Virtual Special Issue – Women in Topology II: Further collaborations in homotopy theory

Directional derivatives and higher order chain rules for abelian functor calculus

Kristine Bauer [a,*], Brenda Johnson [b], Christina Osborne [c], Emily Riehl [d], Amelia Tebbe [e]

[a] *Department of Mathematics and Statistics, University of Calgary, 2500 University Dr. NW, Calgary, Alberta, T2N 1N4, Canada*
[b] *Department of Mathematics, Union College, Schenectady, NY 12308, United States*
[c] *Department of Mathematics, University of Virginia, 141 Cabell Drive, Charlottesville, VA 22904, United States*
[d] *Department of Mathematics, Johns Hopkins University, 3400 N Charles Street, Baltimore, MD 21218, United States*
[e] *Department of Mathematics, Indiana University Kokomo, 2300 S. Washington Street, Kokomo, IN 46902, United States*

A B S T R A C T

In this paper, we consider *abelian functor calculus*, the calculus of functors of abelian categories established by the second author and McCarthy. We carefully construct a category of abelian categories and suitably homotopically defined functors, and show that this category, equipped with the directional derivative, is a *cartesian differential category* in the sense of Blute, Cockett, and Seely. This provides an abstract framework that makes certain analogies between classical and functor calculus explicit. Inspired by Huang, Marcantognini, and Young's chain rule for higher order directional derivatives of functions, we define a higher order directional derivative for functors of abelian categories. We show that our higher order directional derivative is related to the iterated partial directional derivatives of the second author and McCarthy by a Faà di Bruno style formula. We obtain a higher order chain rule for our directional derivatives using a feature of the cartesian differential category structure, and with this provide a formulation for the nth layers of the Taylor tower of a composition of functors $F \circ G$ in terms of the derivatives and directional derivatives of F and G, reminiscent of similar formulations for functors of spaces or spectra by Arone and Ching. Throughout, we provide explicit chain homotopy equivalences that tighten previously established quasi-isomorphisms for properties of abelian functor calculus.

© 2017 Published by Elsevier B.V.

* Corresponding author.
E-mail addresses: bauerk@ucalgary.ca (K. Bauer), johnsonb@union.edu (B. Johnson), cdo5bv@virginia.edu (C. Osborne), eriehl@math.jhu.edu (E. Riehl), antebbe@iu.edu (A. Tebbe).

1. Introduction

1.1. Chain rules

When studying the calculus of functions, the chain rule for the first derivative of a composition of functions is given by the familiar formula

$$(f \circ g)' = (f' \circ g) \cdot g'.$$

By repeatedly applying this formula, one can derive the classical Faà di Bruno formula for the nth derivative of the composite $f \circ g$:

$$(f \circ g)^{(n)}(x) = \sum \frac{n!}{k_1! k_2! \ldots k_n!} f^{(k)}(g(x)) \left(\frac{g'(x)}{1!} \right)^{k_1} \cdots \left(\frac{g^{(n)}(x)}{n!} \right)^{k_n}$$

where the sum is taken over all non-negative integer solutions to

$$k_1 + 2k_2 + \ldots + nk_n = n$$

and $k = k_1 + \ldots + k_n$. An alternative higher order chain rule can be obtained by using the *directional* derivative of f at x along v:

$$\nabla f(v; x) = \lim_{t \to 0} \frac{1}{t} \left[f(x + tv) - f(x) \right].$$

Huang, Marcantognini, and Young [10] define *higher order directional derivatives* $\Delta_n f$, functions of $(n + 1)$ variables with $\Delta_1 f = \nabla f$, with which they obtain a concise formulation for the nth derivative of a composition

$$(f \circ g)^{(n)}(x) = \Delta_n f(g^{(n)}(x), \ldots, g'(x); g(x)), \tag{1.1}$$

when g is a function of a single variable, or

$$\Delta_n(f \circ g) = \Delta_n f(\Delta_n g, \ldots, \Delta_1 g; g), \tag{1.2}$$

in the general case.

For homotopy functors (such as functors from spectra to spectra or spaces to spaces that preserve weak equivalences), Goodwillie's calculus of functors builds a *Taylor tower* of "polynomial" functors and natural transformations:

$$F \to \cdots \to P_n F \to P_{n-1} F \to \cdots \to P_0 F$$

where each $P_n F$ is an n-excisive functor that approximates F in a range of homotopy groups depending on the functor F and the connectivity of the input [8]. The homotopy fiber of the map $P_n F \to P_{n-1} F$ is the nth *layer* of the tower, $D_n F$. For homotopy functors from spectra to spectra, the layers are of the form $D_n F \simeq \partial_n F \wedge_{h\Sigma_n} X^{\wedge n}$, where $\partial_n F$ is a spectrum with an action of the nth symmetric group, Σ_n. This spectrum is called the nth derivative of F.

Several versions of chain rules for the derivatives of functor calculus have been developed. Notably, Arone and Ching [1] derived a chain rule for the derivatives $\partial_n F$ using the fact that for functors of spaces or spectra, the symmetric sequence $\{\partial_n F\}$ is a module over the operad formed by the derivatives of the

identity functor of spaces. This generalized earlier work of Klein and Rognes [16] that established a chain rule for first derivatives. For functors of spectra, a chain rule for the derivatives is given by Ching [5]. For functors of spaces, Yeakel has developed an alternative method for deriving these chain rules that does not require passage to spectra [22].

1.2. Chain rules in abelian functor calculus

The chain rule results of the preceding paragraph use Goodwillie's original formulation of the Taylor tower. The second author and McCarthy [12] defined Taylor towers for not necessarily additive functors between abelian categories using the classical cross effect functors of Eilenberg and Mac Lane [7]. This approach has been generalized and applies to a wide variety of contexts [2], [3]. In [12], Johnson and McCarthy defined a notion of directional derivative $\nabla F(V; X)$ for a functor F valued in an abelian category and proved the analog of Huang, Marcantognini, and Young's chain rule in degree one [12, Proposition 5.6]:

$$\nabla(F \circ G) \simeq \nabla F(\nabla G; G),$$

under the hypothesis that the functor G preserves the zero object.

The present work was motivated by the goal of proving an analog of the higher order directional derivative chain rule of Huang, Marcantognini, and Young for the abelian functor calculus of Johnson and McCarthy. Achieving this goal required first dealing with the question of how to define a higher order directional derivative in this context. To get a sense of the choices involved, consider the functor $\nabla F(V; X)$. This is a functor of two variables, so the first choice to be made in defining a second order directional derivative is whether one should differentiate with respect to one of the variables (a partial derivative) or both variables simultaneously (a total derivative).

For the first possibility, one notes that $\nabla F(V; X)$ is already a linear functor with respect to its first variable, so that differentiating with respect to that variable yields nothing new. So, for the partial derivative approach, we can restrict our attention to taking the directional derivative of $\nabla F(V; X)$ with respect to X to obtain a second order partial directional derivative $\nabla^2 F$ that is a functor of 3 variables. Iterating this process yields functors $\nabla^n F$ of $(n + 1)$-variables defined in [12], which we refer to as *iterated partial directional derivatives.*

If we follow the total derivative approach by treating ∇F as a functor whose domain is a product category $\mathcal{B} \times \mathcal{B}$, then we obtain as a second order total directional derivative a functor $\nabla^{\times 2} F = \nabla(\nabla F)$ whose source category is $\mathcal{B}^4$. But, this leads to some redundancy, as we end up differentiating in some directions repeatedly. The "correct" approach, at least if one wants to define the analog of the Huang–Marcantognini–Young higher order directional derivatives, lies somewhere in between.

To define our second order directional derivative $\Delta_2 F$ of F, we restrict $\nabla^{\times 2} F$ along a diagonal functor $L_2 \colon \mathcal{B}^3 \to \mathcal{B}^4$ that takes the triple of objects (V_2, V_1, X) to the ordered pair of ordered pairs $((V_2, V_1), (V_1, X))$. With this definition we are able to prove that for a pair of composable functors F and G

$$\Delta_2(F \circ G)(V_2, V_1; X) \simeq \Delta_2 F(\Delta_2 G(V_2, V_1; X), \Delta_1 G(V_1; X); G(X)).$$

Inductively, we define the *nth order directional derivative* $\Delta_n F$ and prove the desired analog of the Huang–Marcantognini–Young chain rule equation (1.2):

Theorem 8.1. *For a composable pair of functors of abelian categories, there is a chain homotopy equivalence between the nth directional derivatives*

$$\Delta_n(F \circ G)(V_n, \ldots, V_1; X) \simeq \Delta_n F(\Delta_n G(V_n, \ldots, V_1; X), \ldots, \Delta_1 G(V_1; X); G(X)).$$

Of course, this "in-between" definition of the higher order directional derivatives raises a natural question – how is it related to the more obvious choices for defining higher order directional derivatives? In the case of $\nabla^{\times n} F$, the functor $\Delta_n F$ is obtained by restricting along a diagonal functor L_n as in the degree 2 case. The case of $\nabla^n F$ proves more interesting. We show that $\Delta_n F$ and $\nabla^n F$ satisfy a Faà di Bruno-type relationship:

Theorem 7.7. *For a functor F between abelian categories, there is a chain homotopy equivalence*

$$\Delta_n F(V_n, \ldots, V_1; X) \simeq \bigoplus \frac{n!}{k_1! k_2! \cdots k_n!} \left(\frac{1}{1!}\right)^{k_1} \cdots \left(\frac{1}{n!}\right)^{k_n} \nabla^{k_1 + \cdots + k_n} F(V_n^{k_n}, \ldots, V_1^{k_1}; X)$$

where the sum is over non-negative integer solutions to the equation $k_1 + 2k_2 + \cdots + nk_n = n$.

A more concise formulation of this result, given in §7, makes use of the fact that the Faà di Bruno coefficients count the number of partitions of a set of n elements. Using these results, we define ordinary (non-directional) derivatives for functors of modules over a commutative ring R and prove a chain rule for these derivatives which is analogous to the Huang–Marcantognini–Young rule from equation (1.1):

Corollary 9.4. *For a composable pair of functors of abelian categories, F and G, where the source category of G is a category of modules over a commutative ring R, the nth derivative*

$$\frac{d^n}{dR^n}(F \circ G)(X)$$

is chain homotopy equivalent to

$$\bigoplus \frac{n!}{k_1! k_2! \cdots k_n!} \left(\frac{1}{1!}\right)^{k_1} \cdots \left(\frac{1}{n!}\right)^{k_n} \nabla^{k_1 + \cdots + k_n} F\left(\frac{d^{k_n}}{dR^{k_n}} G(X), \ldots, \frac{d^{k_1}}{dR^{k_1}} G(X); G(X)\right)$$

where the sum is over non-negative integer solutions to the equation $k_1 + 2k_2 + \cdots + nk_n = n$.

1.3. A categorical context for abelian functor calculus

Although our original motivation for this project was to prove higher order chain rules, we found that a significant part of the challenge in doing so involved placing these chain rules and the abelian functor calculus itself in an appropriate categorical context. This was developed in two stages, using the concepts of Kleisli categories and cartesian differential categories.

For a functor $F: \mathcal{B} \to \mathcal{A}$ between two abelian categories, the abelian functor calculus constructs a degree n approximation $P_n F: \mathcal{B} \to \mathsf{Ch}\mathcal{A}$, where $\mathsf{Ch}\mathcal{A}$ is the category of chain complexes in $\mathcal{A}$ concentrated in non-negative degrees. For a composable pair of functors $F: \mathcal{B} \to \mathcal{A}$ and $G: \mathcal{C} \to \mathcal{B}$, this results in a pair of functors $P_n F: \mathcal{B} \to \mathsf{Ch}\mathcal{A}$ and $P_n G: \mathcal{C} \to \mathsf{Ch}\mathcal{B}$ which no longer appear to be composable. In [12], the authors form the composition $P_n F \circ P_n G$ by "prolonging" $P_n F$ to a functor $\mathsf{Ch}\mathcal{B} \to \mathsf{Ch}(\mathsf{Ch}\mathcal{A})$, composing the prolongation with $P_n G$ and taking the total complex of the resulting double complex. We observe in section 3 that this process is precisely the composition in the Kleisli category associated to a particular pseudomonad Ch acting on the 2-category of abelian categories, and use $\mathsf{AbCat}_{\mathsf{Ch}}$ to denote that Kleisli category.

In functor calculus, the use of terms such as "calculus" and "derivatives" has been justified by pointing out strong formal resemblances to the classical constructions from analysis and undergraduate calculus. Although these analogies are compelling, one wonders if there is a deeper justification. In the course of investigating the higher order chain rule for directional derivatives, we discovered that the properties of the directional derivative for abelian functor calculus correspond exactly to the axioms defining the cartesian

differential categories of Blute, Cockett, and Seely [4]. This concept captures the notion of differentiation in a wide variety of contexts (the category of real vector spaces and smooth maps equipped with the usual differential operator is a standard example of a cartesian differential category) and provides an explanation for some of the formal resemblances.

We prove:

Corollary 6.6. *The homotopy category* $\mathsf{HoAbCat_{Ch}}$ *is a cartesian differential category.*

In this statement, $\mathsf{HoAbCat_{Ch}}$ is the homotopy category of $\mathsf{AbCat_{Ch}}$ obtained by inverting pointwise chain homotopy equivalences. The directional derivative ∇ gives this category its cartesian differential category structure. Fitting calculus into this context led to additional insight stemming from the fact that a cartesian differential category is also an example of a tangent category, as defined by Cockett and Cruttwell [6]. Tangent categories are characterized by the existence of an endofunctor T satisfying essential properties of tangent bundles for manifolds. Our first proof of the chain rule of Theorem 8.1 did not make use of the tangent category structure on $\mathsf{HoAbCat_{Ch}}$, but the proof presented here, which does use the endofunctor T, provides a conceptual simplification. As the first-order chain rule can be seen to encode the functoriality of the derivative, the nth order chain rule can be extracted from the functoriality of the iterated endofunctor T^n, which is an easily derived consequence of our Corollary 6.6.

Proving that $\mathsf{HoAbCat_{Ch}}$ is a cartesian differential category required reworking many of the constructions and results for abelian functor calculus established in [12]. Driving these challenges was the fact that chain rules in the abelian context do not come in the form of isomorphisms. For example, applying the definition of ∇ (see Definition 6.1 or 6.2) to $F \circ G$ produces a chain complex that is not isomorphic to $\nabla F(\nabla G, G)$ in general. Instead, when G is reduced ($G(0) \cong 0$), [12] prove that there is a quasi-isomorphism between $\nabla(F \circ G)$ and $\nabla F(\nabla G, G)$.

The axioms of a cartesian differential category require that $\nabla(F \circ G)$ and $\nabla F(\nabla G, G)$ be isomorphic. This suggests that we attempt to realize $\mathsf{AbCat_{Ch}}$ as a cartesian differential category by inverting quasi-isomorphisms. But because the functors we consider between abelian categories are far from being exact, when we invert quasi-isomorphisms in $\mathsf{AbCat_{Ch}}$, the composition of morphisms is no longer well-defined. However, the somewhat delicately defined composition in our Kleisli category preserves chain homotopy equivalences, so we define $\mathsf{HoAbCat_{Ch}}$ to be the category that inverts these instead. Because of this we needed to upgrade many definitions and results of [12] from quasi-isomorphisms to chain homotopy equivalences.

Revising results of [12] to work up to chain homotopy equivalence rather than quasi-isomorphism was a relatively straightforward process, with the notable exception of the chain rule itself. That required proving the following proposition, which refines Lemma 5.7 of [12].

Proposition 5.7. *For any composable pair of functors F and G with G reduced, there is a chain homotopy equivalence*

$$D_1(F \circ G) \simeq D_1 F \circ D_1 G.$$

To prove this, we prove a technical general result that might be of independent interest (or might already be in the literature):

Theorem 7.7. *Let $\iota \colon A_{\bullet,\bullet} \to B_{\bullet,\bullet}$ be a morphism of first-quadrant bicomplexes that admits a row-wise strong deformation retraction. Then ι induces a chain homotopy equivalence of total complexes* $\mathrm{Tot}(A_{\bullet,\bullet}) \to \mathrm{Tot}(B_{\bullet,\bullet})$.

We include a proof of this result in Appendix A. The proof defines an explicit chain homotopy on the total complex that does not come from a chain homotopy of bicomplexes, as defined in [21, 5.7.3], because the row-wise retraction and chain homotopies are not assumed to be natural with respect to the vertical differentials. The precise conditions are stated in Definition A.1. As a corollary, it follows that:

Corollary A.7. *Let $A_{\bullet,\bullet}$ be a first-quadrant bicomplex so that every row except the zeroth row $A_{0,\bullet}$ is contractible. Then the natural inclusion $A_{0,\bullet} \hookrightarrow \mathrm{Tot}(A)_{\bullet}$ is a chain homotopy equivalence.*

The proof of Proposition 5.7 appears in Appendix B. Combining Proposition 5.7 with some other results, we prove

Theorem 6.5. *For any composable pair of functors F and G, there is a chain homotopy equivalence*

$$\nabla(F \circ G)(V; X) \simeq \nabla F(\nabla G(V; X); G(X)).$$

This improves on Proposition 5.6 of [12] in two ways – by replacing the quasi-isomorphism with a chain homotopy equivalence, and by removing the condition that G be a reduced functor.

1.4. Organization of the paper

In section 2, we define and review properties of the cross effect functors. These are the building blocks for the Taylor towers of abelian functor calculus, including the linearization and directional derivative functors. Section 3 is used to define the Kleisli and homotopy categories in which the constructions and main results of the paper take place. In section 4, we show how the Taylor towers of abelian functor calculus are built and establish their fundamental properties. This treatment differs from that of [12] in two important ways. As discussed in section 1.3, we need the fundamental properties of the tower to hold up to chain homotopy equivalence instead of quasi-isomorphism. In addition, we have streamlined the means by which terms in the abelian functor calculus tower are defined. We do so by identifying a comonad C_n on the category of all functors between a fixed pair of abelian categories, and defining P_nF directly as a resolution of F by this comonad. In [12], this approach was only used for the category of reduced functors. An extra step involving a mapping cone was used to extend this construction to all functors.

In section 5, we start working out the main ingredients for our chain rules – we define the linearization functor D_1 and derive its essential properties. We use these to define the directional derivative for abelian functor calculus in section 6, and prove that it gives $\mathsf{HoAbCat}_{\mathsf{Ch}}$ the structure of a cartesian differential category. In section 7, we define the iterated partial directional derivatives and the higher order directional derivatives and prove that they are related by the Faà di Bruno formula of Theorem 7.7. We prove the analog of the higher order chain rule of [10] in section 8, using the tangent category structure of $\mathsf{HoAbCat}_{\mathsf{Ch}}$. Finally, in section 9, we discuss how the directional derivatives are related to a notion of ordinary derivatives for abelian functor calculus and use the main result of section 7 to prove the Faà di Bruno-style formula for these derivatives in Corollary 9.4.

Conventions   This paper introduces a number of constructions on functors (e.g., the higher order directional derivatives) and investigates their behavior with respect to composition, product constructions, and so forth. A convenient context to describe these operations involves a large category whose objects are abelian categories and whose morphisms are functors between them; the precise construction of this category, which is somewhat delicate, is given in §3. There likely exist more than a set's worth of functors between any fixed pair of non-small abelian categories, so if our constructions on functors are to be interpreted globally, they need to take place in some extension of the usual ZFC axioms for set theory: e.g., by assuming there exists a hierarchy of inaccessible cardinals. See [20] for a friendly discussion of the myriad possible choices.

Tslil Clingman suggested that chain complexes might define a monad on abelian categories and Robin Cockett pointed out that non-reduced functors may fail to preserve chain complexes. This is why the monad of §3 is pseudo and not strict. Geoffrey Cruttwell explained the benefits of thinking of a cartesian differential category as a tangent category, a perspective which led to a simplified proof of Theorem 8.1. We thank Randy McCarthy for helpful conversations in starting this project, and for the ideas that inspired it.

2. Cross effects for functors

Classically, the term *cross effect* was used to describe the combined effects of two or more forces, or to describe the difference between the quantities $f(x + y)$ and $f(x) + f(y)$. For *reduced* functions satisfying $f(0) = 0$, the condition that $f(x+y) = f(x) + f(y)$ for all x and y is equivalent to linearity. Thus, the cross effect function, defined by

$$\mathrm{cr}_2 f(x, y) = f(x + y) - f(x) - f(y),$$

measures the failure of a reduced function f to be linear; this failure is called the *deviation* by Eilenberg and Mac Lane [7]. In this section, we will study an analogous notion for functors from a pointed category to an abelian category. The cross effects were first extended to functors of additive categories by Eilenberg and Mac Lane [7], and these became the fundamental building blocks of abelian functor calculus as developed by Johnson and McCarthy [12]. We recall the definition of the cross effect and summarize the properties that we will need for cross effect functors.

For the duration of this section, let $\mathcal{B}$ be a category with a *basepoint*, i.e., an initial object $\star$ which is also terminal, and finite coproducts, denoted by $\vee$. Let $\mathcal{A}$ be an abelian category with zero object 0 and biproducts denoted $\oplus$.

Definition 2.1 *([7])*. The *nth cross effect* of a functor $F \colon \mathcal{B} \to \mathcal{A}$ is the n-variable functor $\mathrm{cr}_n F \colon \mathcal{B}^n \to \mathcal{A}$ defined recursively by

$$F(X) \cong F(\star) \oplus \mathrm{cr}_1 F(X)$$
$$\mathrm{cr}_1 F(X_1 \vee X_2) \cong \mathrm{cr}_1 F(X_1) \oplus \mathrm{cr}_1 F(X_2) \oplus \mathrm{cr}_2 F(X_1, X_2)$$

and in general,

$$\mathrm{cr}_{n-1} F(X_1 \vee X_2, X_3, \ldots, X_n) \cong \mathrm{cr}_{n-1} F(X_1, X_3, \ldots, X_n) \oplus \mathrm{cr}_{n-1} F(X_2, X_3, \ldots, X_n)$$
$$\oplus \mathrm{cr}_n F(X_1, X_2, \ldots, X_n).$$

Despite the asymmetry in this definition, the nth cross effect is symmetric in its n variables [12, Proposition 1.2].

Remark 2.2. In a category with a basepoint and finite coproducts, each coproduct inclusion $X \hookrightarrow X \vee Y$ is a split monomorphism; the relevance of the retraction is that split monomorphisms are preserved by any

functor. In an abelian category, split monomorphisms extend to split short exact sequences. In particular, it makes sense to define the cross effects as the direct sum complements of Definition 2.1. Eilenberg and Mac Lane defined cross effects as the images of certain homomorphisms, but the direct sum complement definition is more useful for the types of computations we wish to do.

Example 2.3. Fix an object A in an abelian category $\mathcal{A}$ and consider the functor $F \colon \mathcal{A} \to \mathcal{A}$ defined by $F(X) := A \oplus X$. Since $F(0) \cong A$, it follows that $\mathrm{cr}_1 F \cong \mathrm{id}$. Now

$$X \oplus Y \oplus \mathrm{cr}_2 F(X,Y) \cong \mathrm{cr}_1 F(X) \oplus \mathrm{cr}_1 F(Y) \oplus \mathrm{cr}_2 F(X \oplus Y) \cong \mathrm{cr}_1 F(X \oplus Y) \cong X \oplus Y$$

implies that $\mathrm{cr}_2 F(X,Y) \cong 0$, so all higher cross effects must also vanish.

Definition 2.1 is *functorial* in the sense that a natural transformation $F \to G$ induces a naturally-defined map $\mathrm{cr}_n F \to \mathrm{cr}_n G$. To precisely describe this functoriality, write:

- $\mathrm{Fun}(\mathcal{B}, \mathcal{A})$ for the category of all functors from $\mathcal{B}$ to $\mathcal{A}$;
- $\mathrm{Fun}_*(\mathcal{B}, \mathcal{A})$ for the category of all *strictly reduced* functors from $\mathcal{B}$ to $\mathcal{A}$, that is, functors for which $F(\star) \cong 0$; and
- $\mathrm{Fun}_*(\mathcal{B}^n, \mathcal{A})$ for the category of all *strictly multi-reduced* functors from $\mathcal{B}^n$ to $\mathcal{A}$, that is, functors for which $F(X_1, \ldots, X_n) \cong 0$ if any $X_i = \star$.

An easy exercise [12, Proposition 1.2] shows that $\mathrm{cr}_n F$ is an object in $\mathrm{Fun}_*(\mathcal{B}^n, \mathcal{A})$.

Lemma 2.4. *For any $n \geq 1$ and any $F \colon \mathcal{B} \to \mathcal{A}$, $\mathrm{cr}_n F \cong \mathrm{cr}_n \mathrm{cr}_1 F$.*

Proof. The isomorphism $F(\star) \cong F(\star) \oplus \mathrm{cr}_1 F(\star)$ implies that $\mathrm{cr}_1 F(\star) \cong 0$. Now

$$\mathrm{cr}_1 F(X) \cong \mathrm{cr}_1 F(\star) \oplus \mathrm{cr}_1(\mathrm{cr}_1 F(X)) \cong \mathrm{cr}_1^2 F(X)$$

proves that $\mathrm{cr}_1 F \cong \mathrm{cr}_1^2 F$. The general case now follows immediately from Definition 2.1 by induction. $\quad \square$

Importantly, the cross effect functors admit adjoints:

Proposition 2.5. *The inclusion of the category of strictly reduced functors into the category of not-necessarily reduced functors admits a right adjoint, namely the first cross effect functor:*

$$\mathrm{Fun}_*(\mathcal{B}, \mathcal{A}) \underset{\mathrm{cr}_1}{\overset{\perp}{\rightleftarrows}} \mathrm{Fun}(\mathcal{B}, \mathcal{A}).$$

The component of the counit of this adjunction at a functor $G \in \mathrm{Fun}(\mathcal{B}, \mathcal{A})$ is the natural inclusion $\mathrm{cr}_1 G \hookrightarrow G$.

Proof. The universal property of the adjunctions asserts that if F is reduced and G is not necessarily reduced then any natural map $\alpha \colon F \to G$ factors uniquely through the inclusion $\mathrm{cr}_1 G \hookrightarrow G$. Let ! denote the natural transformation to the basepoint object in $\mathcal{B}$, $! \colon \mathrm{id} \to \star$, and consider the diagram:

$$
\begin{array}{ccccc}
\mathrm{cr}_1 FX := \ker & \overset{\cong}{\rightarrowtail} & FX & \overset{F!}{\longrightarrow} & F\star \\
{\scriptstyle \exists !}\big\downarrow & & {\scriptstyle \alpha_X}\big\downarrow & & {\scriptstyle \alpha_\star}\big\downarrow \\
\mathrm{cr}_1 GX := \ker & \rightarrowtail & GX & \underset{G!}{\longrightarrow} & G\star
\end{array}
$$

The right-hand square commutes by naturality of α. The kernels of the right-hand horizontal maps define the first cross effects. By commutativity of the right-hand square, there exists a unique map $\mathrm{cr}_1 FX \to \mathrm{cr}_1 GX$ so that the left-hand square commutes. But because $F\star = 0$, the top left-hand map is an isomorphism, which gives us the desired unique factorization. $\square$

Remark 2.6. A full subcategory is *coreflective* when the inclusion admits a right adjoint. Proposition 2.5 asserts that $\mathrm{Fun}_*(\mathcal{B}, \mathcal{A})$ defines a coreflective subcategory of $\mathrm{Fun}(\mathcal{B}, \mathcal{A})$. Because the left adjoint is full and faithful the unit is necessarily an isomorphism. By abstract nonsense, the inclusion of a coreflective subcategory is necessarily *comonadic*, that is, $\mathrm{Fun}_*(\mathcal{B}, \mathcal{A})$ is the category of coalgebras for the idempotent comonad cr_1 acting on $\mathrm{Fun}(\mathcal{B}, \mathcal{A})$. This formalism gives a characterization of the reduced functors: namely F is reduced if and only if the natural map $\mathrm{cr}_1 F \hookrightarrow F$ is an isomorphism.

Johnson and McCarthy observe that when the construction of the nth cross effect is restricted to reduced functors, it is right adjoint to pre-composition with the diagonal functor [12, Example 1.8]. Proposition 2.5 allows us to extend this adjunction to non-reduced functors:

Corollary 2.7. *There is an adjunction*

$$\mathrm{Fun}(\mathcal{B}, \mathcal{A}) \xleftarrow[\mathrm{cr}_n]{\overset{\Delta^*}{\underset{\perp}{\longleftarrow}}} \mathrm{Fun}_*(\mathcal{B}^n, \mathcal{A})$$

between the nth cross effect functor and the functor given by precomposing with the diagonal $\Delta \colon \mathcal{B} \to \mathcal{B}^n$, inducing a comonad C_n on $\mathrm{Fun}(\mathcal{B}, \mathcal{A})$ defined by

$$C_n F(X) := \mathrm{cr}_n F(X, \dots, X).$$

Proof. By Lemma 2.4, $\mathrm{cr}_n \cong \mathrm{cr}_n \mathrm{cr}_1$ for $n \geq 1$. Thus, the adjunction $\Delta^* \dashv \mathrm{cr}_n$ is the composite of the adjunctions

$$\mathrm{Fun}(\mathcal{B}, \mathcal{A}) \xrightarrow[\mathrm{cr}_1]{\underset{\perp}{\longleftarrow}} \mathrm{Fun}_*(\mathcal{B}, \mathcal{A}) \xleftarrow[\mathrm{cr}_n]{\overset{\Delta^*}{\underset{\perp}{\longleftarrow}}} \mathrm{Fun}_*(\mathcal{B}^n, \mathcal{A})$$

of Proposition 2.5 with the adjunction of [12, Example 1.8]. $\square$

Remark 2.8. Note that by [12, Example 1.8], the counit of the adjunction $\Delta^* \dashv \mathrm{cr}_n$ between $\mathrm{Fun}_*(\mathcal{B}, \mathcal{A})$ and $\mathrm{Fun}_*(\mathcal{B}^n, \mathcal{A})$ is

$$\mathrm{cr}_n G(X, \dots, X) \rightarrowtail G(\vee_{i=1}^n X) \xrightarrow{G(+)} G(X)$$

where $+ \colon \vee_n M \to M$ is the "fold" map from the coproduct. Putting this together with Proposition 2.5, we see that the counit of the adjunction of Corollary 2.7 for functors which are not necessarily reduced is the map ϵ defined to be the composite:

$$\mathrm{cr}_n \mathrm{cr}_1 G(X, \dots, X) \rightarrowtail \mathrm{cr}_1 G(\vee_{i=1}^n X) \xrightarrow{G(+)} \mathrm{cr}_1 G(X) \rightarrowtail G(X).$$

As observed in [12], adjoint functors between abelian categories can be used to construct contractible chain complexes; for instance:

Lemma 2.9. *Let* $\mathcal{A} \underset{R}{\overset{L}{\rightleftarrows}} \mathcal{B}$ *define an adjunction between abelian categories inducing a comonad* $C = LR$ *on* $\mathcal{A}$ *with counit* $\epsilon\colon LR \Rightarrow \mathrm{id}$. *Then for each* $A \in \mathcal{A}$ *the chain complex in* $\mathcal{B}$ *with differentials defined to be the alternating sums* $\sum_{i \geq 0}^{k} (-1)^i R(LR)^{\times i}\epsilon$

$$\cdots R(LR)^{\times 3}(A) \xrightarrow{R\epsilon - RLR\epsilon + R(LR)^{\times 2}\epsilon} R(LR)^{\times 2}A \xrightarrow{R\epsilon - RLR\epsilon} RLRA \xrightarrow{R\epsilon} RA$$

with splittings s_2, s_1, s_0

admits a contracting homotopy, and these splittings are natural in $\mathcal{A}$.
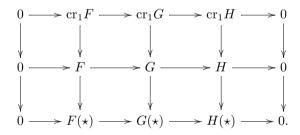
This contracting homotopy comes from a contracting simplicial homotopy: the adjoint pair $L \dashv R$ define an augmented cosimplicial object in $\mathrm{Fun}(\mathcal{A}, \mathcal{B})$ admitting a splitting.

Proof. Define $s_k = \eta R(LR)^{\times k}$ using the unit $\eta\colon \mathrm{id} \Rightarrow RL$ of the adjunction $L \dashv R$. □

The functor categories $\mathrm{Fun}(\mathcal{B}, \mathcal{A})$ and $\mathrm{Fun}_*(\mathcal{B}^n, \mathcal{A})$ are abelian, with kernels and direct sums and so forth defined objectwise in $\mathcal{A}$. In particular, these categories have short exact sequences. A functor from one abelian category to another which preserves exact sequences is called *exact*.

Proposition 2.10. *For each* $n \geq 1$, *the functors* $\mathrm{cr}_n\colon \mathrm{Fun}(\mathcal{B}, \mathcal{A}) \to \mathrm{Fun}_*(\mathcal{B}^n, \mathcal{A})$ *and* $C_n\colon \mathrm{Fun}(\mathcal{B}, \mathcal{A}) \to \mathrm{Fun}(\mathcal{B}, \mathcal{A})$ *are exact.*

Proof. Let $0 \to F \to G \to H \to 0$ be a short exact sequence of functors. Consider the diagram below

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & \mathrm{cr}_1 F & \longrightarrow & \mathrm{cr}_1 G & \longrightarrow & \mathrm{cr}_1 H & \longrightarrow & 0 \\
& & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & F & \longrightarrow & G & \longrightarrow & H & \longrightarrow & 0 \\
& & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
0 & \longrightarrow & F(\star) & \longrightarrow & G(\star) & \longrightarrow & H(\star) & \longrightarrow & 0.
\end{array}
$$

Since the columns are exact, and the bottom two rows are exact, the 3×3 lemma [21, 1.3.2] guarantees that the top row is exact. Then assuming that cr_n is exact, we can apply the same lemma to the diagram

$$
\begin{array}{ccc}
\mathrm{cr}_{n+1}F(A_1, A_2, B_2, \ldots, B_n) & \twoheadrightarrow \mathrm{cr}_{n+1}G(A_1, A_2, B_2, \ldots, B_n) \twoheadrightarrow & \mathrm{cr}_{n+1}H(A_1, A_2, B_2, \ldots, B_n) \\
\downarrow & \downarrow & \downarrow \\
\mathrm{cr}_n F(A_1 \vee A_2, B_2, \ldots, B_n) & \twoheadrightarrow \mathrm{cr}_n G(A_1 \vee A_2, B_2, \ldots, B_n) \twoheadrightarrow & \mathrm{cr}_n H(A_1 \vee A_2, B_2, \ldots, B_n) \\
\downarrow & \downarrow & \downarrow \\
\mathrm{cr}_n F(A_1, B_2, \ldots, B_n) & \mathrm{cr}_n G(A_1, B_2, \ldots, B_n) & \mathrm{cr}_n H(A_1, B_2, \ldots, B_n) \\
\oplus \mathrm{cr}_n F(A_2, B_2, \ldots, B_n) & \longrightarrow \oplus \mathrm{cr}_n G(A_2, B_2, \ldots, B_n) \longrightarrow & \oplus \mathrm{cr}_n H(A_2, B_2, \ldots, B_n)
\end{array}
$$

to conclude that cr_{n+1} is exact. The proof follows by induction. The result for C_n follows immediately. □

3. A categorical context for abelian functor calculus

Abelian functor calculus — also called *additive* or *discrete functor calculus* [2] — considers arbitrary functors valued in an abelian category. The linear approximation defined in [12] satisfies a universal property "up to homotopy." For this to make sense, the target abelian category must be replaced by some sort of homotopical category in which strict universal properties (asserting that certain diagrams commute on the nose) can be replaced by weak ones (where the commutativity is up to some sort of homotopy relation). For an abelian category $\mathcal{A}$, let $\mathsf{Ch}\mathcal{A}$ denote the category of chain complexes on $\mathcal{A}$ concentrated in non-negative degrees. In general, the linear approximation of $F\colon \mathcal{B} \to \mathcal{A}$ defines a functor $D_1F\colon \mathcal{B} \to \mathsf{Ch}\mathcal{A}$; a precise definition of this is given in Definition 5.1. More generally, [12] define a linear approximation $D_1F\colon \mathcal{B} \to \mathsf{Ch}\mathcal{A}$ for any functor $F\colon \mathcal{B} \to \mathsf{Ch}\mathcal{A}$ so that when this construction is applied to a functor concentrated in degree zero it recovers the construction for $F\colon \mathcal{B} \to \mathcal{A}$.

Now consider a composable pair of functors $G\colon \mathcal{C} \to \mathcal{B}$ and $F\colon \mathcal{B} \to \mathcal{A}$ between abelian categories.[1] Their linear approximations $D_1G\colon \mathcal{C} \to \mathsf{Ch}\mathcal{B}$ and $D_1F\colon \mathcal{B} \to \mathsf{Ch}\mathcal{A}$ are not obviously composable. The standard trick, which appears in [12, Lemma 5.7], is to prolong the second functor, applying D_1F degreewise to a chain complex in $\mathcal{B}$ to produce a chain complex in chain complexes in $\mathcal{A}$, and then convert this double complex into a chain complex in $\mathcal{A}$ by means of the totalization. That is, the composite $D_1F \circ D_1G$ is defined to be the functor

$$D_1F \circ D_1G\colon \mathcal{C} \xrightarrow{D_1G} \mathsf{Ch}\mathcal{B} \xrightarrow{\mathsf{Ch}(D_1F)} \mathsf{Ch}\mathsf{Ch}\mathcal{A} \xrightarrow{\mathrm{Tot}} \mathsf{Ch}\mathcal{A}.$$

If G is reduced, then [12, Lemma 5.7] (a stronger version of which appears as Proposition 5.7 below) proves that this composite is quasi-isomorphic to $D_1(F \circ G)$, i.e., that D_1 is "functorial up to quasi-isomorphism" with respect to the composition structure just introduced.

To keep track of which functors should be composable and which are not (and to avoid a proliferation of Ch's) it is convenient to regard both F and D_1F as functors from $\mathcal{B}$ to $\mathcal{A}$. This sort of bookkeeping is effortlessly achieved by the categorical formalism of a Kleisli category for a monad that we now introduce.

Observation 3.1. There is a pseudomonad[2] $\mathsf{Ch}(-)$ acting on the (large) 2-category of abelian categories, arbitrary functors between them, and natural transformations. Here we are not interested in the 2-dimensional aspects, so we instead describe the quotient monad $\mathsf{Ch}(-)$ acting on the 1-category AbCat of abelian categories and isomorphism classes of functors:

- The monad carries an abelian category $\mathcal{A}$ to the category $\mathsf{Ch}\mathcal{A}$ of non-negatively graded chain complexes in $\mathcal{A}$.
- The monad carries a functor $F\colon \mathcal{B} \to \mathcal{A}$ to its prolongation $\mathsf{Ch}F\colon \mathsf{Ch}\mathcal{B} \to \mathsf{Ch}\mathcal{A}$. Because F might not preserve zero maps, the definition of its prolongation is somewhat delicate. Making use of the Dold–Kan equivalence between non-negatively graded chain complexes and simplicial objects, the functor $\mathsf{Ch}F$ is defined to be the composite

$$\mathsf{Ch}F\colon \mathsf{Ch}\mathcal{B} \xrightarrow{\simeq} \mathcal{B}^{\Delta^{\mathrm{op}}} \xrightarrow{F_*} \mathcal{A}^{\Delta^{\mathrm{op}}} \xrightarrow{\simeq} \mathsf{Ch}\mathcal{A},$$

where the action of F on simplicial objects is by post-composition. Note this operation is not strictly functorial: if $G\colon \mathcal{C} \to \mathcal{B}$ and $F\colon \mathcal{B} \to \mathcal{A}$ are composable functors, then $\mathsf{Ch}F \circ \mathsf{Ch}G$ and $\mathsf{Ch}(F \circ G)$

---

[1] As in Section 2, it suffices to assume that the domain of G is a pointed category with finite coproducts, but when considering composite functors it is linguistically convenient to suppose that all of the categories are abelian.

[2] In category theory, the prefix "pseudo" is attached to structures which hold up to specified coherent isomorphism. A *pseudomonad* is given by the same underlying 1-categorical data as a monad, but with structure diagrams which commute only up to specified coherent natural isomorphisms. For a more detailed definition, see [18].

are naturally isomorphic but not identical. This natural isomorphism satisfies additional "coherence" conditions, which imply, in particular, that the composite natural isomorphisms between $\mathsf{Ch}F \circ (\mathsf{Ch}G \circ \mathsf{Ch}H) = (\mathsf{Ch}F \circ \mathsf{Ch}G) \circ \mathsf{Ch}H$ and $\mathsf{Ch}(F \circ (G \circ H)) = \mathsf{Ch}((F \circ G) \circ H)$ coincide. None of this higher coherence structure is visible in the 1-category AbCat where the monad $\mathsf{Ch}(-)$ acts strictly functorially on natural isomorphism classes of functors.

- The components of the unit of the monad are the functors

$$\mathcal{A} \xrightarrow{\ \deg_0\ } \mathsf{Ch}\mathcal{A}$$

that embed $\mathcal{A}$ as the subcategory of chain complexes concentrated in degree zero.
- The components of the multiplication of the monad are the functors

$$\mathsf{Ch}\mathsf{Ch}\mathcal{A} \xrightarrow{\ \mathrm{Tot}\ } \mathsf{Ch}\mathcal{A}$$

that convert a double complex in $\mathcal{A}$ into a chain complex in $\mathcal{A}$ by forming the total complex.[3]

We leave to the reader the straightforward verification that these functors define the components of a monad $\mathsf{Ch}(-)$ on AbCat.

For any category acted upon by a monad there is an associated Kleisli category (see, e.g., [19, 5.2.9]), which we describe explicitly in the special case of $\mathsf{Ch}(-)$ acting on AbCat.

Definition 3.2. There is a (large) category[4] $\mathsf{AbCat}_{\mathsf{Ch}}$ whose:

- objects are abelian categories;
- morphisms $\mathcal{B} \rightsquigarrow \mathcal{A}$ are natural isomorphism classes of functors $\mathcal{B} \to \mathsf{Ch}\mathcal{A}$;
- identity morphisms $\mathcal{A} \rightsquigarrow \mathcal{A}$ are the functors $\deg_0 \colon \mathcal{A} \to \mathsf{Ch}\mathcal{A}$; and in which
- composition of morphisms $\mathcal{C} \rightsquigarrow \mathcal{B}$ and $\mathcal{B} \rightsquigarrow \mathcal{A}$, corresponding to the pair of functors $G \colon \mathcal{C} \to \mathsf{Ch}\mathcal{B}$ and $F \colon \mathcal{B} \to \mathsf{Ch}\mathcal{A}$, is defined by

$$F \circ G : \mathcal{C} \xrightarrow{\ G\ } \mathsf{Ch}\mathcal{B} \xrightarrow{\ \mathsf{Ch}F\ } \mathsf{Ch}\mathsf{Ch}\mathcal{A} \xrightarrow{\ \mathrm{Tot}\ } \mathsf{Ch}\mathcal{A}$$

Note that AbCat defines a subcategory of $\mathsf{AbCat}_{\mathsf{Ch}}$, where a functor $F \colon \mathcal{B} \to \mathcal{A}$ is identified with the morphism $\mathcal{B} \rightsquigarrow \mathcal{A}$ in $\mathsf{AbCat}_{\mathsf{Ch}}$ represented by the functor

$$\mathcal{B} \xrightarrow{\ F\ } \mathcal{A} \xrightarrow{\ \deg_0\ } \mathsf{Ch}\mathcal{A}$$

Note that the results of Section 2 apply to the homs in the Kleisli category $\mathsf{AbCat}_{\mathsf{Ch}}$ by considering the functor categories $\mathrm{Fun}(\mathcal{B}, \mathsf{Ch}\mathcal{A})$, henceforth denoted simply by $\mathrm{Fun}(\mathcal{B}, \mathcal{A})$.

Convention 3.3. Henceforth, we work in the Kleisli category $\mathsf{AbCat}_{\mathsf{Ch}}$ without further comment. Practically, this means that we may simply write "$F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$" to denote what is really a functor $F \colon \mathcal{B} \to \mathsf{Ch}\mathcal{A}$. There is

---

[3] Here a *double complex* is a chain complex of chain complexes, whose squares commute. On account of this convention, a sign must be introduced into the definition of the differentials in the totalization. See, for example, Remark A.4.

[4] It is more categorically natural to describe the Kleisli bicategory for the pseudomonad $\mathsf{Ch}(-)$ acting on the 2-category of abelian categories, functors, and natural transformations. The category $\mathsf{AbCat}_{\mathsf{Ch}}$ is then the quotient 1-category whose morphisms are isomorphism classes of parallel 1-cells. While it is somewhat inelegant to define the morphisms in a 1-category to be isomorphism class of functors, much of our work actually takes place in a further quotient of $\mathsf{AbCat}_{\mathsf{Ch}}$ where naturally chain homotopically equivalent functors are identified; see Definition 3.5.

no ambiguity in the special case where this functor is concentrated in degree zero, i.e., is actually a functor valued in $\mathcal{A}$. Composition of functors is implemented by the Kleisli construction: by prolonging the second functor and taking the totalization of the resulting double complex. In particular, all functors whose target is an abelian category are regarded as chain-complex valued, so for instance we can safely refer to chain homotopy equivalences classes of functors or ask whether a given functor is objectwise contractible.

In everything that follows, it is always possible to relax the hypotheses on the domain category and assume only the existence of finite coproducts and a zero object, but we often refer simply to functors "between abelian categories" to avoid being overly pedantic with our language. Accordingly, we now write "$\oplus$" and "0" for the coproduct and basepoint in our domain categories.

The Kleisli category $\mathsf{AbCat_{Ch}}$ is the appropriate context to study composition relations involving the Johnson–McCarthy polynomial functors $P_n \colon \mathrm{Fun}(\mathcal{B}, \mathcal{A}) \to \mathrm{Fun}(\mathcal{B}, \mathcal{A})$ and linearization functors $D_1 \colon \mathrm{Fun}(\mathcal{B}, \mathcal{A}) \to \mathrm{Fun}(\mathcal{B}, \mathcal{A})$. However, there remains one additional technical issue that needs to be addressed. Many of the key properties of [12] are proven up to pointwise quasi-isomorphism in the codomain $\mathsf{Ch}\mathcal{A}$. In fact, [12] introduce various models of the functors $P_n F$ and $D_1 F$ that are only well-defined up to quasi-isomorphism. Unfortunately, quasi-isomorphisms are only preserved under composition with exact functors, while the functors we include as morphisms in $\mathsf{AbCat_{Ch}}$ are far from being exact. Thus, if we take the point of view that the functor $D_1 G$ is only defined up to quasi-isomorphism, the Kleisli composite $F \circ D_1 G$ will not be well-defined for arbitrary functors F.

However, the composition operation in $\mathsf{AbCat_{Ch}}$ does respect chain homotopy equivalence of functors, and with some care we will show in Sections 4 and 5 that the universal properties up to quasi-isomorphism established in [12] in fact hold up to chain homotopy equivalence. Two functors $H, G \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ are *pointwise chain homotopy equivalent* if the chain complexes $H(X)$ and $G(X)$ are chain homotopy equivalent in $\mathsf{Ch}\mathcal{A}$ for every object $X \in \mathcal{B}$, and *naturally chain homotopy equivalent* if the chain complexes $H(X)$ and $G(X)$ are chain homotopy equivalent in $\mathsf{Ch}\mathcal{A}$, naturally in $X \in \mathcal{B}$. The following lemma proves that composition defined in the Kleisli category $\mathsf{AbCat_{Ch}}$ respects (natural) chain homotopy.

Lemma 3.4. *Suppose $G, H \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ are (naturally) chain homotopic functors. Then for any pair of functors $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and $K \colon \mathcal{D} \rightsquigarrow \mathcal{C}$, the composite functors*

$$FGK, FHK \colon \mathcal{D} \rightsquigarrow \mathcal{A}$$

are again (naturally) chain homotopic.

Proof. It is obvious that if G and H are pointwise or naturally chain homotopic, then so are the restrictions GK and HK. The non-trivial part is to show that a (natural) chain homotopy between G and H is preserved by post-composition with F.

We first argue that if $G, H \colon \mathcal{C} \to \mathsf{Ch}\mathcal{B}$ are pointwise chain homotopic, then the composite functors

$$\mathcal{C} \xrightarrow{\ G\ } \mathsf{Ch}\mathcal{B} \xrightarrow{\ \mathsf{Ch}F\ } \mathsf{Ch}\mathsf{Ch}\mathcal{A} \qquad \text{and} \qquad \mathcal{C} \xrightarrow{\ H\ } \mathsf{Ch}\mathcal{B} \xrightarrow{\ \mathsf{Ch}F\ } \mathsf{Ch}\mathsf{Ch}\mathcal{A}$$

are pointwise chain homotopic, where we define chain homotopies in $\mathsf{Ch}\mathsf{Ch}\mathcal{A}$ as in any category of chain complexes valued in an abelian category (which in this case happens to be $\mathsf{Ch}\mathcal{A}$). This is a consequence of the Dold–Kan prolongation used to define the monad $\mathsf{Ch}(-)$. The Dold–Kan equivalence takes chain homotopies of chain complexes in $\mathsf{Ch}\mathcal{B}$ to simplicial homotopies of simplicial objects in $\mathcal{B}^{\Delta^{\mathrm{op}}}$ and vice-versa. Simplicial homotopies in $\mathcal{B}^{\Delta^{\mathrm{op}}}$ are structurally defined, and so preserved by post-composition with $F \colon \mathcal{B} \to \mathsf{Ch}\mathcal{A}$. Thus, we conclude that a pointwise chain homotopy between G and H is carried to a pointwise chain homotopy between $\mathsf{Ch}F \circ G$ and $\mathsf{Ch}F \circ H$.

To prove that the Kleisli composites $F \circ G, F \circ H : \mathcal{C} \to \mathsf{Ch}\mathcal{A}$ are pointwise chain homotopic, we need only argue that the totalization functor $\mathrm{Tot} : \mathsf{ChCh}\mathcal{A} \to \mathsf{Ch}\mathcal{A}$ preserves chain homotopies. The chain homotopies we are considering in $\mathsf{ChCh}\mathcal{A}$ are a special case of the more general notion of "chain homotopy of bicomplexes," so this follows from [21, 5.7.3].

If G and H are chain homotopic by a natural chain homotopy, then the Dold–Kan prolongation will produce a natural chain homotopy between $\mathsf{Ch}F \circ G$ and $\mathsf{Ch}F \circ H$. Similarly, the totalization functor will preserve naturality, so that $F \circ G$ and $F \circ H$ are naturally chain homotopic. $\quad\square$

Each hom-set in the category $\mathsf{AbCat_{Ch}}$ is equipped with a "homotopy equivalence of functors" relation defined as pointwise chain homotopy equivalence in the codomain of the functor. We choose to use pointwise chain homotopy equivalence to define our homotopy relation rather than natural chain homotopy equivalence in order to make it simpler to establish the chain homotopies needed in what follows. We will often omit the word "pointwise" when we mean "pointwise chain homotopy". Lemma 3.4 implies that the pointwise chain homotopy equivalence classes in each $\mathrm{Fun}(\mathcal{B}, \mathcal{A})$ are respected by the composition operation in $\mathsf{AbCat_{Ch}}$. We denote these equivalence classes by $[\mathcal{B}, \mathcal{A}]$, and we let $\mathsf{HoAbCat_{Ch}}$ denote the category with the same objects as $\mathsf{AbCat_{Ch}}$ and with hom-sets $[\mathcal{B}, \mathcal{A}]$. This is the advertised categorical context for abelian functor calculus.

Definition 3.5 *(A category for functor calculus).* There is a (large) category $\mathsf{HoAbCat_{Ch}}$ whose:

- objects are abelian categories;
- morphisms $\mathcal{B} \rightsquigarrow \mathcal{A}$ are pointwise chain homotopy equivalence classes of functors $\mathcal{B} \to \mathsf{Ch}\mathcal{A}$;
- identity morphisms $\mathcal{A} \rightsquigarrow \mathcal{A}$ are the functors $\deg_0 : \mathcal{A} \to \mathsf{Ch}\mathcal{A}$; and in which
- composition of morphisms $\mathcal{C} \rightsquigarrow \mathcal{B}$ and $\mathcal{B} \rightsquigarrow \mathcal{A}$, corresponding to the pair of functors $G : \mathcal{C} \to \mathsf{Ch}\mathcal{B}$ and $F : \mathcal{B} \to \mathsf{Ch}\mathcal{A}$, is defined by

$$F \circ G : \mathcal{C} \xrightarrow{\ G\ } \mathsf{Ch}\mathcal{B} \xrightarrow{\ \mathsf{Ch}F\ } \mathsf{ChCh}\mathcal{A} \xrightarrow{\ \mathrm{Tot}\ } \mathsf{Ch}\mathcal{A}$$

This category is an identity-on-objects quotient of $\mathsf{AbCat_{Ch}}$ with hom-sets denoted by $[\mathcal{B}, \mathcal{A}]$ defined by taking pointwise chain homotopy equivalence classes of functors in $\mathrm{Fun}(\mathcal{B}, \mathcal{A})$.

4. The Taylor tower in abelian functor calculus

In this section we review the Taylor tower of a functor constructed in Section 2 of [12]. The constructions we will provide are essentially the same as those in [12], though our presentation differs in two ways that are relevant to our treatment of the chain rule. First, we have chosen to emphasize functors which are not necessarily reduced. The results of Section 2 make it clear that both the reduced and non-reduced cases can be treated simultaneously. Second, we refine the notion of *degree n* functors, replacing the requirement of acyclicity (quasi-isomorphism) in [12, Definition 2.9] with a stronger contractibility (chain homotopy equivalence) condition. Proposition 4.4 proves that the up to quasi-isomorphism universal properties of the polynomial approximations proven in [12, 2.13] become up to chain homotopy equivalence universal properties in the present context.

The Taylor tower will consist of a list of polynomial degree n functors with natural transformations between functors of degree n and degree $n - 1$. We begin this section by explaining what it means for a functor to be *degree n*.

Definition 4.1. A functor $F : \mathcal{B} \rightsquigarrow \mathcal{A}$ is *degree n* if $\mathrm{cr}_{n+1}F : \mathcal{B}^{n+1} \rightsquigarrow \mathcal{A}$ is *contractible*, i.e., pointwise chain homotopy equivalent to zero.

We write "$\simeq$" when there exists a pointwise chain homotopy equivalence between functors. Note that if $\operatorname{cr}_k F \simeq 0$, then $\operatorname{cr}_\ell F \simeq 0$ for all $\ell > k$. This follows from Lemma B.4 and the fact that higher cross effects are direct summands of lower cross effects. In particular, degree k functors are also degree ℓ for all $\ell > k$.

Following [12], we now define the universal polynomial degree n approximations to a functor $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$. From any comonad acting on an abelian category and object in that category one can extract a chain complex [12, Definition 2.4]. The following definition is a combination of the two definitions of $P_n F$ for reduced and unreduced functors of [12, Definition 2.8] into a single definition.

Definition 4.2 ([12]). The *nth polynomial approximation* $P_n F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ of a functor $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ is the functor that carries $X \in \mathcal{B}$ to the totalization of the chain complex of chain complexes in $\mathcal{A}$

$$\cdots \qquad C_{n+1}^{\times 3} F(X) \xrightarrow{\epsilon - C_{n+1}\epsilon + C_{n+1}^{\times 2}\epsilon} C_{n+1}^{\times 2} F(X) \xrightarrow{\epsilon - C_{n+1}\epsilon} C_{n+1} F(X) \xrightarrow{\epsilon} F(X)$$

defined by $(P_n F(X))_k := (C_{n+1})^{\times k} F(X)$ for $k \geq 1$ with differentials defined to be the alternating sums $\sum_i (-1)^i C_{n+1}^{\times i} \epsilon$ of the counit map.

Remark 4.3. When $n = 0$, it is easy to give an explicit computation of the chain complex $P_0 F$. By Lemma 2.4, $C_1^{\times k} F \cong \operatorname{cr}_1 F$ for all $k \geq 1$. Thus, we can explicitly compute this functor as the chain complex of functors

$$\cdots \operatorname{cr}_1 F \xrightarrow{0} \operatorname{cr}_1 F \xrightarrow{\operatorname{id}} \operatorname{cr}_1 F \xrightarrow{0} \operatorname{cr}_1 F \xrightarrow{\epsilon} F,$$

where the differentials continue to alternate between 0 and id. Since $F \cong \operatorname{cr}_1 F \oplus F(0)$, we can rewrite the chain complex $P_0 F$ as a direct sum of the two chain complexes

$$\cdots \operatorname{cr}_1 F \xrightarrow{0} \operatorname{cr}_1 F \xrightarrow{\operatorname{id}} \operatorname{cr}_1 F \xrightarrow{0} \operatorname{cr}_1 F \xrightarrow{id} \operatorname{cr}_1 F$$

and

$$\cdots 0 \longrightarrow 0 \longrightarrow 0 \longrightarrow 0 \longrightarrow F(0).$$

Here we have used the fact that the map $\epsilon \colon \operatorname{cr}_1 F \to F$ is the identity on the component $\operatorname{cr}_1 F$ of $F \cong \operatorname{cr}_1 F \oplus F(0)$. The chain complex on the top line is contractible. Thus, we will use the chain complex in the second line as our model for $P_0 F$, and accordingly we will write $P_0 F(X) \cong F(0)$.

Recalling our convention of working in the Kleisli category described in Section 3, the nth polynomial approximation construction defines a functor $P_n \colon \operatorname{Fun}(\mathcal{B}, \mathcal{A}) \to \operatorname{Fun}(\mathcal{B}, \mathcal{A})$. Proposition 2.10 implies that this functor has good homotopical properties.

Proposition 4.4 ([12, 2.13]). *For any $n \geq 0$,*

(i) $P_n \colon \operatorname{Fun}(\mathcal{B}, \mathcal{A}) \to \operatorname{Fun}(\mathcal{B}, \mathcal{A})$ *is exact.*
(ii) P_n *preserves chain homotopies, chain homotopy equivalences, and contractibility.*

Proof. For (i), let $0 \to F \to G \to H \to 0$ be an exact sequence of functors. By Proposition 2.10,

$$0 \to P_n F \to P_n G \to P_n H \to 0$$

is a sequence of chain complexes that is exact in each degree. Hence, the sequence is exact. The properties enumerated in (ii) are consequences of (i): exact functors preserve chain homotopies and the zero chain complex.　□

The functor $P_n F$ receives a natural transformation $p_n \colon F \to P_n F$ defined by inclusion into the degree zero part of the chain complex $P_n F$. The basic properties of the nth polynomial approximation are summarized in the following proposition, an adaptation of [12, Lemma 2.11].

Proposition 4.5. *For $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$,*

(i) *The functor $P_n F$ is degree n.*
(ii) *If F is degree n, then the map $p_n \colon F \to P_n F$ is a chain homotopy equivalence.*
(iii) *The pair $(P_n F, p_n \colon F \to P_n F)$ is universal up to chain homotopy equivalence with respect to degree n functors receiving natural transformations from F.*

Proof. We start with (i), writing F_k for the kth degree component of $F \colon \mathcal{B} \to \mathsf{Ch}\mathcal{A}$. By definition, $P_n F$ is the total complex of the bicomplex whose kth column is $P_n(F_k)$. To prove that $P_n F$ is degree n, we must show that the bicomplex $\mathrm{cr}_{n+1} P_n F$ is contractible. Using [21, Definition 5.7.3], a contraction for a bicomplex is given by horizontal contractions s^h and vertical contractions s^v which satisfy $s^h d^v = d^v s^h$, $s^v d^h = d^h s^v$, and $1 = (s^h d^h + d^h s^h) + (s^v d^v + d^v s^v)$. In particular, when the bicomplex has naturally contractible columns we see that it is contractible by setting $s^h = 0$. Lemma 2.9, applied to the adjunction of Corollary 2.7, supplies a natural contracting homotopy for each of the columns $\mathrm{cr}_{n+1} P_n F_k$. Since totalization preserves chain homotopy equivalence, we conclude that the chain complex $\mathrm{cr}_{n+1} P_n F$ is contractible.

For (ii), recall that $P_n F$ is the totalization of the first-quadrant bicomplex whose kth row is $C_{n+1}^{\times k} F$. The map $p_n \colon F \to P_n F$ is the natural inclusion of the zeroth row into the totalization. By Corollary A.7, to prove that this is a chain homotopy equivalence, it suffices to show that each chain complex $C_{n+1}^{\times k} F$ is contractible for $k \geq 1$. By exactness of C_{n+1}, it suffices to prove that the chain complex $C_{n+1} F$ is contractible. But the hypothesis that F is degree n tells us immediately that $C_{n+1} F(X) := \mathrm{cr}_{n+1} F(X, \dots, X)$ is contractible.

Part (iii) follows formally. Let $\tau \colon F \to G$ be a natural transformation from F to a degree n functor G. The natural map p_n provides a commutative diagram

$$
\begin{array}{ccc}
F & \xrightarrow{\;\tau\;} & G \\
{\scriptstyle p_n}\big\downarrow & & \simeq\big\downarrow{\scriptstyle p_n} \\
P_n F & \xrightarrow[P_n \tau]{} & P_n G
\end{array}
$$

where the right hand map is a chain homotopy equivalence by part (ii). This shows that τ factors through $p_n \colon F \to P_n F$ up to pointwise chain homotopy equivalence. For uniqueness, consider another factorization

$$
\begin{array}{ccccc}
 & & \overset{\tau}{\frown} & & \\
F & \xrightarrow{\;p_n\;} & P_n F & \xrightarrow{\;\sigma\;} & G \\
{\scriptstyle p_n}\big\downarrow & & {\scriptstyle p_n P_n}\big\downarrow{\scriptstyle \simeq} & & \simeq\big\downarrow{\scriptstyle p_n} \\
P_n F & \xrightarrow[P_n p_n]{} & P_n P_n F & \xrightarrow[P_n \sigma]{} & P_n G \\
 & & \underset{P_n \tau}{\smile} & &
\end{array}
$$

The maps $p_n P_n, P_n p_n \colon P_n F \to P_n P_n F$ are not identical, but do agree up to a natural automorphism of $P_n P_n F$. In particular, because $P_n F$ is degree n by (i), $p_n P_n$ is a chain homotopy equivalence by (ii), and

thus so is $P_n p_n$. In this way, we see that σ is determined up to chain homotopy equivalence by $P_n \tau$, and so is unique up to chain homotopy equivalence.   □

Since $P_{n-1} F$ is also degree n, the universal property of Proposition 4.5(iii) provides a factorization

$$
\begin{array}{ccc}
F & \xrightarrow{\ p_n\ } & P_n F \\
& \searrow{\scriptstyle p_{n-1}} & \Big\downarrow{\scriptstyle q_n} \\
& & P_{n-1} F
\end{array}
$$

The resulting tower of functors:

$$
\cdots \xleftarrow{\quad} P_{n+1} F \xrightarrow{\ q_{n+1}\ } P_n F \xrightarrow{\ q_n\ } P_{n-1} F \longrightarrow \cdots \xrightarrow{\ q_1\ } P_0 F
$$

with maps $p_{n+1}, p_n, p_{n-1}, p_0$ from F.

is called the *algebraic Taylor tower* of F.

Remark 4.6. Explicitly, the map $q_n : P_n F \to P_{n-1} F$ is induced by a natural transformation of comonads $\rho_n : C_{n+1} \to C_n$ given on components by the composite

$$
C_{n+1} F(X) = \mathrm{cr}_{n+1} F(X, \ldots, X) \hookrightarrow \mathrm{cr}_n F(X \oplus X, X, \ldots X) \xrightarrow{\mathrm{cr}_n F(\text{fold})} \mathrm{cr}_n F(X, \ldots, X) = C_n F(X)
$$

of the canonical direct summand inclusion followed by the image of the fold map $X \oplus X \to X$ [12, p. 770].

Example 4.7. Recall the functor $F : \mathcal{A} \to \mathcal{A}$ defined by $F(X) = A \oplus X$ from Example 2.3. Since $\mathrm{cr}_1 F(X) = X = \mathrm{id}(X)$ and $\mathrm{cr}_n F \cong 0$ for $n \geq 2$, it follows immediately from Definition 4.2 that

$$
P_n F(X) = (\cdots \to 0 \to 0 \to F(X))
$$

for all $n \geq 1$. When $n = 0$, it follows immediately from Remark 4.3 that $P_0 F(X) \cong A$. Indeed, since the chain homotopy $P_0 F(X) \simeq F(0)$ is given by the chain map which projects $F(X)$ onto $F(0)$ in degree 0, one sees that the map $q_1 : P_1 F(X) \to P_0 F(X)$, or $q_1 : F(X) \to A$, is also this projection map.

5. Linear approximations

In this section we consider the linear approximation to a functor, or the homotopy fiber of the map $q_1 : P_1 F \to P_0 F$. The linearization of a functor will be an important ingredient in defining the directional derivative and the two types of higher order derivatives considered in Sections 7 and 8. Since the linearization of a functor F with codomain $\mathcal{B}$ is a functor $D_1 F$ with codomain $\mathsf{Ch}\mathcal{B}$, the Kleisli category conventions of Section 3 will be particularly useful here, especially when we consider linearization together with composition of functors. In order to maximize the benefits of this structure, we make two main changes in our presentation of the linear approximation from the presentation in [12].

First, we insist that the linear approximation $D_1 F$ be a morphism in $\mathsf{AbCat_{Ch}}$. In particular, this means we insist that $D_1 F$ take values in *non-negatively graded* chain complexes. In [12, 2.14.7], $D_1 F$ (and indeed all of the homotopy fibers of the maps $q_n : P_n F \to P_{n-1} F$) are defined using a particular model for the homotopy fiber given by taking the mapping cone of q_n shifted down one degree. This shift means that the resulting chain complex will have a non-zero object in degree -1. In Definition 5.1 we give an alternate

model of D_1F which is chain homotopic to the model in [12], but which is concentrated in non-negative degrees.

Second, since composition by prolongation does not always preserve quasi-isomorphisms, we have chosen to work up to the stronger notion of chain homotopy equivalence. In particular, we have provided explicit chain homotopy equivalences for the D_1 chain rule in Proposition 5.7, which strengthens [12, Lemma 5.7]. This chain rule is strengthened further in Proposition 5.10, where we present the analogue of the D_1 chain rule for functors which are not necessarily reduced. Indeed, all of the properties of the linear approximation D_1F are developed up to chain homotopy equivalence, and for functors which need not be reduced. These properties occupy the majority of this section.

Definition 5.1. The *linearization* of $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ is the functor $D_1F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ given as the totalization of the explicit chain complex of chain complexes (D_1F_*, ∂_*) where:

$$(D_1F)_k := \begin{cases} C_2^{\times k} F & k \geq 1 \\ \operatorname{cr}_1 F & k = 0 \\ 0 & \text{otherwise.} \end{cases}$$

The chain differential $\partial_1 \colon (D_1F)_1 \to (D_1F)_0$ is given by the map ρ_1 of 4.6, and the chain differential $\partial_k \colon (D_1F)_k \to (D_1F)_{k-1}$ is given by $\sum_{i=0}^{k-1} (-1)^i C_2^{\times i} \epsilon$ when $k \geq 1$. Note that in the case $n = 1$, the map ρ_1 is precisely the counit ϵ for the component $\operatorname{cr}_1 F$ (see Remark 2.8). Since $C_2 F \cong C_2 \operatorname{cr}_1 F$ by Lemma 2.4, the fact that $\partial_1 \circ \partial_2 = 0$ follows immediately by construction.

Remark 5.2.

(i) Using the model of P_0F described in Remark 4.3, the map $q_0 \colon P_1F \to P_0F$ is surjective, so the homotopy fiber D_1F is simply the kernel of q_0. There is an inclusion map from D_1F into the homotopy fiber of q_1, and this inclusion is a chain homotopy equivalence. Thus, this model is chain homotopy equivalent to the one defined in [12].

(ii) Viewed another way, we see that Definition 5.1 is exactly the same chain complex as $P_1(\operatorname{cr}_1 F)$ by Lemma 2.4. The latter is the degree 1 approximation to the reduction of F.

These two observations imply that, up to chain homotopy equivalence, the two processes of reducing a functor and taking its degree 1 approximation produce the same result when applied in either order.

Example 5.3. Returning to the functor $F(X) = A \oplus X$ from Examples 2.3 and 4.7, since $\operatorname{cr}_2 F \cong 0$ and $\operatorname{cr}_1 F \cong \operatorname{id}$, the linearization D_1F is given by the chain complex which is the functor id concentrated in degree 0. Thus, the linearization of F is id.

We now turn our attention to the properties of D_1. First, we see that the linearization construction defines an exact functor:

Proposition 5.4.

(i) $D_1 \colon \operatorname{Fun}(\mathcal{B}, \mathcal{A}) \to \operatorname{Fun}(\mathcal{B}, \mathcal{A})$ *is exact.*
(ii) D_1 *preserves chain homotopies, chain homotopy equivalences, and contractibility.*

Proof. Since D_1F is isomorphic to $P_1(\operatorname{cr}_1 F)$, (i) is a direct consequence of Proposition 4.4, and (ii) follows from exactness. $\square$

Next, we would like to justify the use of the term *linearization* for D_1. We start by defining what it means for a functor to be linear. Recall that a functor is degree one if $cr_2 F$ is contractible, as in Definition 4.1.

Definition 5.5. A functor $F \colon \mathcal{B} \to \mathcal{A}$ is *linear* if it is degree one and also reduced, meaning that $F(0)$ is contractible. Equivalently, F is linear if it preserves finite direct sums up to chain homotopy equivalence.

We will start with reduced functors. If F is reduced, $D_1 F$ is exactly the analog of the linear approximation of a function for functors. That is, a function $f \colon \mathbb{R} \to \mathbb{R}$ whose graph passes through the origin is linear if $f(x + y) = f(x) + f(y)$.

Lemma 5.6.

(i) *For any $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$, the functor $D_1 F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ is strictly reduced, and for any $X, Y \in \mathcal{B}$, the natural map*

$$D_1 F(X) \oplus D_1 F(Y) \hookrightarrow D_1 F(X \oplus Y)$$

is a chain homotopy equivalence. In particular, $D_1 F$ is linear.

(ii) *The functor $D_1 \colon [\mathcal{B}, \mathcal{A}] \to [\mathcal{B}, \mathcal{A}]$ is linear in the sense that $D_1 0 \cong 0$ and for any pair of functors $F, G \in [\mathcal{B}, \mathcal{A}]$,*

$$D_1 F \oplus D_1 G \cong D_1 (F \oplus G).$$

Proof. First, note that $D_1 F$ is strictly reduced because the functors $C_2 F$ and $cr_1 F$ are both strictly reduced functors.

To prove the second part of (i), note that the map

$$i \colon D_1 F(X) \oplus D_1 F(Y) \to D_1 F(X \oplus Y)$$

is the inclusion of a direct summand with complement $cr_2 D_1 F(X, Y)$. But $D_1 F \cong P_1(cr_1 F)$ by Remark 5.2, and $cr_2 P_1(cr_1 F)(X, Y)$ is chain contractible by Proposition 4.5(i). It follows that the map i is a chain homotopy equivalence.

Part (ii) is a consequence of Proposition 5.4: any exact functor preserves finite direct sums up to isomorphism. $\square$

Johnson and McCarthy prove that D_1 is functorial up to quasi-isomorphism whenever the first of a composable pair of functors is reduced. The proof of [12, Lemma 5.7] proceeds by showing that both $D_1(F \circ G)$ and $D_1 F \circ D_1 G$ are quasi-isomorphic to a third chain complex, coming from a tricomplex containing each of the other two. The next proposition, whose technical proof is deferred to Appendix B, shows that in fact $D_1(F \circ G)$ is chain homotopy equivalent to $D_1 F \circ D_1 G$.

Proposition 5.7. *For any composable pair of functors $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ with G reduced, there is a chain homotopy equivalence*

$$D_1(F \circ G) \simeq D_1 F \circ D_1 G.$$

Our next task is to extend Proposition 5.7 to not necessarily reduced functors $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ and $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$. The proof follows easily from the next two lemmas. The first of these indicates that the linearization of F is the same as the linearization of its reduced component, $cr_1 F$. This is analogous to the statement that the

linearizations of the functions $f(x)$ and $f(x) - f(0)$ have the same slope. The second function, $f(x) - f(0)$, is reduced in the sense that its graph goes through the origin. This analogy explains precisely how to linearize unreduced functors.

Lemma 5.8. *For any $F: \mathcal{B} \rightsquigarrow \mathcal{A}$, the natural map $D_1\mathrm{cr}_1 F \to D_1 F$ is an isomorphism.*

Proof. By Lemma 2.4 and Corollary 2.7, $C_2\mathrm{cr}_1 F \cong C_2 F$ and $\mathrm{cr}_1\mathrm{cr}_1 F \cong \mathrm{cr}_1 F$. The result now follows from the definition of $D_1 F$. $\square$

Lemma 5.9. *Let $F: \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G: \mathcal{C} \rightsquigarrow \mathcal{B}$ be composable functors. Then*

$$\mathrm{cr}_1(F \circ G)(X) \cong (\mathrm{cr}_1 F \circ \mathrm{cr}_1 G)(X) \oplus \mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G(X)).$$

Proof. Note that

$$\mathrm{cr}_1(F \circ G)(X) \oplus F(G(0)) \cong F(G(X)) \cong F(G(0) \oplus \mathrm{cr}_1 G(X))$$
$$\cong F(0) \oplus \mathrm{cr}_1 F(G(0) \oplus \mathrm{cr}_1 G(X))$$
$$\cong F(0) \oplus \mathrm{cr}_1 F(G(0)) \oplus \mathrm{cr}_1 F(\mathrm{cr}_1 G(X)) \oplus \mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G(X)).$$

On the other hand, $F(G(0)) \cong F(0) \oplus \mathrm{cr}_1 F(G(0))$. Taking complements we conclude the claimed isomorphism. $\square$

Proposition 5.10. *If $F: \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G: \mathcal{C} \rightsquigarrow \mathcal{B}$ are composable functors, then there is a chain homotopy equivalence*

$$D_1(F \circ G) \simeq D_1 F \circ D_1 G \oplus D_1\mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G).$$

Note that when G is reduced, $\mathrm{cr}_2 F(G(0), G(X)) = \mathrm{cr}_2 F(0, G(X)) \simeq 0$, so the second term vanishes. Thus Proposition 5.10 reduces to Proposition 5.7 in this case.

Proof. Taking the linearization of $F \circ G$, or equivalently, the linearization of $\mathrm{cr}_1(F \circ G)$, we have

$$D_1(F \circ G) \cong D_1\left(\mathrm{cr}_1(F \circ G)\right) \cong D_1\left((\mathrm{cr}_1 F \circ \mathrm{cr}_1 G) \oplus \mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G)\right),$$

by Lemmas 5.8 and 5.9. Using the linearity of D_1 established in Lemma 5.6, this is isomorphic to

$$\cong D_1(\mathrm{cr}_1 F \circ \mathrm{cr}_1 G) \oplus D_1\mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G).$$

Since $\mathrm{cr}_1 G$ is reduced, we can apply Proposition 5.7 to the first summand to obtain a chain homotopy equivalence between the last term and

$$\simeq D_1(\mathrm{cr}_1 F) \circ D_1(\mathrm{cr}_1 G) \oplus D_1\mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G).$$

Applying the isomorphisms $D_1(\mathrm{cr}_1 F) \cong D_1 F$ and $D_1(\mathrm{cr}_1 G) \cong D_1 G$ of Lemma 5.8, we obtain:

$$\cong D_1 F \circ D_1 G \oplus D_1\mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G). \qquad \square$$

We finish this section with a few observations about linearizations of functors of more than one variable. When $F: \mathcal{B}^n \rightsquigarrow \mathcal{A}$ is a functor of n variables, it is possible to linearize F with respect to a subset of the variables.

Convention 5.11. Given a functor $F \colon \mathcal{B}^n \rightsquigarrow \mathcal{A}$, let $F_i \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ be the functor defined by

$$F_i(Y) := F(X_1, \ldots, X_{i-1}, Y, X_{i+1}, \ldots, X_n)$$

where $X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n$ are fixed objects of $\mathcal{B}$. Write $D_1^i F(X_1, \ldots, X_n)$ for $D_1 F_i(X_i)$.

When linearizing F with respect to two or more variables, we have more than one option for how to proceed.

- We can linearize F with respect to two or more variables *simultaneously*. For $i < j$, the linearization of F with respect to its i and j variables simultaneously is the linearization of the functor $F_{i \times j} \colon \mathcal{B} \times \mathcal{B} \rightsquigarrow \mathcal{A}$ defined by

$$F_{i \times j}(Y, Z) := F(X_1, \ldots, X_{i-1}, Y, X_{i+1}, \ldots, X_{j-1}, Z, X_{j+1}, \ldots, X_n)$$

  where the X_k's are fixed objects in $\mathcal{B}$. We denote this simultaneous linearization by $D_1^{i \times j} F$. Similarly, we write $D_1^{i_1 \times \cdots \times i_k} F$ for the linearization with respect to k variables simultaneously.
- We can linearize F with respect to two or more variables *sequentially*. For $i < j$, the linearization of F with respect to its i and j variables sequentially is the linearization of the functor $D_1^i F$ with respect to the jth variables. That is,

$$D_1^j(D_1^i F) := D_1(D_1^i F)_j$$

  where $(D_1^i F)_j$ is the functor obtained from $D_1^i F$ by holding all but the jth variable constant, as in 5.11. The special case in which $F \colon \mathcal{B}^n \rightsquigarrow \mathcal{A}$ has been linearized with respect to each of its n variables sequentially is denoted $D_1^{(n)} F$. That is,

$$D_1^{(n)} F := D_1^n \cdots D_1^1 F.$$

Since the variables of a given multilinear functor are not always ordered, it is possible for confusion to arise when we are multilinearizing a functor either simultaneously or sequentially. To disambiguate, we will use the name of the variable in this case. For example, the notation

$$D_1^{X \times Y} \mathrm{cr}_3 F(Z, Y, X)$$

indicates that we should simultaneously linearize the third cross effect of F with respect to the variables X and Y, which occur in the second and third slots of $\mathrm{cr}_3 F$, respectively. We could equally well use the notation $D_1^{2 \times 3} \mathrm{cr}_3 F$ to indicate the same multilinearization, and context will determine which one is more practical.

The simultaneous linearization of functors of several variables is often a trivialization of the functor, as in the next lemmas.

Lemma 5.12. *Suppose that $H \colon \mathcal{B}^n \rightsquigarrow \mathcal{A}$ is strictly multi-reduced. Then for any $1 \leq i < j \leq n$, $D_1^{i \times j} H(X_1, \ldots, X_n)$ is contractible.*
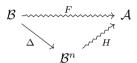
Proof. For simplicity of notation, we prove the lemma for a functor $H(X, Y)$ of two variables. Since $D_1^{1 \times 2} H(X, Y)$ is linear simultaneously in the variables (X, Y), the inclusion

$$D_1^{1 \times 2} H((X_1, Y_1) \oplus (X_2, Y_2)) \to D_1^{1 \times 2} H(X_1, Y_1) \oplus D_1^{1 \times 2} H(X_2, Y_2)$$

is a chain homotopy equivalence by Lemma 5.6. Note that $(X, Y) \cong (X, 0) \oplus (0, Y)$, so as a special case $D_1^{1 \times 2} H(X, Y)$ is chain homotopic to $D_1^{1 \times 2} H(X, 0) \oplus D_1^{1 \times 2} H(0, Y)$, which is zero because H is strictly multi-reduced. Thus $D_1^{1 \times 2} H(X, Y)$ must be contractible as well. □

Lemma 5.12 is often applied in the following form:

Corollary 5.13. *Suppose* $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ *factors as*

where $\Delta \colon \mathcal{B} \to \mathcal{B}^n$ *is the diagonal functor and* H *is strictly multi-reduced. Then* $D_1 F$ *is contractible.*

Proof. The diagonal functor is strictly reduced so by Proposition 5.7, $D_1 F$ is chain homotopic to $D_1 H \circ D_1 \Delta$, but here $D_1 H$ is a simultaneous linearization to which Lemma 5.12 applies, implying that $D_1 H$ is contractible. Hence we have a chain contraction $D_1 F \simeq 0$. □

Example 5.14. The simultaneous linearization of the cross-effects functor $\mathrm{cr}_n F(X_1, \ldots, X_n)$ is contractible by Lemma 5.12. However, the sequential multilinearization

$$D_1^{(n)} \mathrm{cr}_n F(X_1, \ldots, X_n)$$

is generally not contractible. If $D_1^{(n)} \mathrm{cr}_n F(X_1, \ldots, X_n)$ were contractible, then its homotopy orbits $\left(D_1^{(n)} \mathrm{cr}_n F(X_1, \ldots, X_n) \right)_{h \Sigma_n}$ would be acyclic since homotopy orbits preserve quasi-isomorphisms. However, $\left(D_1^{(n)} \mathrm{cr}_n F(X, \ldots, X) \right)_{h \Sigma_n}$ is quasi-isomorphic to the nth layer of the Taylor tower of F by [12, Proposition 3.9]. The nth layer of the Taylor tower of F, $D_n F$, is the homotopy fiber of $q_n \colon P_n F \to P_{n-1} F$. In many cases the nth layers of Taylor towers are known to be non-trivial (see, e.g., [12, Section 6] and [15]).

When linearizing sequentially, it does not matter in which order we linearize the variables.

Lemma 5.15. *For any* $F \colon \mathcal{B}^n \rightsquigarrow \mathcal{A}$, *there is an isomorphism*

$$D_1^i D_1^j F \cong D_1^j D_1^i F.$$

Proof. By using the definition of the cross effects, it is possible to show that $\mathrm{cr}_n(\mathrm{cr}_m F_i)_j \cong \mathrm{cr}_m(\mathrm{cr}_n F_j)_i$, where F_i denotes the functor defined in Convention 5.11. This is a somewhat lengthy, though straightforward, argument. Instead we simply note that both functors are reduced and linear in the ith and jth variables, respectively and indeed that $P_1^i P_1^j (\mathrm{cr}_1^i \mathrm{cr}_1^j F)$ and $P_1^j P_1^i (\mathrm{cr}_1^j \mathrm{cr}_1^i F)$ satisfy the same universal property. □

Finally, it will be useful to record how linearization behaves with respect to products and projections. The remaining results in this section describe this behavior.

Lemma 5.16. *Let* $\pi \colon \mathcal{B} \times \mathcal{B} \to \mathcal{B}$ *be the functor* $\pi \colon (X, Y) \mapsto X$. *Then* $D_1^{X \times Y} \pi \cong \pi$.

Proof. The functor $\pi \colon \mathcal{B} \times \mathcal{B} \to \mathcal{B}$ is strictly linear so is isomorphic to its linearization. □

Corollary 5.17. *For any $F\colon \mathcal{B} \rightsquigarrow \mathcal{A}$, let $\tilde{F}\colon \mathcal{B} \times \mathcal{B} \rightsquigarrow \mathcal{A}$ be the functor defined by*

$$\tilde{F}(X,Y) := F(X).$$

Then $D_1^{X \times Y} \tilde{F}(X) \cong D_1^X F(X)$.

Proof. By direct computation,

$$F(X) \cong \tilde{F}(X,Y) \cong \tilde{F}(0,0) \oplus \mathrm{cr}_1 \tilde{F}(X,Y) \cong F(0) \oplus \mathrm{cr}_1 \tilde{F}(X,Y),$$

from which we conclude that $\mathrm{cr}_1 \tilde{F}(X,Y) \cong \mathrm{cr}_1 F(X)$. Similarly,

$$\mathrm{cr}_1 \tilde{F}((X_1,Y_1) \oplus (X_2,Y_2)) \cong \mathrm{cr}_1 \tilde{F}(X_1,Y_1) \oplus \mathrm{cr}_1 \tilde{F}(X_2,Y_2) \oplus \mathrm{cr}_2 \tilde{F}((X_1,Y_1),(X_2,Y_2))$$
$$= \mathrm{cr}_1 F(X_1) \oplus \mathrm{cr}_1 F(X_2) \oplus \mathrm{cr}_2 \tilde{F}((X_1,Y_1),(X_2,Y_2)).$$

Since $\mathrm{cr}_1 \tilde{F}((X_1,Y_1) \oplus (X_2,Y_2)) \cong \mathrm{cr}_1 F(X_1 \oplus X_2)$, we conclude from the definition of $\mathrm{cr}_2 F$ that $\mathrm{cr}_2 F(X_1,X_2) \cong \mathrm{cr}_2 \tilde{F}((X_1,Y_1),(X_2,Y_2))$. By definition of D_1, we now conclude that $D_1 \tilde{F}(X,Y) \cong D_1 F(X)$. □

Note that $\tilde{F}$ is a composition of the functor F and the reduced functor π; thus it is possible to obtain Corollary 5.17 as an immediate consequence of Lemma 5.16 and Proposition 5.7. However, this result would only allow us to conclude that $D_1 \tilde{F}$ and $D_1 F$ are chain homotopy equivalent. The direct proof provides a stronger result.

Lemma 5.18. *For a functor $\langle F,G \rangle \colon \mathcal{C} \rightsquigarrow \mathcal{A} \times \mathcal{B}$ with components $F\colon \mathcal{C} \rightsquigarrow \mathcal{A}$ and $G\colon \mathcal{C} \rightsquigarrow \mathcal{B}$,*

$$D_1 \langle F,G \rangle \cong \langle D_1 F, D_1 G \rangle.$$

Proof. In the product abelian category $\mathcal{A} \times \mathcal{B}$, direct sums are defined coordinate-wise:

$$(A_1,B_1) \oplus (A_2,B_2) \cong (A_1 \oplus A_2, B_1 \oplus B_2).$$

This implies that the cross effects are defined coordinate-wise, and the result follows. □

6. The first directional derivative

Applying the results of the previous section, we will now define the directional derivative for a functor in analogy with the directional derivative in function calculus. We will show that our definition, which is new, recovers the directional derivatives defined in [12]. We review, generalize, and prove new properties of this functor, in analogy with the directional derivative from functional calculus except that equations are replaced with chain homotopy equivalences. These results can be summarized in a single theorem: the directional derivative satisfies the axioms for a categorical derivative in the sense of [4], equipping HoAbCat$_{\mathsf{Ch}}$ with the structure of a cartesian differential category. In particular, we show that the directional derivative satisfies the chain rule. The properties which we need in order to establish that HoAbCat is a differential category are precisely the properties we will need in order to obtain a higher order chain rule analogous to [10, Theorem 3], which we accomplish in Section 8.

The directional derivative of a differentiable function $f\colon \mathbb{R}^n \to \mathbb{R}^m$ (or more generally, a function between Banach spaces) at the point $x \in \mathbb{R}^n$ in the direction $v \in \mathbb{R}^n$ measures how the value of f at x changes while translating along the infinitesimal vector from x in the direction v. One way to make this idea precise

is to define $\nabla f(v; x)$ to be the derivative of the composite function, substituting the affine linear function $t \mapsto x + tv$ into the argument of f, evaluated at $t = 0$:

$$\nabla f(v; x) = \frac{\partial}{\partial t} f(x + tv)\big|_{t=0}.$$

This construction motivates the following definition for functors of abelian categories.

Definition 6.1. For $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and $X, V \in \mathcal{B}$, define a functor $\nabla F \colon \mathcal{B} \times \mathcal{B} \rightsquigarrow \mathcal{A}$ by

$$\nabla F(V; X) := D_1 F(X \oplus -)(V).$$

Alternatively, [12] define the directional derivative for a functor $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ from a category with finite coproducts and basepoint to an abelian category by using a sequence of analogies with the formula

$$\nabla f(v; x) = \lim_{t \to 0} \frac{1}{t} [f(x + tv) - f(x)]$$

from the calculus of functions of a real variable. Let $X \in \mathcal{B}$ be the 'point', corresponding to the point $x \in \mathbb{R}^n$ at which we will evaluate our derivative. Let $V \in \mathcal{B}$ denote the 'direction'. Sums in $\mathcal{B}$ are coproducts, which we will denote $\oplus$. The difference operation in the expression $f(x + tv) - f(x)$ becomes a kernel operation in our analogy,

$$\ker (F(X \oplus V) \to F(X))$$

where the map $F(X \oplus V) \to F(X)$ in $\mathcal{A}$ is obtained by applying F to the map $X \oplus V \to X$ in $\mathcal{B}$ which sends the V component to the basepoint. The linearization D_1 (in the variable V) produces the degree 1, reduced component of the functor, and hence corresponds to taking the limit as t goes to 0 in the expression for the directional derivative of f.

Definition 6.2. Let $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ be a functor of abelian categories, and let X and V be objects in $\mathcal{B}$. The directional derivative of F at X in the direction V is the bifunctor defined by

$$\nabla F(V; X) := D_1^V \left(\ker(F(X \oplus V) \to F(X)) \right).$$

The following lemma proves that our two definitions agree.

Lemma 6.3. *For a functor $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ between abelian categories and any pair of objects $X, V \in \mathcal{B}$, there is an isomorphism*

$$D_1^V \left(\ker(F(X \oplus V) \to F(X)) \right) \cong D_1 F(V) \oplus D_1^V \operatorname{cr}_2 F(X, V)$$

and a chain homotopy equivalence

$$D_1 F(X \oplus -)(V) \simeq D_1 F(V) \oplus D_1^V \operatorname{cr}_2 F(X, V).$$

Proof.
For the first isomorphism, recall that by the definition of the cross effects,

$$F(X \oplus V) \cong F(0) \oplus \operatorname{cr}_1 F(X) \oplus \operatorname{cr}_1 F(V) \oplus \operatorname{cr}_2 F(X, V) \cong F(X) \oplus \operatorname{cr}_1 F(V) \oplus \operatorname{cr}_2 F(X, V).$$

The direct summand inclusion $F(X) \to F(X \oplus V)$ is a section to the projection $F(X \oplus V) \to F(X)$, so the kernel $\ker (F(X \oplus V) \to F(X))$ is isomorphic to $\mathrm{cr}_1 F(V) \oplus \mathrm{cr}_2 F(X, V)$. By Lemma 5.6(ii), D_1^V is linear, so

$$D_1^V (\ker(F(X \oplus V) \to F(X))) \cong D_1^V (\mathrm{cr}_1 F(V) \oplus \mathrm{cr}_2 F(X, V))$$
$$\cong D_1^V \mathrm{cr}_1 F(V) \oplus D_1^V \mathrm{cr}_2 F(X, V)$$
$$\cong D_1^V F(V) \oplus D_1^V \mathrm{cr}_2 F(X, V),$$

with the last simplification by Lemma 5.8.

For the chain homotopy equivalence, apply Proposition 5.10 to F and $X \oplus - : \mathcal{B} \to \mathcal{B}$ to obtain a chain homotopy equivalence

$$D_1 F(X \oplus -)(V) \simeq (D_1 F) \circ (D_1(X \oplus -))(V) \oplus D_1^V \mathrm{cr}_2 F(X, \mathrm{cr}_1(X \oplus V))$$

We simplify the cross effect term on the right using Lemma 5.8:

$$\simeq (D_1 F) \circ (D_1(X \oplus -))(V) \oplus D_1^2 \mathrm{cr}_2 F(X, D_1^V \mathrm{cr}_1(X \oplus V))$$
$$\cong (D_1 F) \circ (D_1(X \oplus -))(V) \oplus D_1^2 \mathrm{cr}_2 F(X, D_1^V (X \oplus V)).$$

Example 5.3 computes that $D_1(X \oplus -)$ is the identity functor, so this simplifies to:

$$\cong D_1 F(V) \oplus D_1^V \mathrm{cr}_2 F(X, V). \qquad \square$$

In the proofs that follow it will be useful to have both Definitions 6.1 and 6.2 and the third description of Lemma 6.3 to characterize $\nabla F(V; X)$. One important observation about the directional derivative is that it preserves chain homotopy equivalences between functors.

Observation 6.4. If F and G are pointwise chain homotopic, then $\nabla F \simeq \nabla G$. This observation follows immediately from the fact that $\nabla F(V; X) = D_1 F(X \oplus -)(V)$, since $D_1 F \simeq D_1 G$ whenever $F \simeq G$. Thus, ∇ is an operation on chain homotopy equivalence classes of functors from $\mathcal{B}$ to $\mathcal{A}$.

The properties of the directional derivative of a functor are reminiscent of the properties of the directional derivative of a function in classical calculus of several real variables. The analogies between the classical notion of differentiation and this more categorical approach has been axiomatized in the notion of a *cartesian differential category*, introduced in [4] though presented here in the equivalent form of [6, Proposition 4.2]. The following theorem is a collection of seven lemmas which will be very useful to us in proving the higher order chain rule and analyzing the higher order directional derivatives in the next two sections. Collectively, these properties tell us that the directional derivative ∇ equips the category $\mathsf{HoAbCat_{Ch}}$ with a cartesian differential structure.

Theorem 6.5. *The category $\mathsf{AbCat_{Ch}}$, together with the directional derivative ∇, satisfies the following seven properties:*

(i) *For functors $F, G : \mathcal{B} \rightsquigarrow \mathcal{A}$,*

$$\nabla (F \oplus G)(V; X) \cong \nabla F(V; X) \oplus \nabla G(V; X).$$

(ii) *For any $F : \mathcal{B} \rightsquigarrow \mathcal{A}$, ∇F is linear in the direction variable: i.e., there is a chain homotopy equivalence*

$$\nabla F(V \oplus W; X) \simeq \nabla F(V; X) \oplus \nabla F(W; X)$$

and an isomorphism

$$\nabla F(0; X) \cong 0.$$

(iii) *The directional derivative of the identity functor* id: $\mathcal{A} \rightsquigarrow \mathcal{A}$ *is given by projection onto the direction variable:*

$$\nabla \mathrm{id}(V; X) \cong V.$$

(iv) *For a functor* $\langle F, G \rangle \colon \mathcal{C} \rightsquigarrow \mathcal{A} \times \mathcal{B}$ *with components* $F \colon \mathcal{C} \rightsquigarrow \mathcal{A}$ *and* $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$,

$$\nabla \langle F, G \rangle (V; X) \cong \langle \nabla F(V; X), \nabla G(V; X) \rangle.$$

(v) *Whenever* $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ *and* $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ *are composable, there is a chain homotopy equivalence*

$$\nabla(F \circ G)(V; X) \simeq \nabla F(\nabla G(V; X); G(X)).$$

(vi) *For any* $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$, *there is an isomorphism*

$$\nabla(\nabla F)((Z; 0); (0; X)) \cong \nabla F(Z; X).$$

(vii) *For any* $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$, *there is a chain homotopy equivalence*

$$\nabla(\nabla F)((Z; W); (V; X)) \simeq \nabla(\nabla F)((Z; V); (W; X)).$$

Before proving Theorem 6.5, we summarize its statements. A *cartesian differential category* is a category satisfying the following axioms. Firstly, it is *left additive*, meaning that each hom-set is a commutative monoid and pre-composition with any map is a monoid homomorphism. A morphism in a left-additive category is then called *additive* if post-composing with it is also a monoid homomorphism. Secondly, the category has finite products so that the diagonal and product-projection maps are additive, and moreover the product of two additive maps is again additive. Finally, a cartesian differential category is equipped with an operation ∇ that takes a morphism $f \colon b \to a$ and produces a morphism $\nabla f \colon b \times b \to a$ satisfying the properties (i)–(vii) of Theorem 6.5. See [4], [6].

The usual cartesian product of two categories that are abelian is again abelian (see, for example, [17, Exercise VIII.3.2]). One can verify that this construction defines a product in the category $\mathsf{HoAbCat_{Ch}}$. The hom-sets $[\mathcal{B}, \mathcal{A}]$ in this category are (large) commutative monoids, with addition of chain homotopy equivalence classes of functors defined pointwise using the direct sum in $\mathcal{A}$. Pre-composition manifestly respects this structure, as does post-composition with diagonal and product-projection maps. It is similarly easy to verify that the product of two additive maps is additive. Thus Theorem 6.5 can be summarized by the following corollary:

Corollary 6.6. *The homotopy category* $\mathsf{HoAbCat_{Ch}}$ *is a (large) cartesian differential category.*

The proof of Theorem 6.5 will occupy the remainder of this section.

Proof of Theorem 6.5(i). For functors $F, G \colon \mathcal{B} \rightsquigarrow \mathcal{A}$,

$$\ker\left((F \oplus G)(X \oplus V) \to (F \oplus G)(X)\right) \cong \ker\left(F(X \oplus V) \to F(X)\right) \oplus \ker\left(G(X \oplus V) \to G(X)\right).$$

The isomorphism

$$\nabla(F \oplus G)(V; X) \cong \nabla F(V; X) \oplus \nabla G(V; X)$$

now follows from linearity of D_1 in the form of Lemma 5.6(ii). $\quad\square$

Proof of Theorem 6.5(ii). The chain homotopy equivalence

$$\nabla F(V \oplus W; X) \simeq \nabla F(V; X) \oplus \nabla F(W; X)$$

and the isomorphism

$$\nabla F(0; X) \simeq 0$$

are immediate from either definition of ∇F and Lemma 5.6(i). $\quad\square$

Proof of Theorem 6.5(iii). Evidently $\ker(X \oplus V \to X) \cong V$. The linearization of the identity functor is the identity, so by Definition 6.2

$$\nabla \mathrm{id}(V; X) = D_1^V \ker(X \oplus V \to X) \cong D_1^V \mathrm{id}V \simeq V. \qquad \square$$

Proof of Theorem 6.5(iv). Consider a functor $\langle F, G \rangle \colon \mathcal{C} \rightsquigarrow \mathcal{A} \times \mathcal{B}$. By Definition 6.2

$$\nabla \langle F, G \rangle(V; X) := D_1^V \ker(\langle F, G \rangle(X \oplus V) \to \langle F, G \rangle(X))$$
$$\cong D_1^V \ker(\langle F(X \oplus V), G(X \oplus V) \rangle \to \langle F(X), G(X) \rangle)$$

Since limits in product categories are computed coordinate-wise:

$$\cong D_1^V \langle \ker(F(X \oplus V) \to F(X)), \ker(G(X \oplus V) \to G(X)) \rangle$$

By Lemma 5.18:

$$\cong \langle D_1^V \ker(F(X \oplus V) \to F(X)), D_1^V \ker(G(X \oplus V) \to G(X)) \rangle$$
$$=: \langle \nabla F(V; X), \nabla G(V; X) \rangle. \qquad \square$$

Proof of Theorem 6.5(v). Whenever $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ and $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ are composable and G is reduced, [12, Proposition 5.6] proves a quasi-isomorphism

$$\nabla(F \circ G)(V; X) \simeq \nabla F(\nabla G(V; X); G(X)).$$

We will upgrade this to a chain homotopy equivalence and drop the hypotheses on G.

To make use of Lemma 6.3 for $\nabla(F \circ G)$, we first compute the second cross effect of $F \circ G$, which is defined by

$$(F \circ G)(X \oplus V) \cong (F \circ G)(0) \oplus \mathrm{cr}_1(F \circ G)(X) \oplus \mathrm{cr}_1(F \circ G)(V) \oplus \mathrm{cr}_2(F \circ G)(X, V)$$
$$\cong (F \circ G)(X) \oplus \mathrm{cr}_1(F \circ G)(V) \oplus \mathrm{cr}_2(F \circ G)(X, V). \qquad (6.7)$$

This in turn is isomorphic to

$$F(G(X \oplus V)) \cong F(G(X) \oplus \mathrm{cr}_1 G(V) \oplus \mathrm{cr}_2 G(X, V))$$
$$\cong F(0) \oplus \mathrm{cr}_1 F(G(X)) \oplus \mathrm{cr}_1 F(\mathrm{cr}_1 G(V)) \oplus \mathrm{cr}_1 F(\mathrm{cr}_2 G(X, V))$$
$$\oplus \mathrm{cr}_2 F(G(X), \mathrm{cr}_1 G(V)) \oplus \mathrm{cr}_2 F(G(X), \mathrm{cr}_2 G(X, V))$$
$$\oplus \mathrm{cr}_2 F(\mathrm{cr}_1 G(V), \mathrm{cr}_2 G(X, V)) \oplus \mathrm{cr}_3 F(G(X), \mathrm{cr}_1 G(V), \mathrm{cr}_2 G(X, V))$$

The first two terms sum to $F(G(X))$, so we have identified that piece of (6.7). Lemma 5.9 tells us that $\mathrm{cr}_1(F \circ G)(V) \cong \mathrm{cr}_1 F(\mathrm{cr}_1 G(V)) \oplus \mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G(V))$. To separate this piece, we substitute $G(X) \cong G(0) \oplus \mathrm{cr}_1 G(X)$ into the fifth summand and then re-express that second cross effect as a sum of three terms:

$$\cong F(G(X)) \oplus \mathrm{cr}_1 F(\mathrm{cr}_1 G(V)) \oplus \mathrm{cr}_1 F(\mathrm{cr}_2 G(X, V))$$
$$\oplus \mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G(V)) \oplus \mathrm{cr}_2 F(\mathrm{cr}_1 G(X), \mathrm{cr}_1 G(V))$$
$$\oplus \mathrm{cr}_3 (G(0), \mathrm{cr}_1 G(X), \mathrm{cr}_1 G(V)) \oplus \mathrm{cr}_2 F(G(X), \mathrm{cr}_2 G(X, V))$$
$$\oplus \mathrm{cr}_2 F(\mathrm{cr}_1 G(V), \mathrm{cr}_2 G(X, V)) \oplus \mathrm{cr}_3 F(G(X), \mathrm{cr}_1 G(V), \mathrm{cr}_2 G(X, V))$$

Taking complements to cancel these terms in our two formulas for $F(G(X \oplus V))$, we conclude that

$$\mathrm{cr}_2(F \circ G)(X, V) \cong \mathrm{cr}_1 F(\mathrm{cr}_2 G(X, V)) \oplus \mathrm{cr}_2 F(\mathrm{cr}_1 G(X), \mathrm{cr}_1 G(V))$$
$$\oplus \mathrm{cr}_3 (G(0), \mathrm{cr}_1 G(X), \mathrm{cr}_1 G(V)) \oplus \mathrm{cr}_2 F(G(X), \mathrm{cr}_2 G(X, V))$$
$$\oplus \mathrm{cr}_2 F(\mathrm{cr}_1 G(V), \mathrm{cr}_2 G(X, V)) \oplus \mathrm{cr}_3 F(G(X), \mathrm{cr}_1 G(V), \mathrm{cr}_2 G(X, V))$$

Now each of these terms is multi-reduced, so upon applying D_1^V, Corollary 5.13 implies that any term in which V appears more than once will vanish. Hence there is a chain homotopy equivalence:

$$D_1^V \mathrm{cr}_2(F \circ G)(X, V) \simeq D_1^V \mathrm{cr}_1 F(\mathrm{cr}_2 G(X, V)) \oplus D_1^V \mathrm{cr}_2 F(\mathrm{cr}_1 G(X), \mathrm{cr}_1 G(V))$$
$$\oplus D_1^V \mathrm{cr}_3 (G(0), \mathrm{cr}_1 G(X), \mathrm{cr}_1 G(V)) \oplus D_1^V \mathrm{cr}_2 F(G(X), \mathrm{cr}_2 G(X, V))$$

By Proposition 5.10 and this computation, this proves that there is a chain homotopy equivalence

$$\nabla(F \circ G)(V; X) \simeq D_1(F \circ G)(V) \oplus D_1^V \mathrm{cr}_2(F \circ G)(V, X)$$
$$\simeq D_1 F(D_1 G(V)) \oplus D_1^V \mathrm{cr}_2 F(G(0), \mathrm{cr}_1 G(V))$$
$$\oplus D_1^V \mathrm{cr}_1 F(\mathrm{cr}_2 G(X, V)) \oplus D_1^V \mathrm{cr}_2 F(\mathrm{cr}_1 G(X), \mathrm{cr}_1 G(V))$$
$$\oplus D_1^V \mathrm{cr}_3 (G(0), \mathrm{cr}_1 G(X), \mathrm{cr}_1 G(V)) \oplus D_1^V \mathrm{cr}_2 F(G(X), \mathrm{cr}_2 G(X, V))$$

The second, fourth, and fifth terms sum to $D_1^V \mathrm{cr}_2 F(G(X), \mathrm{cr}_1 G(V))$, so this simplifies to:

$$\cong D_1 F(D_1 G(V)) \oplus D_1^V \mathrm{cr}_2 F(G(X), \mathrm{cr}_1 G(V))$$
$$\oplus D_1^V \mathrm{cr}_1 F(\mathrm{cr}_2 G(X, V)) \oplus D_1^V \mathrm{cr}_2 F(G(X), \mathrm{cr}_2 G(X, V))$$

Now let's work out the other side:

$$\nabla F(\nabla G(V; X); G(X)) \cong \nabla F(D_1 G(V) \oplus D_1^V \mathrm{cr}_2 G(X, V); G(X))$$

By Theorem 6.5(ii):

$$\cong \nabla F(D_1 G(V); G(X)) \oplus \nabla F(D_1^V \mathrm{cr}_2 G(X,V); G(X))$$

$$\cong D_1 F(D_1 G(V)) \oplus D_1^2 \mathrm{cr}_2 F(G(X), D_1 G(V))$$

$$\oplus D_1 F(D_1^V \mathrm{cr}_2 G(X,V)) \oplus D_1^2 \mathrm{cr}_2 F(G(X), D_1^V \mathrm{cr}_2 G(X,V))$$

Substituting $D_1 G \cong D_1 \mathrm{cr}_1 G$ in the second term and $D_1 F \cong D_1 \mathrm{cr}_1 F$ in the third term, we can apply Proposition 5.7 to each of the last three terms to obtain a chain homotopy equivalence:

$$\simeq D_1 F(D_1 G(V)) \oplus D_1^V \mathrm{cr}_2 F(G(X), \mathrm{cr}_1 G(V))$$

$$\oplus D_1^V \mathrm{cr}_1 F(\mathrm{cr}_2 G(X,V)) \oplus D_1^V \mathrm{cr}_2 F(G(X), \mathrm{cr}_2 G(X,V))$$

which agrees with the formula for $\nabla(F \circ G)(V; X)$ above. $\square$

To prove the remaining statements Theorem 6.5(vi) and (vii) we will provide a general formula for $\nabla(\nabla F)$ in terms of the cross-effects of F that will prove important for our investigation of higher order chain rules in the next section. Our formula will make use of the following characterization of $\mathrm{cr}_2 \nabla F$.

Lemma 6.8. *For any $F: \mathcal{B} \rightsquigarrow \mathcal{A}$ and objects $Z, W, V, X \in \mathcal{B}$*

$$\mathrm{cr}_2 \nabla F((Z;W), (V;X))$$
$$\cong D_1^Z \mathrm{cr}_2 F(Z,X) \oplus D_1^Z \mathrm{cr}_3 F(Z,W,X) \oplus D_1^V \mathrm{cr}_2 F(V,W) \oplus D_1^V \mathrm{cr}_3 F(V,W,X).$$

Proof. Note that $\nabla F(0;0) \cong 0$, which implies that $\mathrm{cr}_1 \nabla F \cong \nabla F$. Thus $\mathrm{cr}_2 F$ is defined as the direct sum complement:

$$\nabla F(Z \oplus V; W \oplus X) \cong \nabla F(Z;W) \oplus \nabla F(V;X) \oplus \mathrm{cr}_2 \nabla F((Z;W), (V;X)).$$

By Theorem 6.5(ii)

$$\nabla F(Z \oplus V; W \oplus X) \cong \nabla F(Z; W \oplus X) \oplus \nabla F(V; W \oplus X)$$
$$= D_1^Z F(Z) \oplus D_1^Z \mathrm{cr}_2 F(Z, W \oplus X) \oplus D_1^V F(V) \oplus D_1^V \mathrm{cr}_2 F(V, W \oplus X)$$

By linearity of D_1 (Lemma 5.8) and the formula defining the third cross effects (Definition 2.1):

$$\cong D_1^Z F(Z) \oplus D_1^Z \mathrm{cr}_2 F(Z,W) \oplus D_1^Z \mathrm{cr}_2 F(Z,X) \oplus D_1^Z \mathrm{cr}_3 F(Z,W,X)$$
$$\oplus D_1^V F(V) \oplus D_1^V \mathrm{cr}_2 F(V,X) \oplus D_1^V \mathrm{cr}_2 F(V,W) \oplus D_1^V \mathrm{cr}_3 F(V,W,X)$$
$$\cong \nabla F(Z;W) \oplus \nabla F(V;X) \oplus D_1^Z \mathrm{cr}_2 F(Z,X) \oplus$$
$$\oplus D_1^Z \mathrm{cr}_3 F(Z,W,X) \oplus D_1^V \mathrm{cr}_2 F(V,W) \oplus D_1^V \mathrm{cr}_3 F(V,W,X)$$

and the claim follows by cancelling the isomorphic complements. $\square$

With Lemma 6.8 in hand, we can now compute $\nabla(\nabla F)((Z;W); (V;X))$. The next lemma shows that $\nabla(\nabla F)((Z;W); (V;X))$ is chain homotopy equivalent to $\nabla F(Z;X)$ plus an error term, which will be interpreted in the next section; see Proposition 7.5.

Lemma 6.9. *For any $F: \mathcal{B} \rightsquigarrow \mathcal{A}$, there is a chain homotopy equivalence*

$$\nabla(\nabla F)((Z;W); (V;X))$$
$$\simeq D_1^Z F(Z) \oplus D_1^Z \mathrm{cr}_2 F(Z,X) \oplus D_1^W D_1^V \mathrm{cr}_2 F(V,W) \oplus D_1^W D_1^V \mathrm{cr}_3 F(V,W,X).$$

Proof. By Lemma 6.3

$$\nabla(\nabla F)((Z;W);(V;X))$$
$$\cong D_1^{Z\times W}\nabla F(Z;W)\oplus D_1^{Z\times W}\mathrm{cr}_2\nabla F((Z;W);(V;X))$$

By linearity of D_1 and Lemma 6.8

$$\cong D_1^{Z\times W}(D_1^Z F(Z)\oplus D_1^Z\mathrm{cr}_2 F(Z,W)\oplus D_1^Z\mathrm{cr}_2 F(Z,X)$$
$$\oplus D_1^Z\mathrm{cr}_3 F(Z,W,X)\oplus D_1^V\mathrm{cr}_2 F(V,W)\oplus D_1^V\mathrm{cr}_3 F(V,W,X))$$

The functors $D_1^Z\mathrm{cr}_2 F(Z,W)$ and $D_1^Z\mathrm{cr}_3 F(Z,W,X)$ are strictly multi-reduced in the Z and W variables, so after distributing the linearization $D_1^{Z\times W}$ over the direct sum, Lemma 5.12 implies that these two terms are contractible. Thus, the previous term is chain homotopy equivalent to:

$$D_1^{Z\times W}(D_1^Z F(Z)\oplus D_1^Z\mathrm{cr}_2 F(Z,X)\oplus D_1^V\mathrm{cr}_2 F(V,W)\oplus D_1^V\mathrm{cr}_3 F(V,W,X))$$

and by Corollary 5.17 this is isomorphic to

$$D_1^Z F(Z)\oplus D_1^Z\mathrm{cr}_2 F(Z,X)\oplus D_1^W D_1^V\mathrm{cr}_2 F(V,W)\oplus D_1^W D_1^V\mathrm{cr}_3 F(V,W,X). \qquad \square$$

The proofs of Theorem 6.5 (vi) and (vii) follow immediately from Lemma 6.9. However, using Lemma 6.9 will only provide a chain homotopy equivalence between the desired term. When $V=W=0$, we can strengthen this to an isomorphism by reexamining the proof of Lemma 6.9.

Proof of Theorem 6.5(vi). From the proof of Lemma 6.9, recall that

$$\nabla(\nabla F)((Z;W);(V;X))$$
$$\cong D_1^{Z\times W}(D_1^Z F(Z)\oplus D_1^Z\mathrm{cr}_2 F(Z,W)\oplus D_1^Z\mathrm{cr}_2 F(Z,X)$$
$$\oplus D_1^Z\mathrm{cr}_3 F(Z,W,X)\oplus D_1^V\mathrm{cr}_2 F(V,W)\oplus D_1^V\mathrm{cr}_3 F(V,W,X)).$$

When $V=W=0$, Lemma 5.6(i) implies that the second term and the last three terms are 0, so

$$\nabla(\nabla F)((Z;0);(V;X))\simeq D_1^Z F(Z)\oplus D_1^Z\mathrm{cr}_2 F(Z,X)\cong\nabla F(Z;X). \qquad \square$$

The proof of 6.5(vii) is now an immediate consequence of Lemma 6.9.

Proof of Theorem 6.5(vii). By inspecting the formula of Lemma 6.9, Lemma 5.15 and the symmetry of the cross effects implies that there is a chain homotopy equivalence

$$\nabla(\nabla F)((Z;W);(V;X))\simeq\nabla(\nabla F)((Z;V);(W;X)). \qquad \square$$

7. Higher order directional derivatives and the Faà di Bruno formula

The directional derivative $\nabla F\colon \mathcal{B}\times\mathcal{B}\rightsquigarrow\mathcal{A}$ of a functor $F\colon\mathcal{B}\rightsquigarrow\mathcal{A}$ can be iterated to a higher order directional derivative in several different ways. The naïve iteration $\nabla^{\times 2}F:=\nabla(\nabla F)\colon\mathcal{B}^4\rightsquigarrow\mathcal{A}$ has some redundancy stemming from the fact that ∇F is linear in its first variable; see Theorem 6.5(ii) and Lemma 6.9. This problem compounds for higher iterations: in some sense $\nabla^{\times n}F\colon\mathcal{B}^{2n}\rightsquigarrow\mathcal{A}$ has $n-1$ too many variables.

In this section, we'll consider two "more efficient" formulations of the higher directional derivative that bear a very interesting relationship to one another captured by Theorem 7.7. The first of these is a functor $\nabla^n F \colon \mathcal{B}^{n+1} \rightsquigarrow \mathcal{A}$ introduced by Johnson–McCarthy, which we refer to as the nth *iterated partial directional derivative*. The second of these is a new functor $\Delta_n \colon \mathcal{B}^{n+1} \rightsquigarrow \mathcal{A}$ modeled after the higher directional derivatives of Huang, Marcantognini, and Young, which we call the nth *higher order directional derivative*.

Definition 7.1 *([12, Definition 5.8]).* Let $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ be a functor. The nth *iterated partial directional derivative* of F is defined recursively by

$$\nabla^0 F(X) := F(X),$$
$$\nabla^1 F(V; X) := \nabla F(V; X),$$

and for $n \geq 2$ and objects $V_1, \ldots, V_n$ and X in $\mathcal{B}$, the nth iterated partial directional derivative of F at X in the directions $V_1, \ldots, V_n$ is

$$\nabla^n F(V_n, \ldots, V_1; X) := \nabla \left(\nabla^{n-1} F(V_{n-1}, V_{n-2}, \ldots, V_1; -) \right) (V_n; X).$$

Johnson–McCarthy prove that the nth iterated partial directional derivative can be understood as a multilinearization of the nth and $(n+1)$st cross effects. Recall that for a functor of n-variables, $D_1^{(n)}$ indicates the linearization in each of its first n variables sequentially. By Lemma 5.15, the order in which these linearizations are performed does not matter.

Proposition 7.2. *[12] For any $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and variables $V_n, \ldots, V_1, X$*

(i) There is a natural isomorphism

$$\nabla^n F(V_n, \ldots, V_1; X) \cong D^{(n)} \left(\mathrm{cr}_{n+1} F(V_n, \ldots, V_1, X) \oplus \mathrm{cr}_n F(V_n, \ldots, V_1) \right).$$

(ii) For any permutation $\sigma \in \Sigma_n$, there is a natural isomorphism

$$\nabla^n F(V_n, \ldots, V_1; X) \cong \nabla^n F(V_{\sigma(n)}, \ldots, V_{\sigma(1)}; X).$$

Proof. The first statement is [12, Proposition 5.9]. The second statement, appearing as [12, Corollary 5.10], follows immediately from the symmetry of the cross effects and Lemma 5.15. □

On the other hand, the nth higher order directional derivative of [10] suggests that the following definition should be considered.

Definition 7.3. For a functor $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ the nth *higher order directional derivatives* of F are defined recursively by

$$\Delta_0 F(X) := F(X),$$
$$\Delta_1 F(V; X) := \nabla F(V; X),$$

and for $n \geq 2$ and objects $V_1, \ldots, V_n$ and X in $\mathcal{B}$, the nth higher order directional derivative of F at X in the directions $V_1, \ldots, V_n$ is defined to be

$$\Delta_n F(V_n, \ldots, V_1; X) := \nabla(\Delta_{n-1} F) \left((V_n, \ldots, V_2; V_1); (V_{n-1}, \ldots, V_1; X) \right).$$

Note that $\nabla^1 F = \nabla(\nabla^0 F)$ and $\Delta_1 F = \nabla(\Delta_0 F)$, so the recursive formulas for ∇^n and Δ_n given in Definitions 7.1 and 7.3 hold for all $n = 1$ as well.

Remark 7.4. Both the iterated partial directional derivative ∇^n and the higher order directional derivative Δ_n can be thought of as an iteration of the directional derivative ∇ where a new direction is introduced at each instance of differentiation. The difference between these two is that in the first case, the nth iterated partial directional derivative is considered as a functor of a single variable with its first $(n-1)$ "directions" (that is, $V_{n-1}, \ldots, V_1$) fixed. The new single variable functor of X is then differentiated in a new direction V_n. In the second case, the n-variable functor $\Delta_{n-1} F$ is differentiated in the "direction" $(V_n, \ldots, V_2, V_1)$ and at the "point" $(V_{n-1}, \ldots, V_1, X)$. This "point" should be thought of as having recorded the history of how the functor has been differentiated in previous iterations, that is, in the directions $V_{n-1}, \ldots, V_1$, respectively. The effect of this is that the functor Δ_n can be understood as a restriction of $\nabla(\Delta_{n-1})$ along a diagonal functor

$$L_n : \mathcal{B}^{n+1} \to \mathcal{B}^n \times \mathcal{B}^n$$

defined by $L_n(V_n, \ldots, V_1, X) = ((V_n, \ldots, V_1), (V_{n-1}, \ldots, V_1, X))$. Thus, as an operation on functors from $\mathcal{B}$ to $\mathcal{A}$, we have

$$\Delta_n := L_n^* \circ \nabla \circ \Delta_{n-1}$$

where L_n^* is the functor which precomposes with L_n. We can write this diagrammatically as

$$\Delta_n := [\mathcal{B}, \mathcal{A}] \xrightarrow{\Delta_{n-1}} [\mathcal{B}^n, \mathcal{A}] \xrightarrow{\nabla} [\mathcal{B}^n \times \mathcal{B}^n, \mathcal{A}] \xrightarrow{L_n^*} [\mathcal{B}^{n+1}, \mathcal{A}]$$

where $[\mathcal{B}, \mathcal{A}]$ denotes the hom-set in $\mathsf{HoAbCat_{Ch}}$.

By definition, $\nabla^0 F = F = \Delta_0 F$ and $\nabla^1 F = \nabla F = \Delta_1 F$. However, the two possible ways of iterating the directional derivative diverge for $n \geq 2$. Their precise relationship for $n = 2$ is explained in the following proposition.

Proposition 7.5. *For a functor $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and objects (Z, W, V, X) in $\mathcal{B}^4$, there is a chain homotopy equivalence*

$$\nabla(\nabla F)\left((Z; W); (V; X)\right) \simeq \nabla F(Z; X) \oplus \nabla^2 F(W, V; X).$$

Specializing to the object (V_2, V_1, V_1, X) in the image of L_2, we have a chain homotopy equivalence

$$\Delta_2 F(V_2, V_1; X) \simeq \nabla F(V_2; X) \oplus \nabla^2 F(V_1, V_1; X).$$

Proof. By Lemma 6.9, we have a chain homotopy equivalence

$$\nabla(\nabla F)((Z; W); (V; X))$$
$$\simeq D_1^Z F(Z) \oplus D_1^Z \mathrm{cr}_2 F(Z, X) \oplus D_1^W D_1^V \mathrm{cr}_2 F(V, W) \oplus D_1^W D_1^V \mathrm{cr}_3 F(V, W, X).$$

By the characterization of ∇F in Lemma 6.3, $D_1^Z F(Z) \oplus D_1^Z \mathrm{cr}_2 F(Z, X) \cong \nabla F(Z; X)$. We need only show that

$$D_1^W D_1^V \mathrm{cr}_2 F(V, W) \oplus D_1^W D_1^V \mathrm{cr}_3 F(V, W, X) \cong \nabla^2 F(V, W; X),$$

but this is precisely the natural isomorphism of Proposition 7.2(i). $\square$

Proposition 7.5 can be generalized to explain the relationship between Δ_n and ∇^n for all $n \geq 0$. This relationship is surprisingly rich. In particular, its formulation is reminiscent of the Faà di Bruno formula for the nth derivative of $f \circ g$ for composable *functions* f and g. The relationship is stated precisely in the next theorem.

For a partition $\pi = \{S_1, \ldots, S_k\}$ of the set $\{1, \ldots, n\}$, we use the following notation:

$$\nabla^\pi F(V_n, \ldots, V_1; X) := \nabla^{|\pi|} F(V_{|S_1|}, \ldots, V_{|S_k|}; X)$$

where $|S|$ denotes the cardinality of the set S. Note that by Proposition 7.2(ii), the order in which we write the $V_{|S_i|}$'s doesn't matter.

Example 7.6. For the partition $\pi = \{\{1\}, \{2\}\}$, $\nabla^\pi F(V_1, V_2; X) := \nabla^2 F(V_1, V_1; X)$, while for $\pi = \{\{1, 2\}\}$, $\nabla^\pi F(V_1, V_2; X) := \nabla F(V_2; X)$. Not coincidentally, these are the two terms appearing in the formula for $\Delta_2 F(V_2, V_1; X)$ proven in Proposition 7.5.

We are now prepared to state the general result:

Theorem 7.7. *For any $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and objects $V_n, \ldots, V_1, X$, there is a chain homotopy equivalence*

$$\Delta_n F(V_n, \ldots, V_1; X) \simeq \bigoplus_{\pi = \{S_1, \ldots, S_k\}} \nabla^\pi F(V_n, \ldots, V_1; X)$$

where the sum is taken over all partitions π of the set $\{1, 2, \ldots, n\}$.

Theorem 7.7 is a consequence of the following lemma.

Lemma 7.8. *For a partition $\pi = \{S_1, \ldots, S_k\}$ of $\{1, \ldots, n\}$, there is a chain homotopy equivalence*

$$\nabla(\nabla^\pi F)((V_{n+1}, \ldots; V_1); (V_n, \ldots, V_1; X)) \simeq \tag{7.9}$$

$$\nabla^{|\pi|+1} F(V_{|S_1|}, \ldots, V_{|S_k|}, V_1; X) \oplus \bigoplus_{i=1}^{k} \nabla^{|\pi|} F(V_{|S_1|}, \ldots, V_{|S_{i-1}|}, V_{|S_i|+1}, V_{|S_{i+1}|}, \ldots, V_{|S_k|}; X).$$

This lemma permits an expeditious proof of Theorem 7.7.

Proof of Theorem 7.7. The right-hand side of (7.9) can be interpreted as the sum of terms of the form $\nabla^\sigma F(V_{n+1}, \ldots, V_1; X)$ taken over all partitions σ of $\{1, 2, \ldots, n+1\}$ that are obtained from the partition $\pi = \{S_1, \ldots, S_k\}$ of $\{1, \ldots, n\}$ by either adding $\{n+1\}$ as a separate set (contributing the first summand) or adding $n+1$ to one of the S_i's (contributing the remaining k summands). Since all partitions of $\{1, \ldots, n+1\}$ can be obtained in this way from partitions of $\{1, \ldots, n\}$, Theorem 7.7 is proven by a simple induction with Proposition 7.5 as the base case, as explained by Example 7.6, and Lemma 7.8 as the inductive step. $\quad\square$

Next we prove the lemma.

Proof of Lemma 7.8. Let $\pi = \{S_1, \ldots, S_k\}$ be a partition of $\{1, \ldots, n\}$. By Proposition 7.2(i)

$$\nabla^\pi F(V_n, \ldots, V_1; X) := \nabla^k F(V_{|S_1|}, \ldots, V_{|S_k|}; X)$$

$$\cong D_1^{(k)} \mathrm{cr}_k F(V_{|S_1|}, \ldots, V_{|S_k|}) \oplus D_1^{(k)} \mathrm{cr}_{k+1} F(V_{|S_1|}, \ldots, V_{|S_k|}, X).$$

By Definition 6.2, $\nabla(\nabla^\pi F)((\overline{V}_{n+1}, \ldots; \overline{V}_1); (V_n, \ldots, V_1; X))$ is equal to

$$D_1^{\overline{V}_{n+1}\times\cdots\times\overline{V}_1}\ker(\nabla^\pi F(\overline{V}_{n+1}\oplus V_n,\ldots;\overline{V}_1\oplus X)\to\nabla^\pi F(V_n,\ldots,V_1;X)), \tag{7.10}$$

where we have added new variables $\overline{V}_i$ to keep track of where the linearizations are being applied. Recalling Convention 5.11, the notation $D_1^{\overline{V}_{n+1}\times\cdots\times\overline{V}_1}$ indicates the linearization $D_1 H$ of a functor $H\colon\mathcal{B}^{n+1}\to\mathcal{A}$, where H is the functor

$$\ker(\nabla^\pi F(-\oplus V_n,\ldots;-\oplus X)\to\nabla^\pi F(V_n,\ldots,V_1;X))$$

and the V_i's and X are fixed values of $\mathcal{B}$. Because H is a functor of several variables, this linearization involves the simultaneous linearization of the variables $\overline{V}_1,\ldots,\overline{V}_{n+1}$.

Our strategy will be to use Proposition 7.2 to rewrite (7.10) in terms of multilinearizations of cross effects, use Lemma 5.12 and Corollary 5.13 to eliminate many of the summands after taking the kernel and applying $D_1^{(V_{n+1},\ldots,V_1)}$, and use Proposition 7.2(i) to verify that what is left is equivalent to (7.9).

By Proposition 7.2(i), (7.10) is isomorphic to $D_1^{\overline{V}_{n+1}\times\cdots\times\overline{V}_1}$ applied to the kernel of the map

$$D_1^{(k)}\mathrm{cr}_k F(\overline{V}_{|S_1|+1}\oplus V_{|S_1|},\ldots,\overline{V}_{|S_k|+1}\oplus V_{|S_k|})\oplus D_1^{(k)}\mathrm{cr}_{k+1}F(\overline{V}_{|S_1|+1}\oplus V_{|S_1|},\ldots,\overline{V}_1\oplus X)$$

$$\downarrow$$

$$D_1^{(k)}\mathrm{cr}_k F(V_{|S_1|},\ldots,V_{|S_k|})\oplus D_1^{(k)}\mathrm{cr}_{k+1}F(V_{|S_1|},\ldots,V_{|S_k|},X)$$

Since the domain of this map is linear in its first k variables, it is chain homotopy equivalent to an expansion of this expression as illustrated for instance by:

$$D_1^{(k)}\mathrm{cr}_k F(\overline{V}_{|S_1|+1}\oplus V_{|S_1|},\ldots,\overline{V}_{|S_k|+1}\oplus V_{|S_k|})$$
$$\simeq D_1^{(k)}\mathrm{cr}_k F(\overline{V}_{|S_1|+1},\ldots,\overline{V}_{|S_k|+1}\oplus V_{|S_k|})\oplus D_1^{(k)}\mathrm{cr}_k F(V_{|S_1|},\ldots,\overline{V}_{|S_k|+1}\oplus V_{|S_k|})$$
$$\simeq D_1^{(k)}\mathrm{cr}_k F(\overline{V}_{|S_1|+1},\overline{V}_{|S_2|+1},\ldots,\overline{V}_{|S_k|+1}\oplus V_{|S_k|})\oplus D_1^{(k)}\mathrm{cr}_k F(\overline{V}_{|S_1|+1},V_{|S_2|},\ldots,\overline{V}_{|S_k|+1}\oplus V_{|S_k|})$$
$$\oplus D_1^{(k)}\mathrm{cr}_k F(V_{|S_1|},\overline{V}_{|S_2|+1},\ldots,\overline{V}_{|S_k|+1}\oplus V_{|S_k|})\oplus D_1^{(k)}\mathrm{cr}_k F(V_{|S_1|},V_{|S_2|},\ldots,\overline{V}_{|S_k|+1}\oplus V_{|S_k|})$$
$$\simeq\cdots$$

but rather than carrying out these expansions completely, we note that because we are applying $D_1^{\overline{V}_{n+1}\times\cdots\times\overline{V}_1}$, Corollary 5.13 guarantees that any summand that has more than one $\overline{V}_i$ will be contractible. Hence we see that $D_1^{\overline{V}_{n+1}\times\cdots\times\overline{V}_1}(\nabla^\pi F)(\overline{V}_{n+1}\oplus V_n,\ldots;\overline{V}_1\oplus X)$ is chain homotopy equivalent to $D_1^{\overline{V}_{n+1}\times\cdots\times\overline{V}_1}$ of

$$\bigoplus_{i=1}^k D_1^{(k)}\mathrm{cr}_k F(V_{|S_1|},\ldots,V_{|S_{i-1}|},\overline{V}_{|S_i|+1},V_{|S_{i+1}|},\ldots,V_{|S_k|})$$

$$\oplus\bigoplus_{i=1}^k D_1^{(k)}\mathrm{cr}_{k+1}F(V_{|S_1|},\ldots,V_{|S_{i-1}|},\overline{V}_{|S_i|+1},V_{|S_{i+1}|},\ldots,V_{|S_k|},\overline{V}_1\oplus X)$$

$$\oplus D_1^{(k)}\mathrm{cr}_{k+1}F(V_{|S_1|},\ldots,V_{|S_k|},\overline{V}_1\oplus X)\oplus D_1^k\mathrm{cr}_k F(V_{|S_1|},\ldots,V_{|S_k|}).$$

That is, the only terms with a single variable of the form $\overline{V}_t$ appear in the kernel. A term of the form

$$D_1^{(k)}\mathrm{cr}_{k+1}F(V_{|S_1|},\ldots,V_{|S_{i-1}|},\overline{V}_{|S_i|+1},V_{|S_{i+1}|},\ldots,V_{|S_k|},\overline{V}_1\oplus X)$$

can be expanded as

$$D_1^{(k)} \mathrm{cr}_{k+1} F(V_{|S_1|}, \ldots, V_{|S_{i-1}|}, \overline{V}_{|S_i|+1}, V_{|S_{i+1}|}, \ldots, V_{|S_k|}, \overline{V}_1)$$
$$\oplus D_1^{(k)} \mathrm{cr}_{k+1} F(V_{|S_1|}, \ldots, V_{|S_{i-1}|}, \overline{V}_{|S_i|+1}, V_{|S_{i+1}|}, \ldots, V_{|S_k|}, X)$$
$$\oplus D_1^{(k)} \mathrm{cr}_{k+2} F(V_{|S_1|}, \ldots, V_{|S_{i-1}|}, \overline{V}_{|S_i|+1}, V_{|S_{i+1}|}, \ldots, V_{|S_k|}, \overline{V}_1, X).$$

But, again by Corollary 5.13, the first and third summands will be contractible after applying $D_1^{\overline{V}_{n+1} \times \cdots \times \overline{V}_1}$. Hence, after computing the kernel and applying $D_1^{\overline{V}_{n+1} \times \cdots \times \overline{V}_1}$, we see that

$$\nabla(\nabla^\pi F)((V_{n+1}, \ldots; V_1); (V_n, \ldots, V_1; X))$$

is chain homotopy equivalent to

$$\bigoplus_{i=1}^k D_1^{(k)} \mathrm{cr}_k F(V_{|S_1|}, \ldots, V_{|S_{i-1}|}, \overline{V}_{|S_i|+1}, V_{|S_{i+1}|}, \ldots, V_{|S_k|})$$
$$\oplus \bigoplus_{i=1}^k D_1^{(k)} \mathrm{cr}_{k+1} F(V_{|S_1|}, \ldots, V_{|S_{i-1}|}, \overline{V}_{|S_i|+1}, V_{|S_{i+1}|}, \ldots, V_{|S_k|}, X)$$
$$\oplus D_1^{(k+1)} \mathrm{cr}_{k+1} F(V_{|S_1|}, \ldots, V_{|S_k|}, \overline{V}_1) \oplus D_1^{(k)} \mathrm{cr}_{k+2} F(V_{|S_1|}, \ldots, V_{|S_k|}, \overline{V}_1, X).$$

By Proposition 7.2(i), this is isomorphic to:

$$\nabla^{|\pi|+1} F(V_{|S_1|}, \ldots, V_{|S_k|}, V_1; X) \oplus \bigoplus_{i=1}^k \nabla^{|\pi|} F(V_{|S_1|}, \ldots, V_{|S_{i-1}|}, V_{|S_i|+1}, V_{|S_{i+1}|}, \ldots, V_{|S_k|}; X) \qquad \square$$

8. The higher order chain rule

This brings us to the original motivation for this paper, the higher order directional derivative chain rule. Having established a relationship between the nth higher directional derivative and the nth iterated partial directional derivative, our goal is to establish a higher order chain rule in the style of [10] for the former of these two.

Theorem 8.1. *For a composable pair $F: \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G: \mathcal{C} \rightsquigarrow \mathcal{B}$ of functors of abelian categories, the nth directional derivative satisfies*

$$\Delta_n(F \circ G)(V_n, \ldots, V_1; X) \simeq \Delta_n F(\Delta_n G(V_n, \ldots, V_1; X), \ldots, \Delta_1 G(V_1; X); G(X)).$$

For the case $n = 2$, the third and fifth authors carried out a very careful computation verifying this theorem using only properties of linearization and cross effects. A short paper containing this computation is in preparation.

We will explain how this chain rule can be derived as a natural consequence of structures associated to cartesian differential categories. Cockett and Cruttwell introduce a more general notion of a *tangent category* [6] and show that any cartesian differential category gives an example. The key component of a tangent category structure is an endofunctor T. In our case, T is defined as follows.

Definition 8.2. There is a functor $T: \mathsf{HoAbCat}_{\mathsf{Ch}} \to \mathsf{HoAbCat}_{\mathsf{Ch}}$ defined on objects by $\mathcal{A} \mapsto \mathcal{A} \times \mathcal{A}$ and on morphisms $F: \mathcal{B} \rightsquigarrow \mathcal{A}$ by

$$TF := \langle \nabla F, F\pi_R \rangle: \mathcal{B} \times \mathcal{B} \rightsquigarrow \mathcal{A} \times \mathcal{A},$$

where $\pi_R \colon \mathcal{B} \times \mathcal{B} \to \mathcal{B}$ denotes the projection onto the second, or right-hand, factor. That is

$$TF(V, X) := \langle \nabla F(V; X), F(X) \rangle.$$

Importantly, Theorem 6.5(v) implies that T is *functorial* up to chain homotopy equivalence, as asserted in Definition 8.2.

Lemma 8.3. *For a composable pair $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ of functors of abelian categories, $T(F \circ G)$ is chain homotopy equivalent to $TF \circ TG$.*

Proof. We use Theorem 6.5(v) to show that T is a functor up to chain homotopy equivalence. By definition, $TF \circ TG$ is given by

$$\mathcal{C} \times \mathcal{C} \xrightsquigarrow{TG} \mathcal{B} \times \mathcal{B} \xrightsquigarrow{TF} \mathcal{A} \times \mathcal{A}$$
$$(V, X) \longmapsto \langle \nabla G(V; X), G(X) \rangle \longmapsto \langle \nabla F(\nabla G(V; X); G(X)), F(G(X)) \rangle$$

while $T(F \circ G)$ is given by

$$\mathcal{C} \times \mathcal{C} \xrightsquigarrow{T(F \circ G)} \mathcal{A} \times \mathcal{A}$$
$$(V, X) \longmapsto \langle \nabla (F \circ G)(V; X), F \circ G(X) \rangle$$

By Theorem 6.5(v) these outputs are chain homotopy equivalent:

$$\langle \nabla (F \circ G)(V; X), F \circ G(X) \rangle \simeq \langle \nabla F(\nabla G(V; X); G(X)), F(G(X)) \rangle. \qquad \square$$

Lemma 8.3 shows that T is an endofunctor of $\mathsf{HoAbCat}_{\mathsf{Ch}}$. As an immediate corollary, its iterates T^n are again functorial: that is, $T^n(F \circ G) \simeq T^n F \circ T^n G$. Since chain homotopy equivalences are respected by pre- and post-composition in $\mathsf{HoAbCat}_{\mathsf{Ch}}$, the composite functors

$$\mathcal{C}^{n+1} \xrightarrow{d_n^*} \mathcal{C}^{2^n} \underset{T^n F \circ T^n G}{\overset{T^n(F \circ G)}{\simeq}} \mathcal{A}^{2^n} \xrightarrow{\pi_L} \mathcal{A} \qquad (8.4)$$

are again chain homotopy equivalent, where π_L is the projection onto the leftmost component and d_n^* is a diagonal functor to be described below. We claim that the top composite of (8.4) is the functor $\Delta_n(F \circ G)$, while the bottom composite is the functor appearing on the right-hand side of the chain homotopy equivalence of Theorem 8.1. Once we make these identifications, which will be achieved by the following series of combinatorial lemmas, we will be able to conclude that Theorem 8.1 is a consequence of the functoriality of T.

The diagonal functor d_n^* is a special case of a "reindexing functor" in the sense of the following definition.

Definition 8.5. Write $\mathbf{n} = \{0, 1, \cdots, n-1\}$ for the set with n elements, and let $c \colon \mathbf{n} \to \mathbf{k}$ be any function. The product category $\mathcal{B}^n$ can be thought of as the category of functors from the discrete category $\mathbf{n}$ to $\mathcal{B}$, from which perspective pre-composition with $c \colon \mathbf{n} \to \mathbf{k}$ defines a *reindexing functor* $c^* \colon \mathcal{B}^k \to \mathcal{B}^n$. Explicitly, c^* is the functor

$$\langle \pi_{c(0)}, \ldots, \pi_{c(n-1)} \rangle \colon \mathcal{B}^k \to \mathcal{B}^n$$

whose components $\pi_i \colon \mathcal{B}^k \to \mathcal{B}$ project onto the ith factor.

Diagonal functors are reindexing functors which come from surjections of sets. Importantly, reindexing functors commute with the directional derivative ∇ in the sense of the following result.

Lemma 8.6. *Suppose* $K \colon \mathcal{B}^k \rightsquigarrow \mathcal{A}$ *factors as*

up to chain homotopy equivalence, where $c^* \colon \mathcal{B}^k \to \mathcal{B}^n$ *is the reindexing functor associated to some function* $c \colon n \to k$. *Then* ∇K *factors as*

up to chain homotopy equivalence

Proof. Let $V, X \in \mathcal{B}^k$. By [Theorem 6.5](v), $\nabla K(V; X) \simeq \nabla H(\nabla c^*(V; X); c^*(X))$. By [Theorem 6.5](iii) and (iv), $\nabla c^*(V; X) \simeq c^* V$, so $\nabla K(V; X) \simeq \nabla H(c^*(V); c^*(X))$. $\quad\square$

This observation allows us to make the relationship between $T^n F$ and $\nabla^{\times k} F$ explicit, which we describe in the next lemma by examining the coordinates of $T^n F \colon \mathcal{B}^{2^n} \to \mathcal{A}^{2^n}$ one at a time. For convenience, we identify the cardinal 2^n with the set $\mathcal{P}(\mathbf{n})$, the power set of $\mathbf{n}$, which has cardinality 2^n. Under this identification, $T^n F \colon \mathcal{B}^{2^n} \rightsquigarrow \mathcal{A}^{2^n}$ is a collection of functors $(T^n F)_S \colon \mathcal{B}^{2^n} \rightsquigarrow \mathcal{A}$, indexed by elements S of $\mathcal{P}(\mathbf{n})$, i.e., by subsets $S \subset \mathbf{n}$.

Lemma 8.7. *For any* $S \subset \mathbf{n}$, *there is a chain homotopy equivalence*

where $\iota_S \colon \mathcal{P}(S) \hookrightarrow \mathcal{P}(\mathbf{n})$ *is the inclusion.*

Proof. We prove this by induction on n. The case $n = 1$ is given by the definition of T, but we include it in detail here in order to begin to examine the definition of components of T in terms of subsets S of $\mathcal{P}(\mathbf{n})$. When $n = 1$, the power set $\mathcal{P}(\mathbf{1})$ contains the two elements $\emptyset$ and $\{0\}$. This corresponds to the two diagrams

This in turn corresponds directly to the definition of $TF(V, X) = \langle \nabla F(V; X), F(X) \rangle$ where the second coordinate is indexed by $\emptyset$ and the first coordinate is indexed by $\{0\}$. In particular, $TF_S = \nabla F$ when S contains the "top" element $0 \in \mathbf{1}$, and $TF_S = F$ when S does not contain 0.

For the inductive step, first consider those $S \subset \mathbf{n}$ that do not contain the top element. For these components, $(T^n F)_S = T((T^{n-1} F)_S) = (T^{n-1} F)_S \circ \pi_R$, and here $\pi_R \colon \mathcal{C}^{2^n} \to \mathcal{C}^{2^{n-1}}$ is the map defined by restricting along the map $\mathcal{P}(\mathbf{n}-\mathbf{1}) \subset \mathcal{P}(\mathbf{n})$ induced by the inclusion of the first $n-1$ elements. By the inductive hypothesis, $(T^{n-1} F)_S \simeq \nabla^{\times |S|} F \circ \iota_S^*$ and the claim follows.

Now consider the components indexed by a subset $S \subset \mathbf{n}$ that does contain the top element $n-1 \in \mathbf{n}$. Here $(T^n F)_S = \nabla((T^{n-1} F)_{S \backslash n-1})$. By the inductive hypothesis, $(T^{n-1} F)_{S \backslash n-1} \simeq \nabla^{\times |S|-1} F \circ \iota_{S \backslash n-1}^*$, so applying Lemma 8.6

$$(T^n F)_S \simeq \nabla(\nabla^{\times |S|-1} F \circ \iota_{S \backslash n-1}^*) \simeq \nabla^{\times |S|} F \circ (\iota_{S \backslash n-1}^* \times \iota_{S \backslash n-1}^*).$$

The proof is completed by the observation that the map

$$\iota_{S \backslash n-1} \times \iota_{S \backslash n-1} \colon \mathcal{P}(S \backslash n-1) \times \mathcal{P}(S \backslash n-1) \hookrightarrow \mathcal{P}(\mathbf{n}) \times \mathcal{P}(\mathbf{n})$$

coincides with the map $\iota_S \colon \mathcal{P}(S) \hookrightarrow \mathcal{P}(\mathbf{n})$. $\square$

We now construct the function $d_n \colon \mathcal{P}(\mathbf{n}) \to \mathbf{n+1}$ inducing the diagonal functor of (8.4). The function $d_n \colon \mathcal{P}(\mathbf{n}) \to \mathbf{n+1}$ is defined by $d_n(S) = |S|$, where $|S|$ denotes the cardinality of S. Equivalently, if we think of the elements of 2^n as n-tuples of binary digits (the vertices of a unit n-cube), there is a function $d_n \colon 2^n \to \mathbf{n+1}$ that counts the number of 1s. This can be thought of as an order-preserving projection onto the diagonal of the cube.

Recall Remark 7.4, which decomposes the functor $\Delta_n \colon [\mathcal{B}, \mathcal{A}] \to [\mathcal{B}^{n+1}, \mathcal{A}]$ as a composite

$$\Delta_n := [\mathcal{B}, \mathcal{A}] \xrightarrow{\Delta_{n-1}} [\mathcal{B}^n, \mathcal{A}] \xrightarrow{\nabla} [\mathcal{B}^n \times \mathcal{B}^n, \mathcal{A}] \xrightarrow{L_n^*} [\mathcal{B}^{n+1}, \mathcal{A}],$$

first applying the directional derivative ∇ to Δ_{n-1} and then restricting the variables along the diagonal functor $L_n \colon \mathcal{B}^{n+1} \to \mathcal{B}^n \times \mathcal{B}^n$. This diagonal functor is also defined by precomposing with a certain function $a_n \colon \mathbf{2} \times \mathbf{n} \to \mathbf{n+1}$. Writing elements in the domain as ordered pairs of elements in $\mathbf{2} = \{0,1\}$ and $\mathbf{n} = \{0, \ldots, n-1\}$, the function a_n adds the two coordinates.

Lemma 8.8. *When $2^n \cong 2 \times 2^{n-1}$ is decomposed into its left-most coordinate paired with its remaining coordinates, the following diagram commutes*

$$
\begin{array}{ccc}
2^n & \xrightarrow{\;\;d_n\;\;} & \mathbf{n+1} \\
{\scriptstyle\cong}\big\downarrow & & \big\uparrow{\scriptstyle a_n} \\
2 \times 2^{n-1} \cong 2^{n-1} + 2^{n-1} & \xrightarrow{d_{n-1}+d_{n-1}} & \mathbf{n} + \mathbf{n} \cong \mathbf{2} \times \mathbf{n}
\end{array}
$$

Proof. For a binary n-tuple $(e_1, \ldots, e_n)$, by definition $d_n(e_1, \ldots, e_n) = \sum_{i=1}^n e_i$. The lower composite sends $(e_1, \ldots, e_n)$ first to $(e_1, (e_2, \ldots, e_n))$, then to $(e_1, \sum_{i=2}^n e_i)$ then to $e_1 + \sum_{i=2}^n e_i$. $\square$

Lemma 8.9. *The higher order directional derivative Δ_n is chain homotopy equivalent to the composite*

$$[\mathcal{B}, \mathcal{A}] \xrightarrow{\nabla^{\times n}} [\mathcal{B}^{2^n}, \mathcal{A}] \xrightarrow{(d_n^*)^*} [\mathcal{B}^{n+1}, \mathcal{A}]$$

which restricts $\nabla^{\times n}$ along the diagonal $d_n^* \colon \mathcal{B}^{n+1} \to \mathcal{B}^{2^n}$.

Proof. We prove this by induction on n. For the base case $n = 1$,

$$\nabla F \circ d_1^*(V_1; X) = \nabla F(V_1; X) = \Delta_1 F(V_1; X).$$

For $n > 1$, consider the diagram below:

$$
\begin{array}{ccccccc}
[\mathcal{B}, \mathcal{A}] & \xrightarrow{\nabla^{\times n-1}} & [\mathcal{B}^{2^{n-1}}, \mathcal{A}] & \xrightarrow{\nabla} & [\mathcal{B}^{2^n}, \mathcal{A}] & \xrightarrow{(d_n^*)^*} & [\mathcal{B}^{n+1}, \mathcal{A}] \\
& \searrow{\scriptstyle \Delta_{n-1}} & \downarrow{\scriptstyle (d_{n-1}^*)^*} & & \downarrow{\scriptstyle (d_{n-1}^* \times d_{n-1}^*)^*} & \nearrow{\scriptstyle (a_n^*)^*} & \\
& & [\mathcal{B}^n, \mathcal{A}] & \xrightarrow{\nabla} & [\mathcal{B}^{2n}, \mathcal{A}] & &
\end{array}
$$

Composition from $[\mathcal{B}, \mathcal{A}]$ to $[\mathcal{B}^{n+1}, \mathcal{A}]$ along the top is $(d_n^*)^* \circ \nabla^{\times n}$, whereas composition along the bottom is Δ_n, since $a_n^* = L_n$. The right triangle commutes by Lemma 8.8. The middle square commutes up to chain homotopy equivalence by Lemma 8.6. By the inductive hypothesis, the left triangle commutes up to chain homotopy equivalence, so the result follows by induction. $\square$

Write π_R (respectively π_L) for the projection from a product category onto the product formed by its k rightmost (respectively leftmost) variables, where the correct arity k is determined by the context. For example, the diagonal functor $L_n \colon \mathcal{B}^{n+1} \to \mathcal{B}^{2n}$ can be understood as the pairing of functors $\pi_L \colon \mathcal{B}^{n+1} \to \mathcal{B}^n$ and $\pi_R \colon \mathcal{B}^{n+1} \to \mathcal{B}^n$, i.e.,

$$
\begin{aligned}
L_n(V_n, \ldots, V_1, X) &= (\pi_L(V_n, \ldots, V_1, X), \pi_R(V_n, \ldots, V_1, X)) \\
&= ((V_n, \ldots, V_1), (V_{n-1}, \ldots, X)).
\end{aligned}
$$

Let

$$K_n \colon [\mathcal{C}, \mathcal{B}] \to [\mathcal{C}^{n+1}, \mathcal{B}^{n+1}]$$

be the functor defined by

$$K_n G := (\Delta_n G, \Delta_{n-1} G \circ \pi_R, \ldots, G \circ \pi_R).$$

Note that this expression appears as the argument for the $\Delta_n F$ appearing on the right-hand side of the chain homotopy equivalence of Theorem 8.1. The next lemma tells us that $K_n G \colon \mathcal{C}^{n+1} \rightsquigarrow \mathcal{B}^{n+1}$ and $T^n G \colon \mathcal{C}^{2^n} \rightsquigarrow \mathcal{B}^{2^n}$ are related by restricting the domain of $T^n G$ and the codomain of $K_n G$ along the diagonal functor d_n^*.

Lemma 8.10. *For $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$, $d_n^* \circ K_n G \simeq T^n G \circ d_n^*$ as functors $\mathcal{C}^{n+1} \to \mathcal{B}^{2^n}$.*

Proof. Again it is convenient to identify 2^n with $\mathcal{P}(n)$. It suffices to prove that these functors $\mathcal{C}^{n+1} \rightsquigarrow \mathcal{B}^{2^n}$ have chain homotopy equivalent components indexed by each subset $S \subset n$.

Lemma 8.7 tells us that $(T^n G)_S \simeq \nabla^{\times |S|} G \circ \iota_S^*$. As functors from $\mathcal{C}^{n+1}$ to $\mathcal{C}^{2^{|S|}}$, $\iota_S^* \circ d_n^* = d_{|S|}^* \circ \pi_R$. Hence,

$$(T^n G)_S \circ d_n^* \simeq \nabla^{\times |S|} G \circ \iota_S^* \circ d_n^* \simeq \nabla^{\times |S|} G \circ d_{|S|}^* \circ \pi_R \simeq \Delta_{|S|} G \circ \pi_R,$$

the last chain homotopy equivalence by Lemma 8.9.

On the other hand, immediately from the definition of $K_n G$,

$$(d_n^* \circ K_n G)_S = \Delta_{|S|} G \circ \pi_R. \qquad \square$$

Proof of Theorem 8.1. Consider the diagram below.

$$
\begin{array}{ccccc}
\mathcal{C}^{2^n} & \xrightarrow{\ T^n G\ } & \mathcal{B}^{2^n} & \xrightarrow{\ T^n F\ } & \mathcal{A}^{2^n} \\
{\scriptstyle d_n^*}\Big\uparrow & \simeq & {\scriptstyle d_n^*}\Big\uparrow & \simeq & \Big\downarrow{\scriptstyle \pi_L} \\
\mathcal{C}^{n+1} & \xrightarrow[K_n G]{} & \mathcal{B}^{n+1} & \xrightarrow[\Delta_n F]{} & \mathcal{A}
\end{array}
$$

The left and right squares commute up to chain homotopy equivalence by Lemmas 8.10 and 8.9, respectively. Using this diagram and Lemma 8.3 we see that

$$
\begin{aligned}
\Delta_n F(\Delta_n G, \Delta_{n-1}G \circ \pi_R, \ldots, G \circ \pi_R) &= \Delta_n F \circ K_n G \\
&\simeq \pi_L \circ T^n F \circ T^n G \circ d_n^* \\
&\simeq \pi_L \circ T^n(FG) \circ d_n^*.
\end{aligned}
$$

By Lemma 8.9, we have

$$
\pi_L \circ T^n(FG) \circ d_n^* \simeq \Delta_n(FG),
$$

which completes the proof. $\square$

9. Derivatives

So far, we have considered the functors $\nabla F(V; X)$ as analogs of the directional derivatives. In classical analysis, the derivative of a single variable function $f \colon \mathbb{R} \to \mathbb{R}$ is a special case of the directional derivative in which the canonical 'direction' is the positive direction, or the direction 1. Indeed, we see that setting $v = 1$ in the definition of the directional derivative yields the following formula:

$$
\nabla f(v; x) = \lim_{t \to 0} \frac{1}{t}[f(x+t) - f(x)]
$$

which recovers the usual single-variable derivative $f'(x)$.

In [13], this idea is extended from functions $f \colon \mathbb{R} \to \mathbb{R}$ to functors $F \colon \mathcal{C} \to \mathcal{D}$ by using generating objects of $\mathcal{C}$ in place of the unit $1 \in \mathbb{R}$. The canonical example of a generating object is the object R in the category of R-modules $\mathcal{M}\mathrm{od}_R$, where R is a commutative ring with unit. The object R is a "generating object" in the sense that all other R-modules can be obtained from finitely generated free R-modules using free resolutions and colimits. In this section, we consider functors $F \colon \mathcal{M}\mathrm{od}_R \rightsquigarrow \mathcal{A}$ and use the generating object R in $\mathcal{M}\mathrm{od}_R$ to define a derivative $\frac{d}{dR} F$. We use this derivative to obtain analogs of the chain rules for compositions of functions involving at least one function of a single variable. We note that these results could be generalized to functors $F \colon \mathcal{C} \rightsquigarrow \mathcal{A}$ by following the definition of $\frac{d}{dC} F$ in [13, Definition 2.13], however we have chosen to work with the category $\mathcal{M}\mathrm{od}_R$ for the sake of exposition.

Definition 9.1. [13, Definition 2.13] For a functor $F \colon \mathcal{M}\mathrm{od}_R \rightsquigarrow \mathcal{A}$ and $n \geq 1$, the nth *derivative* of F at X is

$$
\frac{d^n}{dR^n} F(X) := \nabla^n F(R, \ldots, R; X).
$$

The following lemma indicates another way to define $\frac{d^n}{dR^n} F(X)$.

Lemma 9.2. *For a functor $F \colon \mathcal{M}\mathrm{od}_R \rightsquigarrow \mathcal{A}$, there is a chain homotopy equivalence*

$$
\nabla^n F(R, \ldots, R; X) \simeq \Delta_n F(0, \ldots, 0, R; X).
$$

Proof. Theorem 7.7 expresses $\Delta_n F(V_n, \ldots, V_1; X)$ as a direct sum, indexed by partitions $\pi = \{S_1, \ldots, S_k\}$ of $\{1, \ldots, n\}$, of terms $\nabla^k F(V_{|S_1|}, \ldots, V_{|S_k|}; X)$. Because $\nabla^k F$ vanishes if any of its "direction" variables is zero, the only partition that contributes a non-zero summand when $V_2 = \cdots = V_n = 0$ is $\{\{1\}, \ldots, \{n\}\}$, the partition into one-element subsets. Thus

$$\Delta_n F(0, \ldots, 0, R; X) \simeq \nabla^n F(R, \ldots, R; X)$$

as claimed. $\square$

As a consequence of Theorem 8.1 we obtain an analog of [10, Theorem 2], which is stated in the introduction in equation (1.1).

Theorem 9.3. *For a composable pair $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G \colon \mathcal{M}\text{od}_R \rightsquigarrow \mathcal{B}$ of functors of abelian categories there is a chain homotopy equivalence*

$$\frac{d^n}{dR^n}(F \circ G)(X) \simeq \Delta_n F\left(\frac{d^n}{dR^n}G(X), \ldots, \frac{d}{dR}G(X); G(X)\right).$$

Proof. By Theorem 8.1

$$\Delta_n(F \circ G)(0, \ldots, 0, R; X) \simeq \Delta_n F(\Delta_n G(0, \ldots, 0, R; X), \ldots, \Delta_1 G(R; X); G(X)).$$

Applying Lemma 9.2 this becomes

$$\frac{d^n}{dR^n}(F \circ G)(X) \simeq \Delta_n F\left(\frac{d^n}{dR^n}G(X), \ldots, \frac{d}{dR}G(X); G(X)\right). \qquad \square$$

Combining Theorems 9.3 and 7.7 gives the next corollary, a Faà di Bruno-style characterization of the derivatives. We note that Arone and Ching have obtained similar results for functors of spaces and spectra, by expressing such characterizations of derivatives in terms of composition products of symmetric sequences of derivatives (see [1, Theorem 0.2] and [5, Theorem 1.15]).

Corollary 9.4. *For a composable pair $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G \colon \mathcal{M}\text{od}_R \rightsquigarrow \mathcal{B}$ of functors of abelian categories there is a chain homotopy equivalence*

$$\frac{d^n}{dR^n}(F \circ G)(X) \simeq \bigoplus_{\pi = \{S_1, \ldots, S_k\}} \nabla^\pi F\left(\frac{d^n}{dR^n}G(X), \ldots, \frac{d}{dR}G(X); G(X)\right)$$

where the sum is taken over all partitions π of $\{1, \ldots, n\}$.

The derivatives $\frac{d^n}{dR^n}F$ play an important role in functor calculus in that they classify homogeneous degree n functors, that is, degree n functors F with the property that $P_k F \simeq 0$ for $k < n$. For a functor of abelian categories $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$, Johnson–McCarthy define the nth layer of its Taylor tower to be the homogeneous degree n functor $D_n F := \operatorname{hofiber}(P_n F \to P_{n-1}F)$, i.e., the mapping cone shifted down one degree. This defines $D_n F$ up to quasi-isomorphism. The next few results relate our work to their characterizations of this functor, which hold up to quasi-isomorphism and are denoted with the symbol $\simeq_{qi}$. There is a quasi-isomorphism

$$D_n F(X) \simeq_{qi} (D_1^{(n)} \operatorname{cr}_n F(X, \ldots, X))_{h\Sigma_n}$$
$$\simeq_{qi} \nabla^n F(X, \ldots, X; 0)_{h\Sigma_n}$$

by [12, Proposition 3.9, Corollary 5.11]. The Σ_n-action is the natural action that permutes the variables of $\mathrm{cr}_n F$. These homogeneous layers are analogous to the homogeneous terms $f^{(n)}(x)/n!$ of the Maclaurin series of a function f in the calculus of functions. In terms of the derivative of Definition 9.1, this yields the following corollary.

Corollary 9.5. *For $F \colon \mathcal{M}od_R \rightsquigarrow \mathcal{A}$ there is a quasi isomorphism:*

$$D_n F(R) \simeq (\nabla^n F(R, \dots, R; 0))_{h\Sigma_n} \simeq_{qi} \left(\frac{d^n}{dR^n} F(0) \right)_{h\Sigma_n}.$$

The objects $D_n F(R)$ are classifying objects for homogeneous degree n functors by [14, Theorem 5.13]. Thus, combining these results with Theorem 9.3 and Corollary 9.4, we have the following formula for the homogeneous layers of a composition.

Corollary 9.6. *For functors $G \colon \mathcal{M}od_R \to \mathcal{B}$ and $F \colon \mathcal{B} \to \mathcal{A}$, there is a quasi-isomorphism*

$$D_n(F \circ G)(R) \simeq_{qi} \left(\Delta_n F \left(\frac{d^n}{dR^n} G(0), \dots, \frac{d}{dR} G(0); G(0) \right) \right)_{h\Sigma_n}.$$

Furthermore, there is a quasi-isomorphism

$$D_n(F \circ G)(R) \simeq_{qi} \left(\bigoplus_{\pi = \{S_1, \dots, S_k\}} \nabla^\pi F \left(\frac{d^n}{dR^n} G(0), \dots, \frac{d}{dR} G(0); G(0) \right) \right)_{h\Sigma_n}$$

where the sum is taken over all partitions π of $\{1, \dots, n\}$.

A more thorough investigation of the relationship between Corollary 9.6, the classification of functors of $\mathcal{M}od_R$, and modules over these is warranted. This will be the subject of future work.

Appendix A. A general bicomplex retraction

Throughout this paper, and especially in Sections 4 and 5, we have replaced many quasi-isomorphisms related to properties of polynomial approximation and linearization by explicit chain homotopy equivalences. In several cases (especially Proposition 5.7) we need to have such equivalences between the total complexes of bicomplexes whose rows are chain homotopy equivalent. In this appendix, we establish general criteria under which the total complexes of such bicomplexes are chain homotopy equivalent under weaker conditions than the "chain homotopy of bicomplexes" used elsewhere in this paper. In Appendix B, we provide a proof of Proposition 5.7 as an application.

Let $A_{\bullet,\bullet}$ and $B_{\bullet,\bullet}$ be first-quadrant bicomplexes. Contrary to the conventions in use elsewhere, in this appendix we assume that bicomplexes have anti-commutative squares, as this convention will simplify the signs.

Definition A.1. A morphism of first-quadrant bicomplexes $\iota \colon A_{\bullet,\bullet} \to B_{\bullet,\bullet}$ *admits a row-wise strong deformation retraction* if for all $p \geq 0$ there is a map of chain complexes $f_{p,\bullet} \colon B_{p,\bullet} \to A_{p,\bullet}$

(i) that is a retraction for $\iota_{p,\bullet}$ and
(ii) so that $\iota_{p,\bullet}$ and $f_{p,\bullet}$ induce a strong chain homotopy equivalence between $A_{p,\bullet}$ and $B_{p,\bullet}$. That is, there are morphisms $s \colon B_{p,q} \to B_{p,q+1}$ such that $ds + sd = 1 - \iota_{p,q} f_{p,q}$ and $s\iota_{p,q} = 0$.

Our aim is to prove:

Theorem A.2. *Let $\iota\colon A_{\bullet,\bullet} \to B_{\bullet,\bullet}$ be a morphism of first-quadrant bicomplexes that admits a row-wise strong deformation retraction. Then ι induces a chain homotopy equivalence of total complexes $\mathrm{Tot}(A_{\bullet,\bullet}) \to \mathrm{Tot}(B_{\bullet,\bullet})$.*

We prove Theorem A.2 by constructing an explicit retraction in Proposition A.5 and chain homotopy in Proposition A.6. We first establish our notation. Explicitly, the bicomplex $A_{\bullet,\bullet} = \{A_{p,q} \mid p, q \geq 0\}$ has:

- horizontal differentials $d^A\colon A_{p,q} \to A_{p,q-1}$ so that $(d^A)^2 = 0$, and
- vertical differentials $e^A\colon A_{p,q} \to A_{p-1,q}$ so that $(e^A)^2 = 0$,
- so that $e^A d^A + d^A e^A = 0$ (squares anti-commute).

Similarly, the bicomplex $B_{\bullet,\bullet} = \{B_{p,q} \mid p, q \geq 0\}$ has:

- horizontal differentials $d^B\colon B_{p,q} \to B_{p,q-1}$ so that $(d^B)^2 = 0$, and
- vertical differentials $e^B\colon B_{p,q} \to B_{p-1,q}$ so that $(e^B)^2 = 0$,
- so that $e^B d^B + d^B e^B = 0$ (squares anti-commute).

We will omit superscripts and simply use d and e when the context is clear. In addition to the chain homotopies described explicitly in the statement of Theorem A.2 it is convenient to allow additional maps $d\colon B_{p,0} \to B_{p,-1}$ and $s\colon B_{p,-1} \to B_{p,0}$ all equal to zero. In particular, $ds\colon B_{p,0} \to B_{p,0}$ is the map $1 - \iota f$.

The next lemma records some commutativity relations that follow from the defining relations and conditions of A.2:

$$d^2 = 0 \qquad e^2 = 0 \qquad de + ed = 0 \qquad ds + sd = 1 - \iota f \qquad s\iota = 0$$
$$f\iota = 1 \qquad fd = df \qquad \iota d = d\iota \qquad \iota e = e\iota.$$

Lemma A.3. *For any $k \geq 0$,*

(i) $ef(-es)^k + df(-es)^{k+1} = f(-es)^k e + f(-es)^{k+1}d$,
(ii) $es(-es)^k + ds(-es)^{k+1} = -\iota f(-es)^{k+1} - sd(-es)^{k+1}$,
(iii) $s(-es)^{k+1}d + s(-es)^k e = -(-se)^{k+1}ds$, and
(iv) $sd(-es)^{k+1} = -(-se)^{k+1}ds$.

Here $(-es)^k = (-1)^k(es)^k$.

Proof. We prove (i) and leave the others as exercises for the reader. We prove this by induction on k. When $k = 0$, the left hand side is

$$ef + df(-es) = ef + feds$$
$$= ef + fe(1 - sd - \iota f)$$
$$= ef + fe - fesd - fe\iota f.$$

Using the facts that ι and e commute, and $f\iota = 1$, this becomes

$$= ef + fe - fesd - ef$$

$$= fe - fesd,$$

as desired.

Assuming the statement holds for k, we have

$$ef(-es)^{k+1} + df(-es)^{k+2} = (ef(-es)^k + df(-es)^{k+1})(-es)$$
$$= (f(-es)^k e + f(-es)^{k+1} d)(-es).$$

Since $e^2 = 0$ and d and e anticommute, this is

$$= (-1)^{k+1} f(es)^{k+1} eds$$
$$= (-1)^{k+1} f(es)^{k+1} e(1 - sd - \iota f).$$

Since e and ι commute and $s\iota = 0$, this is simply

$$= f(-es)^{k+1} e + f(-es)^{k+2} d$$

and the proof is complete. $\square$

Recall the *total complex* $\mathrm{Tot}(C)$ of a first-quadrant bicomplex $C_{\bullet,\bullet}$ is the chain complex with

$$\mathrm{Tot}(C)_n := C_{n,0} + \cdots + C_{0,n}$$

(where "$+$" is less-cluttered notation for "$\oplus$") and with differential given by the matrix

$$\mathrm{Tot}(C)_n := C_{n,0} + \cdots + C_{0,n} \xrightarrow{\begin{pmatrix} e & d & 0 & \cdots & \cdots & 0 \\ 0 & e & d & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & e & d & 0 \\ 0 & \cdots & \cdots & 0 & e & d \end{pmatrix}} C_{n-1,0} + \cdots + C_{0,n-1} =: \mathrm{Tot}(C)_{n-1}.$$

See [19, 3.1.27] for an explanation of the matrix notation for a map between finite direct sums.

Remark A.4. The total complex defined above assumes we have a bicomplex with anti-commutative squares. When working with a bicomplex with commutative squares, one must introduce signs to define the total complex. A standard convention is to multiply all of the maps in every odd row by -1 to obtain a bicomplex with anti-commutative squares. As this does not affect the conditions needed for Theorem A.2, we can apply it to bicomplexes with commutative squares as well.

Proposition A.5. *For a morphism of first-quadrant bicomplexes $\iota : A_{\bullet,\bullet} \to B_{\bullet,\bullet}$ satisfying the conditions of Theorem A.2, the induced morphism of total complexes $\mathrm{Tot}(\iota) : \mathrm{Tot}(A)_\bullet \to \mathrm{Tot}(B)_\bullet$ admits a retraction $\rho : \mathrm{Tot}(B)_\bullet \to \mathrm{Tot}(A)_\bullet$ defined in degree n by*

$$\operatorname{Tot}(B)_n := B_{n,0} + \cdots + B_{0,n} \xrightarrow{\begin{pmatrix} f & 0 & \cdots & \cdots & 0 & 0 \\ f(-es) & f & 0 & \ddots & \ddots & 0 \\ f(-es)^2 & f(-es) & f & 0 & \ddots & \vdots \\ f(-es)^3 & f(-es)^2 & f(-es) & f & 0 & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ f(-es)^{n-1} & f(-es)^{n-2} & \cdots & \cdots & f & 0 \\ f(-es)^n & f(-es)^{n-1} & \cdots & \cdots & \cdots & f \end{pmatrix}} A_{n,0} + \cdots + A_{0,n} =: \operatorname{Tot}(A)_n.$$

That is, the $(n+1) \times (n+1)$ matrix defining ρ_n is the lower triangular matrix whose entry in the ith row, jth column, $j \le i$, is $f_{n-(i-1),i-1}(-es)^{i-j}$. Here $(-es)^k = (-1)^k(es)^k$.

Proof. We need to show that ρ is both a chain map and a retraction for ι. We start by showing that ρ is chain map. For each $n \ge 1$, we must verify that $\partial_n \rho_n = \rho_{n-1} \partial_n$ where ∂_n is the total complex differential. We do so by verifying that corresponding entries in the $(n-1) \times n$ matrices defining $\partial_n \rho_n$ and $\rho_{n-1} \partial_n$ agree. Using i for the row number and j for the column number of the matrix (starting from the upper left corner), we consider three cases: $j > i+1$, $j = i+1$, and $j \le i$. We note that the entry in row i, column j of these matrices is the component of our map from $B_{n-j+1,j-1}$ to $A_{n-i,i-1}$. For each case, we display a square diagram whose horizontal arrows are components of ρ and vertical arrows are components of the total complex differential ∂. The upper-right composite is the component of the map $\partial\rho$ and the lower-left composite is the component of the map $\rho\partial$.

When $j > i+1$, the diagram is

$$
\begin{array}{ccc}
B_{n-j+1,j-1} & \xrightarrow{\begin{pmatrix} 0 \\ 0 \end{pmatrix}} & A_{n+1-i,i-1} + A_{n-i,i} \\
{\scriptstyle \begin{pmatrix} d \\ e \end{pmatrix}} \downarrow & & \downarrow {\scriptstyle (e\ d)} \\
B_{n-j+1,j-2} + B_{n-j,j-1} & \xrightarrow[\begin{pmatrix} 0 & 0 \end{pmatrix}]{} & A_{n-i,i-1}.
\end{array}
$$

As this diagram clearly commutes, we have $\partial_n \rho_n = \rho_{n-1}\partial_n$ in this case.

When $j = i+1$, the diagram is

$$
\begin{array}{ccc}
B_{n-i,i} & \xrightarrow{\begin{pmatrix} 0 \\ f_{n-i,i} \end{pmatrix}} & A_{n+1-i,i-1} + A_{n-i,i} \\
{\scriptstyle \begin{pmatrix} d \\ e \end{pmatrix}} \downarrow & & \downarrow {\scriptstyle (e\ d)} \\
B_{n-i,i-1} + B_{n-i-1,i} & \xrightarrow[\begin{pmatrix} f_{n-i,i-1} & 0 \end{pmatrix}]{} & A_{n-i,i-1}.
\end{array}
$$

The composition through the upper right corner is $df_{n-i,i}$ and the composition through the lower left corner is $f_{n-i,i-1}d$. Since f is a map of chain complexes between the rows of $B_{\bullet,\bullet}$ and $A_{\bullet,\bullet}$, these are equal and we have $\partial_n \rho_n = \rho_{n-1}\partial_n$ in this case as well.

Finally, when $j \leq i$, we must verify that the diagram below commutes:

$$
\begin{array}{ccc}
B_{n-j+1,j-1} & \xrightarrow{\left(\begin{smallmatrix} f(-es)^{i-j} \\ f(-es)^{i-j+1} \end{smallmatrix}\right)} & A_{n+1-i,i-1} + A_{n-i,i} \\
{\scriptstyle \left(\begin{smallmatrix} d \\ e \end{smallmatrix}\right)}\downarrow & & \downarrow {\scriptstyle \left(\begin{smallmatrix} e & d \end{smallmatrix}\right)} \\
B_{n-j+1,j-2} + B_{n-j,j-1} & \xrightarrow{\left(\begin{smallmatrix} f(-es)^{i-j+1} & f(-es)^{i-j} \end{smallmatrix}\right)} & A_{n-i,i-1},
\end{array}
$$

that is, that

$$
f(-es)^{i-j+1}d + f(-es)^{i-j}e = ef(-es)^{i-j} + df(-es)^{i-j+1}.
$$

But this is part (i) of Lemma A.3, so ρ is a chain map.

To confirm that $\rho_n \iota_n = 1$ where

$$
\iota_n = \begin{pmatrix}
\iota_{n,0} & 0 & 0 & \cdots & 0 \\
0 & \iota_{n-1,1} & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & 0 \\
0 & 0 & 0 & \cdots & \iota_{0,n}
\end{pmatrix}
$$

we note that the entry in row i, column j of $\rho_n \iota_n$ is

  (i) 0, if $j > i$,
 (ii) $f\iota$, if $i = j$, and
(iii) $f(-es)^{i-j}\iota$, if $i > j$.

Since $f\iota = 1$ and $s\iota = 0$, ρ_n is a retraction for ι_n. $\square$

Proposition A.6. *The composite map* $\iota\rho\colon \mathrm{Tot}(B)_\bullet \to \mathrm{Tot}(B)_\bullet$ *is chain homotopic to the identity via the chain homotopy* $\sigma\colon \mathrm{Tot}(B)_\bullet \to \mathrm{Tot}(B)_{\bullet+1}$ *defined in degree* n *by*

$$
\mathrm{Tot}(B)_n := B_{n,0} + \cdots + B_{0,n} \xrightarrow{\begin{pmatrix}
0 & 0 & \cdots & \cdots & 0 & 0 \\
s & 0 & \ddots & \ddots & \ddots & 0 \\
s(-es) & s & 0 & \ddots & \ddots & \vdots \\
s(-es)^2 & s(-es) & s & 0 & \ddots & \vdots \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
s(-es)^{n-1} & s(-es)^{n-2} & \cdots & s(-es) & s & 0 \\
s(-es)^n & \cdots & \cdots & \cdots & s(-es) & s
\end{pmatrix}} B_{n+1,0} + \cdots + B_{0,n+1} =: \mathrm{Tot}(B)_{n+1}.
$$

Proof. Writing $\partial_n \colon \mathrm{Tot}(B)_n \to \mathrm{Tot}(B)_{n-1}$ for the total complex differential, we must verify that $\partial_{n+1}\sigma_n + \sigma_{n-1}\partial_n \colon \mathrm{Tot}(B)_n \to \mathrm{Tot}(B)_n$ is the matrix

$$\mathrm{Tot}(B)_n := B_{n,0} + \cdots + B_{0,n} \xrightarrow{\begin{pmatrix} 1-\iota f & 0 & \cdots & \cdots & 0 \\ -\iota f(-es) & 1-\iota f & 0 & \cdots & 0 \\ -\iota f(-es)^2 & -\iota f(-es) & 1-\iota f & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\iota f(-es)^{n-1} & \cdots & -\iota f(-es) & 1-\iota f & 0 \\ -\iota f(-es)^n & \cdots & -\iota f(-es)^2 & -\iota f(-es) & 1-\iota f \end{pmatrix}} B_{n,0} + \cdots + B_{0,n} =: \mathrm{Tot}(B)_n.$$

i.e., that

$$\partial_{n+1}\sigma_n + \sigma_{n-1}\partial_n = \mathrm{id}_{\mathrm{Tot}(B)_n} - \iota_n\rho_n.$$

For each pair $1 \leq i, j \leq n+1$, we must verify that the component of $\partial_{n+1}\sigma_n + \sigma_{n-1}\partial_n$ from $B_{n-j+1,j-1}$ to $B_{n-i+1,i-1}$ agrees with the matrix entry in the jth column and ith row, counting from the left and from the top. As in the proof of Proposition A.5, we do this in several cases. For each case, we display a square diagram whose horizontal arrows are components of σ and vertical arrows are components of the total complex differential ∂. The upper-right composite is the component of the map $\partial\sigma$ and the lower-left composite is the component of the map $\sigma\partial$ whose sum we are computing.

We begin with the case where $j > i$. In this case, the entries represent maps $B_{n-j+1,j-1} \to B_{n-i+1,i-1}$. We must show that the sum of the two composites

$$
\begin{array}{ccc}
B_{n-j+1,j-1} & \xrightarrow{\begin{pmatrix} 0 \\ 0 \end{pmatrix}} & B_{n-i+2,i-1} + B_{n-i+1,i} \\
\left.\begin{pmatrix} d \\ e \end{pmatrix}\right\downarrow & & \downarrow\left(\begin{smallmatrix} e & d \end{smallmatrix}\right) \\
B_{n-j+1,j-2} + B_{n-j,j-1} & \xrightarrow{\left(\begin{smallmatrix} 0 & 0 \end{smallmatrix}\right)} & B_{n-i+1,i-1}
\end{array}
$$

is 0, which is evident.

We next verify the diagonal entries $B_{n-j+1,j-1} \to B_{n-j+1,j-1}$ for $1 \leq j \leq n+1$. Here we must show that the sum of the maps

$$
\begin{array}{ccc}
B_{n-j+1,j-1} & \xrightarrow{\begin{pmatrix} 0 \\ s \end{pmatrix}} & B_{n-j+2,j-1} + B_{n-j+1,j} \\
\left.\begin{pmatrix} d \\ e \end{pmatrix}\right\downarrow & & \downarrow\left(\begin{smallmatrix} e & d \end{smallmatrix}\right) \\
B_{n-j+1,j-2} + B_{n-j,j-1} & \xrightarrow{\left(\begin{smallmatrix} s & 0 \end{smallmatrix}\right)} & B_{n-j+1,j-1}
\end{array}
$$

is $1 - \iota f \colon B_{n-j+1,j-1} \to B_{n-j+1,j-1}$, which is so because $ds + sd = 1 - \iota f$.

It remains to verify the lower diagonal entries. Here we wish to show that for $i > j > 0$ the sum of the maps

$$B_{n-j+1,j-1} \xrightarrow{\begin{pmatrix} s(-es)^{i-j-1} \\ s(-es)^{i-j} \end{pmatrix}} B_{n-i+2,i-1} + B_{n-i+1,i}$$

$$\begin{pmatrix} d \\ e \end{pmatrix} \Big\downarrow \qquad\qquad\qquad\qquad \Big\downarrow \begin{pmatrix} e & d \end{pmatrix}$$

$$B_{n-j+1,j-2} + B_{n-j,j-1} \xrightarrow{\begin{pmatrix} s(-es)^{i-j} & s(-es)^{i-j-1} \end{pmatrix}} B_{n-i+1,i-1}$$

is the map $-\iota f(-es)^{i-j} \colon B_{n-j+1,j-1} \to B_{n-i+1,i-1}$. Here the sum of the composites is

$$es(-es)^{i-j-1} + ds(-es)^{i-j} + s(-es)^{i-j}d + s(-es)^{i-j-1}e.$$

Parts (ii), (iii), and (iv) of Lemma A.3 guarantee that this is equal to $-\iota f(-es)^{i-j}$. $\quad\square$

As an immediate corollary of A.2, we have the following.

Corollary A.7. *Let $A_{\bullet,\bullet}$ be a first-quadrant bicomplex so that every row except the zeroth row $A_{0,\bullet}$ is contractible. Then the natural inclusion $A_{0,\bullet} \hookrightarrow \mathrm{Tot}(A)_\bullet$ is a chain homotopy equivalence.*

Proof. In this case, the evident inclusion $\iota \colon A_{0,\bullet} \hookrightarrow \mathrm{Tot}(A)_\bullet$ of chain complexes admits a retraction $\rho \colon \mathrm{Tot}(A)_\bullet \to A_{0,\bullet}$ defined in degree n by

$$\mathrm{Tot}(A)_n := A_{n,0} + \cdots + A_{0,n} \xrightarrow{\begin{pmatrix} (-es)^n & (-es)^{n-1} & \cdots & (-es)^2 & (-es) & 1 \end{pmatrix}} A_{0,n}. \qquad\square$$

Appendix B. Proof of Proposition 5.7

We now apply Theorem A.2 to prove:

Proposition 5.7. *For any composable pair of functors $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ with G reduced, there is a chain homotopy equivalence*

$$D_1(F \circ G) \simeq D_1 F \circ D_1 G.$$

In fact, it suffices to assume that $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ is also reduced. To see this, note that by direct computation:

Lemma B.1. *Given $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ and $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ with G reduced, then $\mathrm{cr}_1(FG) \cong \mathrm{cr}_1 F \circ G$.*

Proof. Because G is reduced, both terms are direct sum complements of $FG(0)$ in FG. $\quad\square$

Lemma B.1 has the following corollary:

Corollary B.2. *For G reduced and F not necessarily reduced $D_1(F \circ G) \simeq D_1(\mathrm{cr}_1 F \circ G)$. In particular, to prove that $D_1 F \circ D_1 G \simeq D_1(F \circ G)$ it suffices to assume that F is also reduced.*

Proof. By Lemma 5.8, $D_1 F \cong D_1(\mathrm{cr}_1 F)$ for all F. So

$$D_1(F \circ G) \cong D_1(\mathrm{cr}_1(F \circ G)) \cong D_1(\mathrm{cr}_1 F \circ G),$$

as claimed. Since $D_1 F \circ D_1 G \cong D_1 \mathrm{cr}_1 F \circ D_1 G$ it suffices to consider the case when both functors are reduced. $\quad\square$

On account of Corollary B.2, for the remainder of this section, we restrict our consideration to functors that are reduced.

In the Kleisli category, the composition of $G: \mathcal{C} \rightsquigarrow \mathcal{B}$ with $F: \mathcal{B} \rightsquigarrow \mathcal{A}$ is defined by prolonging the functor $F: \mathcal{B} \to \mathsf{Ch}\mathcal{A}$ to a functor $\mathsf{Ch}(F): \mathsf{Ch}\mathcal{B} \to \mathsf{ChCh}\mathcal{A}$ via the Dold–Kan equivalence and then applying the totalization $\mathrm{Tot}: \mathsf{ChCh}\mathcal{A} \to \mathsf{Ch}\mathcal{A}$. When F is strictly reduced, there is a functor $\tilde{F}: \mathsf{Ch}\mathcal{B} \to \mathsf{ChCh}\mathcal{A}$ defined by simply applying the functor directly to each term in the chain complex. When F is also linear, [11, Lemma 5.4] show that the two prolongations $\mathrm{Tot}(\tilde{F})$ and $\mathrm{Tot}(\mathsf{Ch}(F))$ are quasi-isomorphic. Lemma B.7 shows that when F is a linearization D_1H of some functor $H: \mathcal{B} \rightsquigarrow \mathcal{A}$, in which case Lemma 5.6(i) shows that F is strictly reduced and linear in the sense of Definition 5.5 (preserving finite direct sums up to natural chain homotopy equivalence), these procedures are in fact chain homotopy equivalent. We will use this in the proof of Proposition 5.7 to construct the composite $D_1F \circ D_1G$ using the simpler form of the prolongation.

We note that the definition of $\tilde{F}$ only makes sense when F is strictly reduced. Otherwise, applying F degreewise to the objects and maps in a chain complex may not produce a chain complex, as $F(\partial \circ \partial)$ may no longer be 0.

The proof of Lemma B.7 will proceed in several steps, corresponding to the verification of the hypotheses of Theorem A.2. Write $K: \mathsf{Ch}\mathcal{B} \to \mathcal{B}^{\Delta^{\mathrm{op}}}$ and $N: \mathcal{B}^{\Delta^{\mathrm{op}}} \to \mathsf{Ch}\mathcal{B}$ for the functors in the Dold–Kan equivalence. The functor N is naturally chain homotopy equivalent to another functor $M: \mathcal{B}^{\Delta^{\mathrm{op}}} \to \mathsf{Ch}\mathcal{B}$ with a simpler definition: for a simplicial object $Y \in \mathcal{B}^{\Delta^{\mathrm{op}}}$, $M(Y)_n = Y_n$ with differentials defined to be the alternating sum of the face maps; see [21, §8.4] and [9, III.2.4] for definitions of these functors and the chain homotopy equivalence. Let $F_*: \mathcal{B}^{\Delta^{\mathrm{op}}} \to \mathsf{Ch}\mathcal{A}^{\Delta^{\mathrm{op}}}$ be the functor defined by post-composition with $F: \mathcal{B} \to \mathsf{Ch}\mathcal{A}$. Recall that $\mathsf{Ch}(F) = N \circ F_* \circ K$ defines the prolongation.

Remark B.3. For a chain complex $X \in \mathsf{Ch}\mathcal{B}$, there is a canonical comparison morphism $\iota_*: K\tilde{F}(X) \to F_*(KX)$ in $\mathsf{Ch}\mathcal{A}^{\Delta^{\mathrm{op}}}$ whose component in simplicial degree n is the natural map

$$\iota: \bigoplus_{[n] \twoheadrightarrow [k]} F(X_k) \to F\left(\bigoplus_{[n] \twoheadrightarrow [k]} X_k \right)$$

where the sums are over all surjections $[n] \twoheadrightarrow [k]$ in Δ. Composing with the functor M, this defines a canonical map of bicomplexes $\iota := M(\iota_*): MK\tilde{F}(X) \to MF_*(KX)$, where the latter functor is chain homotopy equivalent to $\mathsf{Ch}F$.

The first step is to show that this map is a chain homotopy equivalence in each row. This follows from the next two lemmas.

Lemma B.4. *The inclusion of chain complexes $\iota: A \to A \oplus B$ is a chain homotopy equivalence if and only if B is contractible.*

Proof. If ι is a chain homotopy equivalence, then there is a map $\rho: A \oplus B \to A$ such that $\iota \circ \rho \simeq \mathrm{id}_{A \oplus B}$. In other words, there is a chain homotopy

$$A_n \oplus B_n \xrightarrow{\;s = \begin{pmatrix} s_{AA} & s_{BA} \\ s_{AB} & s_{BB} \end{pmatrix}\;} A_{n+1} \oplus B_{n+1}$$

so that

$$ds + sd = \mathrm{id}_{A \oplus B} - \iota \circ \rho: A_n \oplus B_n \to A_n \oplus B_n. \tag{B.5}$$

The component of the map $\iota \circ \rho \colon A_n \oplus B_n \to A_n \oplus B_n$ from B_n to B_n is the zero map. Thus, the component from B_n to B_n of the relation (B.5) asserts that $d_B s_{BB} + s_{BB} d_B = \mathrm{id}_B$, where d_B is the differential in B. Thus, the component $s_{BB} \colon B_n \to B_{n+1}$ of s provides the desired contracting homotopy.

On the other hand, if B is contractible then there is a chain homotopy equivalence $h : B_n \to B_{n+1}$ so that

$$d_B h + h d_B = id_B.$$

Then, the map defined by

$$A_n \oplus B_n \xrightarrow{\ s=\begin{pmatrix} 0 & 0 \\ 0 & h=s_{BB} \end{pmatrix}\ } A_{n+1} \oplus B_{n+1}$$

is a chain homotopy equivalence between $id_{A\oplus B}$ and $\iota\pi$ where π is the canonical projection map. By direct verification,

$$ds + sd = d_B h + h d_B = id_{A\oplus B} - \iota\pi. \qquad \square$$

Lemma B.6. *Let $F = D_1 H$. Then the chain complex $\mathrm{cr}_n F$ is chain contractible for all $n \geq 2$.*

Proof. We proceed by induction on n. When $n = 2$, the lemma is a consequence of Lemma 2.9 and Proposition 4.5(i), since $D_1 H \cong P_1(\mathrm{cr}_1 H)$ as in Remark 5.2(ii). Now suppose that $\mathrm{cr}_n F$ is chain contractible. By definition of the cross effects, $\mathrm{cr}_{n+1} F$ is a direct summand of $\mathrm{cr}_n F$:

$$\mathrm{cr}_n F(X \oplus Y, Z_2, \ldots, Z_n) \cong \mathrm{cr}_n F(X, Z_2, \ldots, Z_n) \oplus \mathrm{cr}_n F(Y, Z_2, \ldots, Z_n) \oplus \mathrm{cr}_{n+1} F(X, Y, Z_2, \ldots, Z_n).$$

Since $\mathrm{cr}_n F(X, Z_2, \ldots, Z_n)$ and $\mathrm{cr}_n F(Y, Z_2, \ldots, Z_n)$ are contractible by hypothesis, the inclusion $\mathrm{cr}_{n+1} F(X, Y, Z_2, \ldots, Z_n) \hookrightarrow \mathrm{cr}_n F(X \oplus Y, Z_2, \ldots, Z_n)$ is a chain homotopy equivalence by Lemma B.4. The result follows since $\mathrm{cr}_n F(X \oplus Y, Z_2, \ldots, Z_n)$ is contractible by hypothesis. $\square$

An immediate consequence of Lemmas B.6 and B.4 is that the nth row of the bicomplex obtained by applying F degreewise to a chain complex X is chain homotopy equivalent to the nth row of the bicomplex obtained by prolonging F when evaluating on X. This, together with Theorem A.2, proves the following lemma.

Lemma B.7. *Let $H \colon \mathcal{B} \rightsquigarrow \mathcal{A}$ be a functor and $F = D_1 H$. Then $\mathrm{Tot}(\mathsf{Ch}(F))$ and $\mathrm{Tot}(\tilde{F})$ are chain homotopy equivalent.*

Proof. We first show that $\mathrm{Tot}\, M \circ F_* \circ K$ and $\mathrm{Tot}\, \tilde{F}$ are chain homotopy equivalent. In particular, we show that the totalization of the canonical comparison map $\iota \colon MK\tilde{F} \to MF_*K$ of Remark B.3 is a chain homotopy equivalence. Note that ι is induced by the inclusion maps $X_k \to \oplus_{[n] \twoheadrightarrow [k]} X_k$; for this reason, ι_* commutes with the face maps of the simplicial objects $K\tilde{F}$ and F_*K. Thus, ι itself is a map of bicomplexes. The map ι has a retraction, ρ, in each row (where a row of $MK\tilde{F}$ or MF_*K corresponds to a fixed simplicial degree in $K\tilde{F}$ or F_*K, respectively), given by projection. In particular, since

$$F\left(\bigoplus_{[n] \twoheadrightarrow [k]} X_k \right) \cong \bigoplus_{[n] \twoheadrightarrow [k]} F(X_k) \oplus_{k=1}^{n} \left(\oplus_{j_1 < \ldots < j_k} \mathrm{cr}_k F(X_{j_1}, \ldots, X_{j_k}) \right),$$

there is a canonical projection map $\rho_* \colon F_* KX \to K\tilde{F}$ given by mapping each of the $\mathrm{cr}_k F$ terms to 0. This induces a map $\rho := M(\rho_*)$ of chain complexes on each row. Since ι and ρ are inclusion and projection of direct summands, the composite $\rho \circ \iota$ is the identity on each row of $MK\tilde{F}$.

We must show that $\iota \circ \rho$ is chain homotopic to the identity map 1_{MF_*K} in each row. Let $B = \oplus_{k=1}^{n} \left(\oplus_{j_1 < \ldots < j_k} \mathrm{cr}_k F(X_{j_1}, \ldots, X_{j_k}) \right)$. By Lemma B.6, in each row there is a chain contraction $h \colon B \to B$ with $sd_B + d_B s = 1_B$. As in the proof of Lemma B.4, the map

$$s \colon \bigoplus_{[n] \to [k]} F(X_k) \oplus B \to \bigoplus_{[n] \to [k]} F(X_k) \oplus B$$

defined by

$$s = \begin{pmatrix} 0 & 0 \\ 0 & h \end{pmatrix}$$

is the desired chain homotopy equivalence. Note that s satisfies $s\iota = 0$ since s is defined to be 0 on $\bigoplus_{[n] \to [k]} F(X_k)$.

By Theorem A.2, s induces a chain homotopy equivalence between the total complexes $\mathrm{Tot}\, M \circ F_* \circ K$ and $\mathrm{Tot}\, \tilde{F}$.

To complete the proof, we use the fact that N and M are chain homotopy equivalent. We have the following commuting square in which the left, right, and bottom arrows are chain homotopy equivalences:

$$
\begin{array}{ccc}
\mathrm{Tot} NK\tilde{F}(X) & \longrightarrow & \mathrm{Tot} NF_*(KX) \\
\simeq \downarrow & & \downarrow \simeq \\
\mathrm{Tot} MK\tilde{F}(X) & \overset{\simeq}{\longrightarrow} & \mathrm{Tot} MF_*(KX)
\end{array}
$$

This implies that the top arrow is a chain homotopy equivalence. The result follows by noting that NK is isomorphic to the identity, and $NF_*(KX) = \mathsf{Ch}(F)(X)$. $\quad\square$

Since we can safely assume that both F and G are reduced, the chain complexes $D_1 F$, $D_1 G$ and $D_1(FG)$ are defined as comonad resolutions for the second cross effect comonad C_2. For readability, we simplify our notation and write "C" for the comonad C_2 and "$L \dashv R$" for the adjoint functors of Corollary 2.7 defining the comonad $C = LR$. Recall L is the diagonal functor, which has an important property:

Lemma B.8. *For any $G \colon \mathcal{C} \rightsquigarrow \mathcal{B}$ and $F \colon \mathcal{B} \rightsquigarrow \mathcal{A}$*

$$
\begin{array}{ccc}
\mathrm{Fun}_*(\mathcal{C}^2, \mathcal{B}) & \overset{L}{\longrightarrow} & \mathrm{Fun}_*(\mathcal{C}, \mathcal{B}) \\
F_* \downarrow & & \downarrow F_* \\
\mathrm{Fun}_*(\mathcal{C}^2, \mathcal{A}) & \overset{L}{\longrightarrow} & \mathrm{Fun}_*(\mathcal{C}, \mathcal{A})
\end{array}
\qquad
\begin{array}{ccc}
\mathrm{Fun}_*(\mathcal{B}^2, \mathcal{A}) & \overset{L}{\longrightarrow} & \mathrm{Fun}_*(\mathcal{B}, \mathcal{A}) \\
G^* \downarrow & & \downarrow G^* \\
\mathrm{Fun}_*(\mathcal{C}^2, \mathcal{A}) & \overset{L}{\longrightarrow} & \mathrm{Fun}_*(\mathcal{C}, \mathcal{A})
\end{array}
$$

That is the diagonal functor commutes with both pre- and post-composition.

In particular, $LR(F) \circ G \cong L(RF \circ G)$ and $F \circ LR(G) \cong L(F \circ RG)$. The left adjoint L also preserves direct sums of functors as does the comonad $C = LR$, by Proposition 2.10. We are now prepared to embark upon the proof.

Proof of Proposition 5.7. Our strategy, following the proof of [12, Lemma 5.7], is to show that the natural inclusions of $D_1(F \circ G)$ and $D_1 F \circ D_1 G$ into the double totalization of a first octant tricomplex $T(F, G)$ are chain homotopy equivalences.[5] Consider the tricomplex $T(F, G)$ defined by

$$T(F, G)_{p,q,r} := (LR)^r((LR)^p F \circ (LR)^q G)$$

and its levelwise totalization

$$A(F, G)_{n,r} := \bigoplus_{p+q=n} (LR)^r \left((LR)^p F \circ (LR)^q G\right) \cong (LR)^r \left(\bigoplus_{p+q=n} (LR)^p F \circ (LR)^q G\right).$$

The $n = 0$ column of $A(F, G)_{n,r}$ is $D_1(FG)$ while the $r = 0$ row is chain homotopy equivalent to $D_1 F \circ D_1 G$ by Lemma B.7. The rth row is $(LR)^r(D_1 F \circ D_1 G)$.

Lemma B.8 and the fact that L preserves sums implies that the columns indexed by each $n > 0$ are each the comonad resolution chain complexes for LR applied to a functor of the form LH for some functor H.

Thus, our bicomplex $A(F, G)$ is of the form:

In particular, each column except for the 0th is contractible, using the unit η of the adjoint $L \dashv R$. Applying Corollary A.7 it follows immediately that the canonical inclusion $D_1(FG) \hookrightarrow \mathrm{Tot}(A(F, G))$ is a chain homotopy equivalence.

By Corollary A.7, to show that $D_1 F \circ D_1 G \hookrightarrow \mathrm{Tot}(A(F, G))$ is a chain homotopy equivalence it suffices to prove that the rows $C^r(D_1 F \circ D_1 G)$ of $A(F, G)$ are also contractible for each $r > 0$. By exactness of C, it suffices to prove that the first row $C(D_1 F \circ D_1 G)$ is contractible.

By Proposition 4.5(i), the chain complexes $CD_1 G(X)$ and $CD_1 F(Y)$ are contractible. It follows that the inclusions

$$D_1 G(X) + D_1 G(X) \hookrightarrow D_1 G(X + X) \qquad D_1 F(Y) + D_1 F(Y) \hookrightarrow D_1 F(Y + Y),$$

which admit $CD_1 G(X)$ and $CD_1 F(Y)$ as direct sum complements, are chain homotopy equivalences.

By Lemma 3.4, these chain homotopy equivalences are preserved by pre- and post-composition. In particular, there is a composite chain homotopy equivalence

$$D_1 F(D_1 G(X)) + D_1 F(D_1 G(X)) \hookrightarrow D_1 F(D_1 G(X) + D_1 G(X)) \to D_1 F(D_1 G(X + X)).$$

It follows that the direct sum complement $C(D_1 F \circ D_1 G)(X)$ is contractible. $\square$

---

[5] Associativity of the chain complex monad asserts the double totalization of a tricomplex is well-defined up to isomorphism.

References

[1] G. Arone, M. Ching, Operads and Chain Rules for the Calculus of Functors, Astérisque, vol. 338, 2011, vi–158 pp.

[2] K. Bauer, B. Johnson, R. McCarthy, Cross effects and calculus in an unbased setting, Trans. Am. Math. Soc. 367 (9) (2015) 6671–6718.

[3] M. Basterra, K. Bauer, A. Beaudry, R. Eldred, B. Johnson, M. Merling, S. Yeakel, Unbased calculus for functors to chain complexes, women in topology: collaborations in homotopy theory, Contemp. Math. 641 (2015) 29–48.

[4] R.F. Blute, J.R.B. Cockett, R.A.G. Seely, Cartesian differential categories, Theory Appl. Categ. 22 (2009) 622–672.

[5] M. Ching, A chain rule for Goodwillie derivatives of functors from spectra to spectra, Trans. Am. Math. Soc. 362 (1) (2010) 399–426.

[6] J.R.B. Cockett, G.S.H. Cruttwell, Differential structure, tangent structure, and SDG, Appl. Categ. Struct. 22 (2) (2014) 331–417.

[7] S. Eilenberg, S. Mac Lane, On the groups $H(\Pi, n)$. II. Methods of computation, Ann. Math. (2) 60 (1954) 49–139.

[8] T. Goodwillie, Calculus. III. Taylor series, Geom. Topol. 7 (2003) 645–711.

[9] P. Goerss, J. Jardine, Simplicial Homotopy Theory, Progress in Mathematics, Birkhäuser-Verlag, 1999.

[10] H.-N. Huang, S.A.M. Marcantognini, N.J. Young, Chain rules for higher derivatives, Math. Intell. 28 (2) (2006) 61–69.

[11] B. Johnson, R. McCarthy, Linearization, Dold–Puppe stabilization, and Mac Lane's Q-construction, Trans. Am. Math. Soc. 350 (4) (1998) 1555–1593.

[12] B. Johnson, R. McCarthy, Deriving calculus with cotriples, Trans. Am. Math. Soc. 356 (2) (2004) 757–803.

[13] B. Johnson, R. McCarthy, A classification of degree n functors. I, Cah. Topol. Géom. Différ. Catég. 44 (1) (2003) 2–38.

[14] B. Johnson, R. McCarthy, A classification of degree n functors. II, Cah. Topol. Géom. Différ. Catég. 44 (3) (2003) 163–216.

[15] B. Johnson, R. McCarthy, Taylor towers of symmetric and exterior powers, Fundam. Math. 201 (3) (2008) 197–216.

[16] J. Klein, J. Rognes, A chain rule in the calculus of homotopy functors, Geom. Topol. 6 (2002) 853–887.

[17] S. Mac Lane, Categories for the Working Mathematician, 2nd ed., Graduate Texts in Mathematics, Springer-Verlag, New York, 1998.

[18] F. Marmolejo, Doctrines whose structure forms a fully faithful adjoint string, Theory Appl. Categ. 3 (2) (1997) 22–42.

[19] E. Riehl, Category Theory in Context, Aurora: New Mathematical Monographs, Dover Press, 2016.

[20] M. Shulman, Set theory for category theory, arXiv:0810.1279, 2010.

[21] C. Weibel, An Introduction to Homological Algebra, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 1994.

[22] S. Yeakel, A Monoidal Model for Goodwillie's Derivatives, Doctoral dissertation, University of Illinois, Urbana-Champaign, 2016.