Cartesian differential categories

J.R.B. Cockett

Department of Computer Science
University of Calgary
Alberta, Canada

robin@cpsc.ucalgary.ca

(work with Robert Seely and Rick Blute)

Ottawa, April/May 2009

Contents:

Getting going

Left-additive categories

Cartesian Differential Categories

Term Logic

Faà di Bruno

Getting going ...

- Differential categories = Seely category + differential operator
- Simple categorical axiomatization
- Easy to recognize (in retrospect)!
- Abstract framework for differentiation: lots of models
- Inspired by Ehrhard's work: Köthe spaces, finiteness spaces and (with Regnier) the differential λ-calculus.
- Many consequences ...
- Useful part of first year calculus ...

Getting going ...

Why is this not enough?

- Classical differential calculus is in the coKleisli category of a differential category ..

  WHAT DO THESE coKLEISLI CATEGORIES LOOK LIKE?

- Differential calculus rapidly moves onto manifolds ..

  HOW DO YOU GET MANIFOLDS FROM COKLEISLI CATEGORIES?

  What about Orbifolds?

- Integration is as central as differentiation

  WHAT IS INTEGRATION ABSTRACTLY?

Getting going ...

WHAT DO THESE coKLEISLI CATEGORIES LOOK LIKE?

... just write down the equations ...

HOW HARD CAN THAT BE?

Getting going ...

Well three years later we still had not got it right!

WHY?

A. We were idiotic?
B. Too much academic baggage ...
C. Too much calculus for the masses ...
D. Analysts stress solving problems ...
E. The structure of the area has been trampled on with:
    - Preconceptions: what does dx mean?
    - Manipulations without algebraic basis ..
    - Didactic short-cuts to “help” students.
F. The axioms are quite tricky!

Getting going ...

The good news:

We have got the *basic* axiomatization right!

FINALLY!

Getting going ...

The bad news:

How do we know?

THE BEAST HAS MANY HEADS!..

Left-additive categories

A category $\mathbb{X}$ is a **left-additive category** in case:

- Each hom-set is a commutative monoid $(0,+)$
- $f(g+h) = (fg) + (fh)$ and $f0 = 0$.

$$A \xrightarrow{f} B \overset{g}{\underset{h}{\rightrightarrows}} C$$

A map h is said to be **additive** if it also preserves the additive structure on the right $(f+g)h = (fh) + (gh)$ and $0h = 0$.

$$A \overset{f}{\underset{g}{\rightrightarrows}} B \xrightarrow{h} C$$

NOTE: additive maps will be the exception ...

Left-additive categories

Lemma

In any left additive category:

- (i) 0 *maps are additive;*
- (ii) *additive maps are closed under addition;*
- (iii) *additive maps are closed under composition;*
- (iv) *identity maps are additive;*
- (v) *if g is a retraction which is additive and the composite gh is additive then h is additive;*
- (vi) *if f is an isomorphism which is additive then f^{-1} is additive.*

Additve maps form a subcategory ...

Left-additive categories

Example

(i) The category whose objects are commutative monoids CMon but whose maps need not preserve the additive structure.

(ii) Real vector spaces with smooth maps.

(iii) The coKleisli category for any comonad on an additive category. (Note: the functor need not be left-additive)

Left-additive categories

A **Cartesian left-additive category** is a left-additive category with products such that:

- the maps π_0, π_1, and Δ are additive;
- whenever f and g are additive then $f \times g$ (pairing preserves additivity).

Lemma

The following are equivalent:

(i) A Cartesian left-additive category;

(ii) A left-additive category for which $\mathbb{X}_+$ has biproducts and the the inclusion $\mathcal{I} : \mathbb{X}_+ \longrightarrow \mathbb{X}$ creates products;

(iii) A Cartesian category $\mathbb{X}$ in which each object is equipped with a chosen commutative monoid structure $(+_A : A \times A \longrightarrow A, 0_A : 1 \longrightarrow A)$ such that $+_{A\times B} = \langle (\pi_0 \times \pi_0)+_A, (\pi_1 \times \pi_1)+_B \rangle$ and $0_{A\times B} = \langle 0_A, 0_B \rangle$.

Left-additive categories

Lemma

In a Cartesian left-additive category:

(i) *f is additive if and only if*

$$(\pi_0 + \pi_1)f = \pi_0 f + \pi_1 f : A \times A \longrightarrow B \quad \textit{and} \quad 0f = 0 : 1 \longrightarrow B;$$

(ii) *$g : A \times X \longrightarrow B$ is additive in its second argument if and only if*

$$1 \times (\pi_0 + \pi_1)g = (1 \times \pi_0)g + (1 \times \pi_1)g : A \times X \times X \longrightarrow B \quad \textit{and} \quad (1 \times 0)g$$

"Multi-additive maps" are maps which are aditive in each argument.

All our earlier example are Cartesian left-additive categories!

Left-additive categories

A functor between Cartesian left-additive categories is **Cartesian left-additive** in case

- $F(f+g) = F(f) + F(g)$ and $F(0) = 0$;
- F preserves products strictly $F(A \times B) = F(A) \times F(B)$.

Lemma

A Cartesian left-additive functor, $F : \mathbb{X} \longrightarrow \mathbb{Y}$, preserves additive maps.and multi-additive maps.

We shall call the category of all cartesian left-additive categories and cartesian left-additive functors CLAdd.

Cartesian Differential Categories

An operator $D_\times$ on the maps of a Cartesian left-additive category

$$\frac{X \xrightarrow{f} Y}{X \times X \xrightarrow[D_\times[f]]{} Y}$$

is a **Cartesian differential operator** in case it satisfies:

[CD.1] $D_\times[f+g] = D_\times[f] + D_\times[g]$ and $D_\times[0] = 0$;

[CD.2] $\langle (h+k), v\rangle D_\times[f] = \langle h, v\rangle D_\times[f] + \langle k, v\rangle D_\times[f]$;

[CD.3] $D_\times[1] = \pi_0$, $D_\times[\pi_0] = \pi_0\pi_0$, and $D_\times[\pi_1] = \pi_0\pi_1$;

[CD.4] $D_\times[\langle f, g\rangle] = \langle D_\times[f], D_\times[g]\rangle$ (and $D_\times[\langle\rangle] = \langle\rangle$);

[CD.5] $D_\times[fg] = \langle D_\times[f], \pi_1 f\rangle D_\times[g]$.

[CD.6] $\langle\langle f, 0\rangle, \langle h, g\rangle\rangle D_\times[D_\times[f]] = \langle f, h\rangle D_\times[f]$;

[CD.7] $\langle\langle 0, f\rangle, \langle g, h\rangle\rangle D_\times[D_\times[f]] = \langle\langle 0, g\rangle, \langle f, h\rangle\rangle D_\times[D_\times[f]]$

A Cartesian left-additive category with a differential operator is a **Cartesian differential category**.

Cartesian Differential Categories

What was so hard about that?

ANSWER: the last two rules!!

They are independent ...
They involve higher differentials ...
... where do they come from in differential categories?

Cartesian Differential Categories

[CD.1] $D_\times[f+g] = D_\times[f] + D_\times[g]$ and $D_\times[0] = 0$;
(operator preserves additive structure)

[CD.2] $\langle (h+k), v\rangle D_\times[f] = \langle h, v\rangle D_\times[f] + \langle k, v\rangle D_\times[f]$
(always additive in first argument);

[CD.3] $D_\times[1] = \pi_0$, $D_\times[\pi_0] = \pi_0\pi_0$, and $D_\times[\pi_1] = \pi_0\pi_1$
(coherence maps are linear -differential constant);

[CD.4] $D_\times[\langle f, g\rangle] = \langle D_\times[f], D_\times[g]\rangle$ (and $D_\times[\langle\rangle] = \langle\rangle$)
(opertor preserves pairing);

[CD.5] $D_\times[fg] = \langle D_\times[f], \pi_1 f\rangle D_\times[g]$ (chain rule);

[CD.6] $\langle\langle f, 0\rangle, \langle h, g\rangle\rangle D_\times[D_\times[f]] = \langle f, h\rangle D_\times[f]$
(differentials are linear in first argument);

[CD.7] $\langle\langle 0, f\rangle, \langle g, h\rangle\rangle D_\times[D_\times[f]] = \langle\langle 0, g\rangle, \langle f, h\rangle\rangle D_\times[D_\times[f]]$
(partial differentials commute);

Cartesian Differential Categories

Real vector spaces with smooth maps are the "standard" example of a Cartesian differential category.

$$\frac{\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} f_1(x_1,..,x_n) \\ \vdots \\ f_m(x_1,..,x_n) \end{pmatrix}}{\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \right) \mapsto \begin{pmatrix} \frac{\mathrm{d}f_1(\tilde{x})}{\mathrm{d}x_1}(x_1) \cdot u_1 + \ldots + \frac{\mathrm{d}f_1(\tilde{x})}{\mathrm{d}x_n}(x_n) \cdot u_n \\ \vdots \\ \frac{\mathrm{d}f_m(\tilde{x})}{\mathrm{d}x_1}(x_1) \cdot u_1 + \ldots + \frac{\mathrm{d}f_m(\tilde{x})}{\mathrm{d}x_n}(x_n) \cdot u_n \end{pmatrix}} \; \mathrm{D}$$

Cartesian Differential Categories

Every simple slice $\mathbb{X}[A]$ of a cartesian differential category, $\mathbb{X}$, is a cartesian differential category.

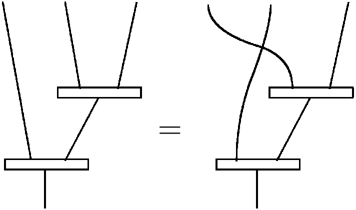
Free Cartesian differential categories on certain graphs ...

Cofree Cartesian differential categories from Faà di Bruno ...

Cartesian Differential Categories

The coKleisli category of a (symmetric) differential category is a cartesian differential category.

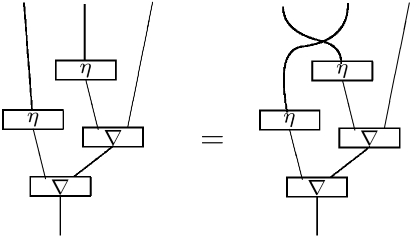
The extra "symmetric" requirement is:



$$1 \otimes d_\otimes; d_\otimes = c; 1 \otimes d_\otimes; d_\otimes$$

Cartesian Differential Categories

This extra symmetric property is always true in a *storage* differential category (has a bialgebra modality and codereliction).



$$1 \otimes d_\otimes ; d_\otimes = c; 1 \otimes d_\otimes ; d_\otimes$$

Cartesian Differential Categories

A map in a Cartesian differential category is said to be **linear** in case $D_\times[f] = \pi_0 f$.

Lemma

(i) *Linear maps are additive:* 0,1, π_0,π_1 *are linear map, and if* f *and* g *are linear then* $f + g$ *is linear;*

(ii) *Linear maps compose, include the identity maps and if* f *and* g *are then* $f + g$ *is linear;*

(iii) $\langle 1, 0\rangle D_\times[f]$ *is linear (uses* **[CD.6]***);*

(iv) *If* a *and* b *are linear then the following inference holds:*

$$\begin{array}{ccc} A & \xrightarrow{f} & B \\ {\scriptstyle a}\downarrow & & \downarrow{\scriptstyle b} \\ A' & \xrightarrow[f']{} & B' \end{array} \quad \Rightarrow \quad \begin{array}{ccc} A \times A & \xrightarrow{D_\times[f]} & B \\ {\scriptstyle a\times a}\downarrow & & \downarrow{\scriptstyle b} \\ A' \times A' & \xrightarrow[D_\times[f']]{} & B' \end{array}$$

(v) *If* f *is an isomorphism and linear then* f^{-1} *is linear.*

Cartesian Differential Categories

STRUCTURAL QUESTION:

When do the linear maps form a differential category?
When is the category the coKleisli category of the linear maps?

Still not completely resolved ...

Term logic

Cartesian differential categories have a term logic ... which makes them much easier (possible) to work in.

Aim to make the term logic “look” like the standard notation for differential calculus ...

Term logic

First the basic structural judgements:

$$\frac{}{\Gamma, x : T \vdash x : T}\ \text{Proj} \qquad \frac{\Gamma \vdash t' : T'}{\Gamma, x : T \vdash t' : T'}\ \text{Weak}$$

$$\frac{\Gamma \vdash t' : T'}{\Gamma, () : 1 \vdash t' : T'}\ \text{Unit} \qquad \frac{\Gamma, x : T_1, y : T_2 \vdash t' : T'}{\Gamma, (x, y) : T_1 \times T_2 \vdash t' : T'}\ \text{Pair}$$

$$\frac{\Gamma \vdash t_1 : T_1 \quad \Gamma \vdash t_2 : T_2}{\Gamma \vdash (t_1, t_2) : T_1 \times T_2}\ \text{Tuple} \qquad \frac{}{\Gamma \vdash () : 1}\ \text{UnitTuple}$$

Term logic

$$\frac{\Gamma \vdash t_1 : T \quad \Gamma \vdash t_2 : T}{\Gamma \vdash t_1 + t_2 : T}\ \text{Add} \qquad \frac{}{\Gamma \vdash 0 : T}\ \text{Zero}$$

$$\frac{\{\Gamma \vdash t_i : T_i\}_{i=1,..,n} \quad f \in \Omega(T_1, ..., T_n; T)}{\Gamma \vdash f(t_1, ..., t_n) : T}\ \text{Fun}$$

$$\frac{\Gamma, x : S \vdash t : T \quad \Gamma \vdash s : S \quad \Gamma \vdash u : S}{\Gamma \vdash \frac{\partial t}{\partial x}(s) \cdot u : T}\ \text{Diff}$$

$$\frac{\Gamma \vdash t_1 : T \quad \Gamma, x : T \vdash t_2 : T'}{\Gamma \vdash t_2[t_1/x] : T'}\ \text{Cut}$$

Note the differential term is a "binding"/"quantification" ...
No infinitesimals ....

Term logic

[Dt.1] $\dfrac{\partial(t_1+t_2)}{\partial p}(s)\cdot u = \dfrac{\partial t_1}{\partial p}(s)\cdot u + \dfrac{\partial t_2}{\partial p}(s)\cdot u$ and
$\dfrac{\partial 0}{\partial p}(s)\cdot u = 0$;

[Dt.2] $\dfrac{\partial t}{\partial p}(s)\cdot(u_1+u_2) = \dfrac{\partial t}{\partial p}(s)\cdot u_1 + \dfrac{\partial t}{\partial p}(s)\cdot u_2$ and
$\dfrac{\partial t}{\partial p}(s)\cdot 0 = 0$;

[Dt.3] $\dfrac{\partial x}{\partial x}(s)\cdot u = u$, $\dfrac{\partial t}{\partial(p,p')}(s,s')\cdot(u,0) = \dfrac{\partial t[s'/p']}{\partial p}(s)\cdot u$ and
$\dfrac{\partial t}{\partial(p,p')}(s,s')\cdot(0,u') = \dfrac{\partial t[s/p]}{\partial p'}(s')\cdot u'$;

[Dt.4] $\dfrac{\partial(t_1,t_2)}{\partial p}(s)\cdot u = \left(\dfrac{\partial t_1}{\partial p}(s)\cdot u, \dfrac{\partial t_1}{\partial p}(s)\cdot u\right)$;

Term logic

[Dt.5] $$\frac{\partial t[t'/p']}{\partial p}(s)\cdot u = \frac{\partial t}{\partial p'}(t'[s/p])\cdot\left(\frac{\partial t'}{\partial p}(s)\cdot u\right)$$
(The chain rule: no variable of p occur in t);

[Dt.5] $$\frac{\partial \frac{\partial t}{\partial p}(s)\cdot p'}{\partial p'}(r)\cdot u = \frac{\partial t}{\partial p}(s)\cdot u.$$

[Dt.5] $$\frac{\partial \frac{\partial t}{\partial p_1}(s_1)\cdot u_1}{\partial p_2}(s_2)\cdot u_2 = \frac{\partial \frac{\partial t}{\partial p_2}(s_2)\cdot u_2}{\partial p_1}(s_1)\cdot u_1$$
(Independence of partial derivatives: s_1, u_1, s_2, u_2 do not contain variables from p_1 or p_2)

The term logic is standard calculus!

Faà di Bruno

Francesco Faà di Bruno (1825-1888) was an Italian of noble birth, a soldier, a mathematician, and a priest. In 1988 he was beatified by Pope John Paul II for his charitable work teaching young women mathematics. As a mathematician he studied with Cauchy in Paris. He was a tall man with a solitary disposition who spoke seldom and, when teaching class, not always successfully. Perhaps his most significant mathematical contribution concerned the combinatorics of the higher-order chain rules. These results where the cornerstone of "combinatorial analysis": a subject which never really took off.

Our interest is in the higher-order chain rule ...

Faà di Bruno

Higher-order derivatives are defined recursively:

$$\frac{\mathrm{d}^{(1)} t}{\mathrm{d}x}(p) \cdot u = \frac{\mathrm{d}t}{\mathrm{d}x}(p) \cdot u$$

$$\frac{\mathrm{d}^{(n)} t}{\mathrm{d}x}(p) \cdot u_1 \cdot \ldots \cdot u_n = \frac{\mathrm{d}\frac{\mathrm{d}^{(n-1)} t}{\mathrm{d}x}(x) \cdot u_1 \cdot \ldots \cdot u_{n-1}}{\mathrm{d}x}(p) \cdot u_n$$

QUESTION:
What do the higher-order chain rule look like?

$$\frac{\mathrm{d}^{(n)} g(f(x))}{\mathrm{d}x}(p) \cdot u_1 \cdot \ldots \cdot u_n = ????$$

The answer involves some combinatorics ...

Faà di Bruno

The second-order chain rule:

$$
\begin{aligned}
&\frac{\mathrm{d}^{(2)} f(g(x))}{\mathrm{d}x}(p) \cdot u_1 \cdot u_2 \\
&= \frac{\mathrm{d}\frac{\mathrm{d}f(g(y))}{\mathrm{d}y}(x) \cdot u_1}{\mathrm{d}x}(p) \cdot u_2 \\
&= \frac{\mathrm{d}\frac{\mathrm{d}f(z)}{\mathrm{d}z}(g(x)) \cdot \left(\frac{\mathrm{d}g(y)}{\mathrm{d}y}(x) \cdot u_1\right)}{\mathrm{d}x}(p) \cdot u_2 \\
&= \frac{\mathrm{d}^{(2)} f(z)}{\mathrm{d}z}(g(p)) \cdot \left(\frac{\mathrm{d}g(y)}{\mathrm{d}y}(p) \cdot u_1\right) \cdot \left(\frac{\mathrm{d}g(x)}{\mathrm{d}x}(p) \cdot u_2\right) \\
&\quad + \frac{\mathrm{d}f(x)}{\mathrm{d}x}(g(p)) \cdot \left(\frac{\mathrm{d}^{(2)} g(x)}{\mathrm{d}x}(p) \cdot u_1 \cdot u_2\right)
\end{aligned}
$$

As n increases the expressions become much more complex!

Faà di Bruno

A **symmetric tree** of depth $n \geq 0$ and in variables $x_1, \dots, x_m$ is:

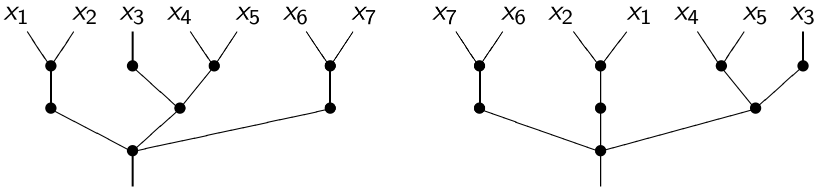
- The only symmetric tree of height 0 has width 1 and is a variable y;
- A symmetric tree of height $n \geq 1$ in the variables $x_1, \dots, x_m$, that is of width m, is an expression $\bullet_r(t_1, \dots, t_r)$ where each t_i is a symmetric tree of height $n-1$ in the variables X_i, where $\bigsqcup_{i=1}^r X_i = X$.

Note that the inductive step involves splitting the variables into r disjoint non-empty subsets. The combinatorics of this is described by Stirling numbers, of the second kind.
The operations at the nodes are viewed as being **symmetric**, or commutative:

$$\bullet_r(t_1, \dots, t_r) = \bullet_r(t_{\sigma(1)}, \dots, t_{\sigma(r)})$$

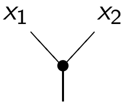
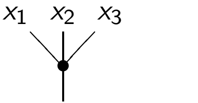
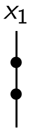
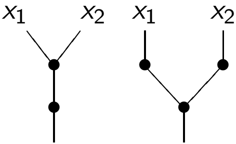
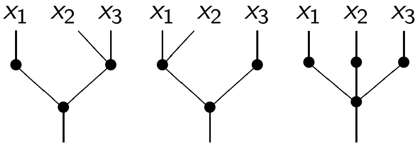
Faà di Bruno

Here are two representations of the same symmetric tree:



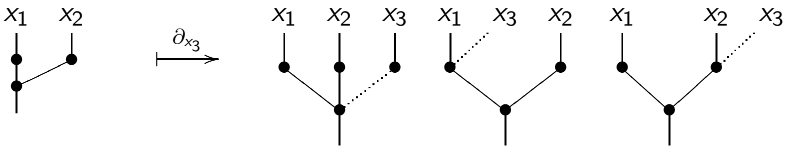
Faà di Bruno

A classification of the first few symmetric trees by height and width:

	width 1	width 2	width 3
hgt 1			
hgt 2			

Faà di Bruno

The differential of a symmetric tree τ of height n and width r produces a bag of m trees of height n and width $r+1$, where m is the number of nodes of τ. The new trees of the differential are produced by picking a node and adding a "limb" to the new variable. The limb consists of a series of unary nodes applied to the new variable: the unary nodes retain the uniform height of the tree.



All symmetric trees of a given height and width can be obtained by differentiating the unique tree of width one of the same height, ι_h.

Faà di Bruno

The Faà di Bruno (bundle) category, Faà($\mathbb{X}$).

Objects: pairs of objects of the original category (A, X) (diagonal case (A, A));

Maps: $f : (A, X) \longrightarrow (B, Y)$ are infinite sequences of *symmetric forms*

$$f = (f_*, f_1, f_2, ...) : (A, X) \longrightarrow (B, X)$$

Where $f_* : X \longrightarrow Y$ is a map in $\mathbb{X}$ and, for $r > 1$,

$$f_r : \underbrace{A \times ... \times A}_{r} \times X \longrightarrow B$$

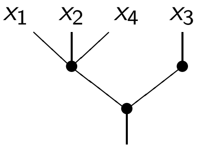
is additive in each of the first r arguments and symmetric in these arguments.

Identities: $(1, \pi_0, ...) : (A, X) \longrightarrow (A, X)$

Composition: Faà di Bruno convolution ...

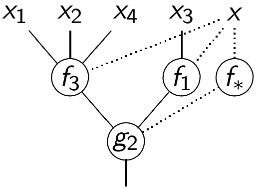
Faà di Bruno

Faà di Bruno convolution ... when τ is the following tree



then

$$(f,g)\star\tau(x) = (((x_1,x_2,x_4,z)f_3,(x_3,z)f_1,f_*(x))g_2 : \underbrace{A\times\ldots\times A}_{4}\times X \longrightarrow C.$$



Notice that $(f,g)\star\tau(x)$ is additive in each argument except the last when f and g have this property.

Faà di Bruno

Faà di Bruno convolution:

$$(fg)_n = \sum_{\tau \in \mathcal{T}_2^n} (f, g) \star \tau$$

where $\mathcal{T}_2^n$ is all symmetric trees of height 2 and width n. This gives an associative composition with unit.

Observations:

- Faà : CLAdd $\longrightarrow$ CLAdd is a functor;
- ε : Faà$(\mathbb{X}) \longrightarrow \mathbb{X}; (f_*, f_1, f_2, ...) \mapsto f_*$ is a fibration and a natural transformation in CLAdd;
- A differential Cartesian category has a section to this fibration: $f \mapsto (f, f^{(1)}, f^{(2)}, ...)$

Faà di Bruno

In fact we are currently filling in the details of:

Theorem
Faà : $\mathsf{CLAdd} \longrightarrow \mathsf{CLAdd}$ *gives a comonad on* CLAdd *which (when restricted to diagonal objects) has coalgebras which are exactly cartesian differential categories.*

More proof of the pudding ...

END