lambda$ -calculus.



Manzonetto, G., 2012. **What is a Categorical Model of the Differential and the Resource  $\lambda$ -Calculus?**

## Example

Bauer, Johnson, Osborne, Riehl, and Tebbe (BJORT) constructed an Abelian functor calculus model of a Cartesian differential category.



Bauer, K., Johnson, B., Osborne, C., Riehl, E. and Tebbe, A., 2018. **Directional derivatives and higher order chain rules for abelian functor calculus.**

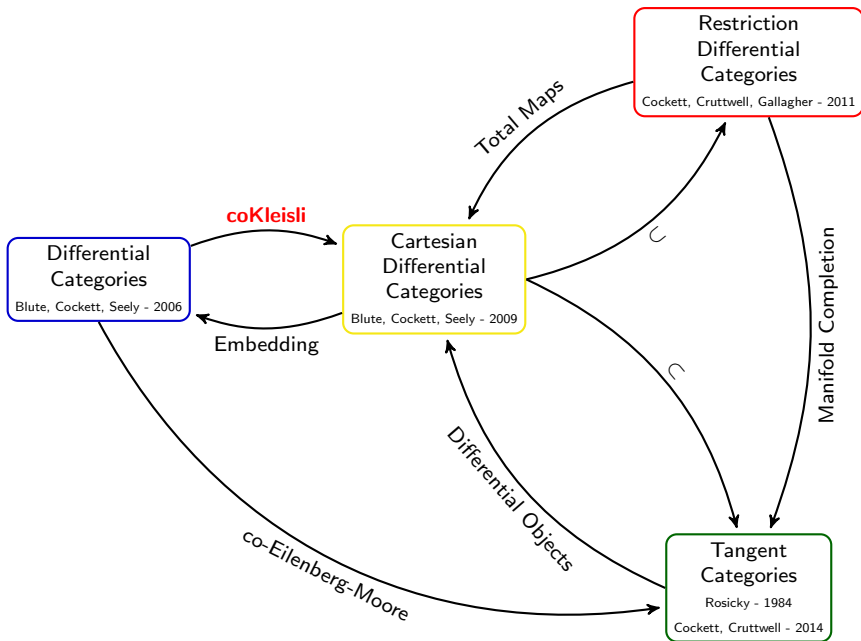
## Example

There is a couniversal construction of Cartesian differential categories, known as the Faa di Bruno construction, that is, for every Cartesian left additive category  $\mathbb{X}$  there is a cofree Cartesian differential category over  $\mathbb{X}$ .



Cockett, J.R.B. and Seely, R.A.G., 2011. **The Faa di bruno construction.**

# The Differential Category World: It's all connected!



# The coKleisli Category of a Differential Category I

Consider a differential category  $\mathbb{X}$  with a coalgebra modality  $!$ :

$$!A \xrightarrow{\delta} !!A$$

$$!A \xrightarrow{\varepsilon} A$$

$$!A \xrightarrow{\Delta} !A \otimes !A$$

$$!A \xrightarrow{e} K$$

and deriving transformation:

$$!A \otimes A \xrightarrow{d} !A$$

and finite products  $\times$  (which are actually biproducts by the additive structure of  $\mathbb{X}$ ).

Let  $\mathbb{X}_!$  be the coKleisli category and we are going to use interpretation brackets  $\llbracket - \rrbracket$ .

$$\frac{f : A \rightarrow B \text{ in } \mathbb{X}_!}{\llbracket f \rrbracket : !A \rightarrow B}$$

$$\llbracket 1 \rrbracket = !A \xrightarrow{\varepsilon} A$$

$$\llbracket g \circ f \rrbracket = !A \xrightarrow{\delta} !!A \xrightarrow{!(\llbracket f \rrbracket)} !B \xrightarrow{\llbracket g \rrbracket} C$$

So how do we make  $\mathbb{X}_!$  into a Cartesian differential category?



## The coKleisli Category of a Differential Category II

For the product structure:

- On objects,  $A \times B$
- Projections:

$$\llbracket \pi_i \rrbracket := !(A_0 \times A_1) \xrightarrow{\varepsilon} A_0 \times A_1 \xrightarrow{\pi_i} A_i$$

For a comonad on a category with finite products, the coKleisli category has finite products.

For the additive structure:

- The sum of maps:  $\llbracket f + g \rrbracket := \llbracket f \rrbracket + \llbracket g \rrbracket$
- Zero maps:  $\llbracket 0 \rrbracket := 0$

For a comonad on an additive category, the coKleisli category is **ONLY** a left additive category, because coKleisli composition does not preserve the additive structure. However, every coKleisli map of the form  $f \circ \varepsilon$  is additive.

For a comonad on an additive category with finite products, the coKleisli category is a Cartesian left additive category.

# The coKleisli Category of a Differential Category III

Recall that last time we defined the differential of  $\llbracket f \rrbracket : !A \rightarrow B$  as:

$$!A \otimes A \xrightarrow{d} !A \xrightarrow{\llbracket f \rrbracket} B$$

But this is not a coKleisli map!

The differential combinator  $\llbracket D[f] \rrbracket : !(A \times A) \rightarrow B$  is defined as follows:

$$!(A \times A) \xrightarrow{\Delta} !(A \times A) \otimes !(A \times A) \xrightarrow{!(\pi_0) \otimes !(\pi_1)} !A \otimes !A \xrightarrow{1 \otimes \varepsilon} !A \otimes A \xrightarrow{d} !A \xrightarrow{\llbracket f \rrbracket} B$$

## Theorem

*For a differential category with finite products, its coKleisli category is a Cartesian differential category.*

Every coKleisli map of the form  $f \circ \varepsilon$  is linear!

(This is an if and only if when one has the Seelye isomorphisms)

## Some examples

### Example

Consider the differential category  $\text{VEC}_k^{op}$  with  $!(V) = \text{Sym}(V)$  from last time. Then  $\text{POLY}_k$  is a sub-CDC of the coKleisli category  $(\text{VEC}_k^{op})_{\text{Sym}}$ .

### Example

Consider the differential category  $\text{VEC}_{\mathbb{R}}^{op}$  with  $!(V) = S^{\infty}(V)$  from last time. Then  $\text{SMOOTH}$  is a sub-CDC of the coKleisli category  $(\text{VEC}_{\mathbb{R}}^{op})_{S^{\infty}}$ .

More explicit examples are described in:



Bucciarelli, A. and Ehrhard, T. and Manzonetto, G. **Categorical models for simply typed resource calculi.**

which include the relational model and the finiteness space model

## The other direction: Cartesian differential storage categories



Blute, R., Cockett, J.R.B. and Seely, R.A., 2015. **Cartesian differential storage categories**.

“... it was not obvious how to pass from Cartesian differential categories back to monoidal differential categories. This paper provides natural conditions under which the linear maps of a Cartesian differential category form a monoidal differential category. ... The purpose of this paper is to make precise the connection between the two types of differential categories. ”

Main idea: While not every Cartesian differential category is the coKleisli category of a differential category, **Cartesian differential storage categories** are precisely the coKleisli categories of differential categories.

### Theorem

*For a Cartesian differential storage category, its category of linear maps form a differential category with finite products and the Seely isomorphisms. Conversely, for a differential category with finite products and the Seely isomorphisms, it's coKleisli category is a Cartesian differential storage category.*

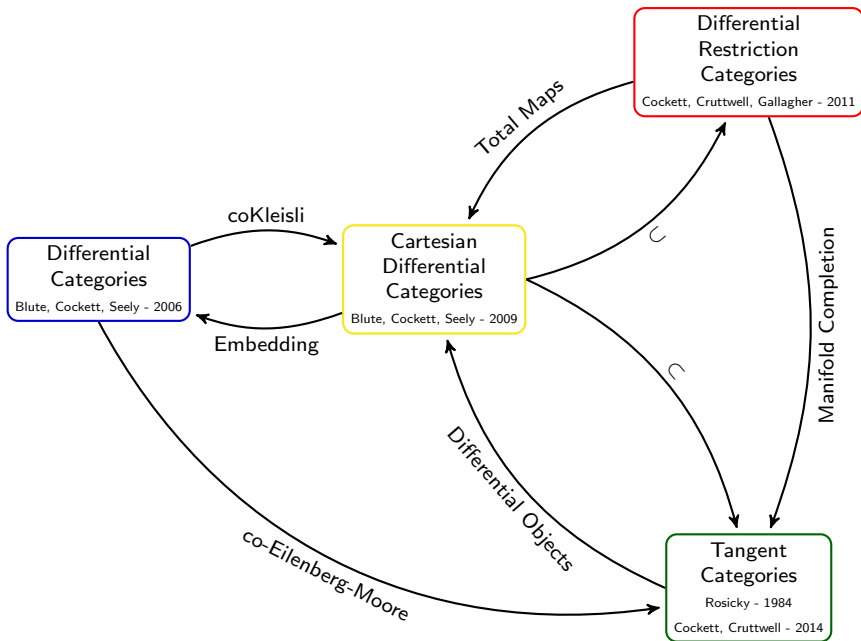


Garner, R, and Lemay, J-S P. **Cartesian differential categories as skew enriched categories.**

### Theorem

*Every (small) Cartesian differential category embeds into the coKleisli category of a differential category.*

# The Differential Category World: It's all connected!



# A quick word on Differential Restriction Categories

A **restriction category** is a category equipped with a restriction operator

$$\frac{f : A \rightarrow B}{\bar{f} : A \rightarrow A}$$

where you should think of  $\bar{f}$  as capturing the domain of definition of  $f$ . Restriction categories allow us to work with partially defined functions.



Lack, S., and Cockett, R. [Restriction Categories \(I - III\)](#).

A **differential restriction category** is **NAIVELY** a Cartesian differential category with a restriction operator such that the differential operator and restriction operator are compatible.



Cockett, R., Cruttwell, G., and Gallagher, J. [Differential Restriction Categories](#).

## Example

- The category of smooth functions defined on open subsets is a differential restriction category.
- Any Cartesian differential category is a differential restriction category where  $\bar{f} = 1$ , so every map is total.
- Conversely, the subcategory of maps such that  $\bar{f} = 1$  in a differential restriction category is a Cartesian differential category.

# The Differential Category World: It's all connected!

Hope you enjoyed it!  
Thanks for listening!  
Merci!

