QUANTITATIVE DIFFERENTIATION

Jeff Cheeger

What is differentiability?

The condition that

$$f : [0, 1] \to \mathbb{R}$$

is *differentiable* at $\underline{x}$ can be expressed as follows.

Suppose we restrict f to the interval

$$[\underline{x} - \delta, \underline{x} + \delta]$$

and rescale the domain and target by the factor δ^{-1}.

Then in the limit as $\delta \to 0$, the rescalings of the function $f(x) - f(\underline{x})$ converge to the *unique linear function*

$$f'(\underline{x})(x - \underline{x}) \,.$$

Separate elements.

The following are *separate* elements.

1) Special structure (linear structure) .

2) Existence of the limit.

Note that even if we make the normalization,

$$|f'| \leq 1\,,$$

it is not possible to control the *rate* at which the rescalings converge to the limit as $\delta \to 0$.

This would follow from a bound on $|f''|$.

Quantitative differentiation.

The idea of "quantitative differentiation" is to concentrate solely on quantifying the *approximate linearity* of the rescalings of

$$f(x) - f(\underline{x})\,.$$

Under the normalization $|f'| \leq 1$, this turns out to be possible, *provided* we do *not* insist that the relevant approximating linear function is

$$f'(\underline{x})(x - \underline{x})\,.$$

The main assertion.

There is a natural measure $\mathcal{C}$ on the collection of all subintervals $J \subset [0, 1]$.

To simplify the exposition, we will use a discrete version in which we consider only diadic intervals $I_{n,j}$, and put

$$\mathcal{C}(I_{n,j}) := |I_{n,j}| = 2^{-n} \, .$$

Thus, for any fixed n,

$$\mathcal{C}\left(\bigcup_j I_{n,j}\right) = \sum_j |I_{n,j}| = 1 \, ,$$

and summing this over all n gives ∞.

However: For all $\epsilon > 0$, there is a collection of diadic subintervals, $\mathcal{G}_\epsilon$, whose complement has at most a definite *finite* measure (depending on ϵ) such that:

$$f \,|\, I_{n,j} \text{ looks ``}\epsilon\text{-linear'' for } I_{n,j} \in \mathcal{G}_\epsilon.$$

Why we can't always use $f'(\underline{x})$.

We return to the basic case

$$f : [0,1] \to \mathbb{R}\,,$$

$$|f'(x)| \leq 1\,.$$

Thus, f has a first order Taylor expansion,

$$f(\underline{x}) + f'(\underline{x})(x - \underline{x})$$

for all $\underline{x}$.

However, even on a short interval J the derivative f' might vary rapidly.

In such a case, f will not be well approximated by $f(\underline{x})+f'(\underline{x})(x-\underline{x})$, for most $\underline{x} \in J$.

A basic example (Semmes).

For $k >> 1$, put

$$f_k(x) = \frac{1}{k}\sin kx\,.$$

If on a subinterval J of length $|J|$, we have

$$\frac{1}{k} << |J|\,,$$

then an essentially optimal choice ℓ for a linear approximation of f is:

$$\ell \equiv 0\,.$$

If $|J| \sim \frac{1}{k}$ then $f\,|\,J$, does *not* look linear, but this includes is only a *finite* number of scales.

If $|J| << k$, f is well approximated by its first order Taylor expansion.

Some history.

For the case of functions, $f : \mathbb{R}^n \to \mathbb{R}$, the basic case of what we call *quantitative differentiation* appears in a paper of Peter Jones from 1988.

Related results of Dorronsoro are from 1985.

See also the book of David-Semmes 1993.

See also R. Schul 2009.

A general perspective.

Our goal is to demonstrate that the model case is a particular instance of a general phenomenon which is present in many different geometric analytic contexts.

A less precise discussion is given in (Cheeger-Kleiner-Naor) [arxiv:2006]; see Section 14.

Here we point out that in a each potential instance, to obtain a quantitative differention theorem, one must verify a single estimate, which we term:

Coercivity of relative defects.

In so far as we can tell this general perspective is a *basic* principle which has been overlooked.

At least in the model case, it could be covered in standard courses.

Applications.

Here are some recent applications that are included in the general paradigm.

A quantitative generalized Rademacher theorem as in [GAFA, 1999] (Cheeger).

Lipschitz maps from the Heisenberg group to L_1; (Cheeger-Kleiner-Naor), [arxiv:2006].

Related estimates for sets of finite perimeter (Cheeger).

Curvature estimates for Einstein manifolds; (Cheeger-Nabor), [arxiv:2011].

Estimates for harmonic maps, minimal surfaces and other energies; (Cheeger-Nabor) [to appear].

Significantly, the mechanisms responsible for the coercivity of relative defects, vary considerably from case to case.

Measure on the space of metric balls.

Given a metric space (X, d^X) we put

$$B_r(x) = \{x' \mid d^X(x, x') < r\} \,.$$

We regard $\{B_r(x)\} = X \times \mathbb{R}_+$.

If in addition X carries a measure μ we give $\{B_r(x)\}$ the measure

$$\mathcal{C} = r^{-1}\, dr \times \mu \,.$$

We have

$$\mathcal{C}(\{B_r(x) \mid x \in U, r_1 \leq r \leq r_2\}) = \log \frac{r_2}{r_1} \cdot \mu(U) \,.$$

Thus, all scales carry the same amount of measure and as a consequence,

$$\mathcal{C}(\{B_r(x) \mid x \in U, < r \leq r_0\}) = \infty \,.$$

Remark. Compare the appendix by Semmes in Gromov's "green book".

The discrete version for $X = \mathbb{R}$.

For purposes of exposition, we consider the discrete (diadic) version.

Given an interval I, we partition it into 2^n intervals $I_{n,j}$ with disjoint interiors, such that

$$I_{n,j} = 2^{-n} \cdot |I|\,.$$

In particular, for each n, N, j,

$$I_{n,j} = \bigcup_{I_{n+N,k} \subset I_{n,j}} I_{n+N,k}\,.$$

As above, for each $I_{n,j}$, there are 2^N intervals $I_{n+N,k}$, with $I_{n+N,k} \subset I_{n,j}$, and

$$|I_{n+N,k}| = 2^{-N} \cdot |I_{n,j}|\,.$$

The measure on $\{I_{n,j}\}$.

Put

$$\mathcal{C}(I_{n,j}) = |I_{n,j}| = 2^{-n} \cdot |I| \,.$$

Thus, for all n.

$$\mathcal{C}\left(\bigcup_{j=1}^{2^n} I_{n,j}\right) = \sum_j |I_{n,j}| = |I| \,.$$

As in the continuous case, the mass of $\mathcal{C}$ is infinite:

$$\mathcal{C}\left(\bigcup_{n,j} I_{n,j}\right) = \sum_{n,j} |I_{n,j}| = \infty \,.$$

Deviation from linearity of $f \mid I$.

Let ℓ denote a generic affine linear function.

For $f : I \to \mathbb{R}$ put:

$$\alpha(f, I) = |I|^{-1} \cdot \inf_{\ell} \sup_{x \in I} |f(x) - \ell(x)| \,.$$

Clearly, $\alpha(\,\cdot\,,\,\cdot\,)$ is translation invariant.

Remark. If in place of fuctions with bounded derivative, we were to consider $H^{1,p}$ $(p > 1)$ functions, the sup norm above should be replaced by the L_p norm.

Scale invariance.

Also, $\alpha(\,\cdot\,,\,\cdot\,)$ is *scale invariant* in the following sense.

If $I = [a, b]$ and for any $t > 0$, we define

$$tI = [ta, tb]\,,$$

$$f_t : [ta, tb] \to \mathbb{R}$$

by

$$f_t(x) = tf(t^{-1}x)\,,$$

then

$$f_t'(x) = f'(t^{-1}x)$$

and

$$\alpha(f_t, tI) = \alpha(f, I)\,.$$

Theorem. Let

$$f : I \to \mathbb{R}\,,$$

$$|f'| \leq 1\,.$$

Then

$$\mathcal{C}(\{I_{n,j} \,|\, \alpha(f \,|\, I_{n,j}) \geq \epsilon\}) \leq 5|\log_2 \epsilon| \cdot \epsilon^{-2} \cdot |I|\,. \tag{1}$$

Equivalently,

$$\sum_{I_{n,j} \,|\, \alpha(f \,|\, I_{n,j}) \geq \epsilon} |I_{n,j}| \leq 5|\log_2 \epsilon| \cdot \epsilon^{-2} \cdot |I|\,.$$

Remark. Since they are scale invariant, the above estimates apply to every subinterval of I as well.

The energy.

Define the energy $\mathcal{E}(f,I)$ by

$$\mathcal{E}(f,I) = \frac{1}{|I|} \cdot \int_I (f')^2$$

Let

$$h_I = \frac{1}{|I|} \int_I h \,.$$

Note that for functions agreeing with f on ∂I, the minimum value of $\mathcal{E}(f,I)$ is

$$\frac{1}{|I|} \cdot \int_I (f_I')^2 \,.$$

It is achieved precisely for linear functions:

$$\alpha(f,I) = 0 \,.$$

The defect.

Define the *defect* of $f \mid I$ by:

$$\widehat{V}(f,I) = \frac{1}{|I|} \cdot \int_I (f')^2 - (f_I')^2$$

$$= \frac{1}{|I|} \cdot \int_I (f' - f_I')^2 \qquad (2)$$

$$\geq 0\,.$$

Thus, $\widehat{V}(f,I)$ can be viewed as the amount by which the energy exceeds its minimum possible value.

The minimum is taken on precisely when $f' \equiv f_I'$, or equivalently, when $\alpha(f,I) = 0$.

Define the *total defect* $V(f,I)$ by

$$V(f,I) = |I| \cdot \widehat{V}(I,f)\,.$$

Monotonicity and the defect family.

Define the *total defect on scale* 2^{-n} by

$$V_{2^{-n}}(f,I) = \sum_{j=1}^{2^n} |I_{n,j}| \cdot \widehat{V}(f, I_{n,j})\,.$$

Thus,

$$V(f,I) = V_1(f,I)$$

and

$$V_{2^{-n}}(f,I) = \sum_{j=1}^{2^n} \int_{I_{n,j}} (f')^2 - (f'_{I_{n,j}})^2\,.$$

It is clear that as $2^{-n} \to 0$,

$$V_{2^{-n}}(f,I) \searrow 0\,. \tag{3}$$

In particular, the total defect is a *monotone* function of 2^{-n}.

The relative defect.

We define the *total relative defect*,

$$\begin{aligned} D_{2^{-N}}(f,I) &= V(f,I) - V_{2^{-N}}(f,I) \\ &\geq 0\,, \end{aligned}$$

and the *relative defect*,

$$\widehat{D}_{2^{-N}}(I,f) = \frac{1}{|I|} \cdot D_{2^{-N}}(I,f)\,. \qquad (4)$$

Roughly speaking, when $\widehat{D}_{2^{-N}}(f,I)$ is small and $V(f,I)$ is not, the defect is concentrated below scale 2^{-N}.

As previously indicated, it will be crucial that in a suitable sense: *the family* $\widehat{D}_{2^{-N}}(f,I)$ *of relative defects is coercive*.

Bound for the relative defect $D_{2^{-N}}(f, I)$.

By inspection, for all n, N,

$$V_{2^{-(n+N)}}(I) = \sum_{j=1}^{2^n} V_{2^{-N}}(I_{n,j})\,. \qquad (5)$$

From (**??**), (**??**), we get

$$\sum_{j=1}^{2^n} D_{2^{-N}}(f, I_{n,j}) = V_{2^{-n}}(f, I) - V_{2^{-(n+N)}}(f, I)\,. \qquad (6)$$

Summing (**??**) over j and taking into account cancellations gives

$$\begin{aligned}\sum_{I_{n,j}} D_{2^{-N}}(I_{n,j}) &= \sum_{n=0}^{\infty} V_{2^{-n}}(f, I) - V_{2^{-(n+N)}}(f, I) \\ &= V(f, I) + \cdots + V_{2^{-(N-1)}}(f, I)\,. \\ &\leq N \cdot V(f, I)\,. \end{aligned} \qquad (7)$$

The Markov bound for $\widehat{D}_n, \mathcal{C}$.

If $|f'| \leq 1$, we have

$$V(f, I) \leq |I| \, .$$

This, together with (**??**) gives

$$\sum_{I_{n,j}} D_{2^{-N}}(I_{n,j}) \leq N \cdot |I| \, . \tag{8}$$

By Markov's inequality, for any $\eta > 0$,

$$\sum_{I_{n,j} \,|\, \widehat{D}_{2^{-N}}(I_{n,j}) > \eta} |I_{n,j}| < \eta^{-1} \cdot N \cdot |I| \, . \tag{9}$$

Equivalently,

$$\mathcal{C}(\{I_{n,j} \,|\, \widehat{D}_{2^{-N}}(I_{n,j}) > \eta\}) < \eta^{-1} \cdot N \cdot |I| \, . \tag{10}$$

Coercivity of relative defects.

Proposition. (Coercivity) Let $|f'| \leq 1$ and assume

$$2^{-N} \leq \epsilon\,,$$

$$\widehat{D}_{2^{-N}}(J) \leq \frac{\epsilon^2}{4}\,. \tag{11}$$

Then

$$\alpha(f, J) \leq \epsilon\,. \tag{12}$$

Proof of theorem. By taking

$$\eta = \frac{\epsilon^2}{4}$$

in (**??**), (**??**), the proposition suffices to complete the proof of the theorem.

To prove the proposition we need a lemma.

Lemma.

$$\widehat{D}_{2^{-N}}(f,J) \geq \left(\frac{1}{|J|} \sum_{j=1}^{2^N} \int_{J_{N,j}} \left| f'_J - f'_{J_{N,j}} \right| \right)^2 . \tag{13}$$

Proof. Recall that

$E_{2^{-N}}(f,J)$

$$= \int_J (f')^2 - (f'_J)^2 - \sum_{j=1}^{2^N} \int_{J_{N,j}} (f')^2 - (f'_{J_{N,j}})^2 .$$

Write the first integral on the right-hand side of (**??**) as a sum of integrals over the intervals $J_{N,j}$ and use (**??**) to get

$$D_{2^{-N}}(f,J) = \sum_{j=1}^{2^N} \int_{J_{N,j}} (f'_J - f'_{J_j})^2 .$$

By the Schwarz inequality, we get (**??**).

□

Proof of Proposition. Without essential loss of generality we consider $J = [0,1]$.

Define an affine linear function ℓ by

$$\ell(x) = f(0) + x(f(1) - f(0)) \,.$$

Since $|f'| \le 1$ and the set $x_j = j \cdot 2^{-N}$ is $\frac{\epsilon}{2}$-dense, it suffices to show

$$|f(x_j) - \ell(x_j)| \le \frac{\epsilon}{2} \qquad \text{(for all } i\text{)} \,.$$

We have $\ell(0) = f(0)$ and by integration,

$$f(x_i) = f(0) + \sum_{j=1}^{i} \int_{J_j} f'_J \, dx \,.$$

$$\ell(x_i) = f(0) + \sum_{j=1}^{i} \int_{J_j} f'_{J_j} \, dx \,.$$

which imply

$$|f(x_i) - \ell(x_i)| \le \sum_{j=1}^{m} \int_{J_j} |f'_J - f'_{J_j}| \, dx \,, \quad (14)$$

Relations (**??**), (**??**) and (**??**) yield the proposition. □

Quantitative differentiation in general.

A general setting for quantitative differentiation is the following:

A class of (doubling) metric measure spaces for which the metric balls can be equipped with an additional structure which we call a *configuration* f.

In the above case $X = \mathbb{R}$, the balls are intervals and the configuration is the function

$$f : I \to \mathbb{R}\,.$$

The configuration f on a ball $B_r(x)$ can be restricted to a sub-ball and then rescaled to unit size.

Energy with coercive relative defects.

There is a scale invariant lower semicontinuous energy functional $\mathcal{E}(f)$ for which the minimal energy configurations with constant energy density can be characterized.

Assume that that the family of relative defects is coercive in the sense in the following sense:

For all $\epsilon > 0$, there exist $\eta(\epsilon) > 0, \delta(\epsilon) > 0$ such that if

$$\widehat{D}_{\delta(\epsilon)}(f, B_r(x)) \leq \eta(\epsilon)\,,$$

then:

f is ϵ-close in the appropriate sense to a minimizer with constant energy density; compare of Proposition (Coercivity).

Remark. Typically, coercivity of relative defects is nontrivial to establish.

Sufficient conditions.

1) There is a **lower semicontinuous** locally defined energy functional $\mathcal{E}(f)$ whose minimizers with constant energy density can be characterized.

2) The associated **family of relative defects is coercive**.

3) There is an **assumed bound** on the configuration $f \,|\, B_1(x)$ which can be shown to imply a bound on $\mathcal{E}(f)$.

Markov and quantitative differentiation.

Given 1)–3), **Markov's inequality** yields a quantitative differentiation theorem.

The theorem states that given $\epsilon > 0$, there is a set of balls $\mathcal{G}_\epsilon$ such that if $B_s(y) \in \mathcal{G}_\epsilon$,

$$f \,|\, B_r(y) \subset B_1(x)$$

is ϵ-close to that of a special minimizer in the appropriate scale invariant sense.

Moreover, the complementary set of balls has at most a definite *finite* measure with respect to $\mathcal{C}$.

In the case we have treated, the *assumed bound* on the configuration is

$$|f'| \leq 1\,.$$

Relation (**??**) is the bound on the integral of the relative defects with respect to $\mathcal{C}$.

The conclusion of the quantiative differention theorem is (**??**).

Below we mention some recent applications.

The goal is to briefly indicate their diversity and some of their special features.

We hope that the discussion is (at least somewhat) intelligible; for additional details see the relevant references.

Generalized Rademacher theorem.

Rademacher's theorem, asserts the almost everywhere differentiability of real valued Lipschitz functions on $\mathbb{R}^n$.

It has been generalized to metric measure spaces which satisfy a doubling condition and a Poincaré inequality; [GAFA,1999].

As is shown there: almost everywhere at the infinitesimal level, any Lipschitz function f looks like a function that is linear in a generalized sense.

As emphasized in [GAFA,1999], the general principle which is responsible for this circumstance is lower semicontinuity of the energy.

The generalized linear structure is an instance of the "special minimizers" which appear on slide 29.

Thus, lower semicontinuity of energy is included in Sufficient Condition 1) on slide 28.

This leads to the appearance of minimizers with special structure in the general context of quantitative differentiation.

In particular, the discussion of [GAFA,1999] can be supplemented by a quantitative differentiation theorem.

In that theorem, coercivity of relative defects amounts to a strengthened version of the Poincaré inequality in which one assumes that the differential of f is bounded in norm and quantitatively small the weak sense.

The inequality asserts that in this case, the left-hand side is quantitatively small as well.

Lipschitz maps to L_1

Lipschitz maps from $f : \mathbb{R} \to L_1$ need not be differentiable anywhere.

However, if $B_1(x)$ is contained in $\mathbb{R}^n$ or the Heisenberg group with its Carnot-Cartheéodory metric, and $f : B_1(x) \to L_1$ is Lipschitz, then the *induced metric* is differentiable in a certain generalized sense; (Cheeger-Kleiner) [Ann. Math. 2010].

This depends on the so-called *cut metric* description of the induced metric.

Roughly speaking, a "cut" is a subset $E \subset B_1(x)$ of the domain of f, which is a top dimensional submanifold, whose boundary has finite area in the appropriate sense; more precisely, a *finite perimeter* subset.

The basic assertion of [CK] is that almost everywhere, at the infinitesimal level E looks like $H \cap B_1(x)$, for some half-space H.

The corresponding quantitative differentiation theorem is proved in (Cheeger-Kleiner-Naor) [arxiv:2011].

There, and in the closely related case of individual sets of finite perimeter, the *assumed bound* is a bound on the *perimeter*; see 3) on slide 28.

On the other hand, the energy $\mathcal{E}$ is a quantity called the *nonmonotonicity*.

The *kinematic formula* is used to show that a bound on the nonmonotonicity follows from a bound on the perimeter.

Remark. In the case of sets of individual finite perimeter, for scaling reasons, one multiplies the measure $\mathcal{C}$ by a factor r^{-1}.

Curvature estimates for Einstein manifolds.

In (Cheeger-Naber) [arxiv:2011], quantitative differentiation is used together with ϵ-regularity theorems of (Cheeger-Colding-Tian) [GAFA 2002] to obtain curvature estimates for Einstein manifolds off sets with quantitatively small volume.

These are quantitative versions of known estimates on Hausdorff dimension or Hausdorff measure.

In this case, the configuration is a (pointed) ball $B_1(x)$ in a manifold with a definite lower bound on Ricci curvature.

The word "pointed" is meant to convey that we only restrict $B_1(x)$ to a concentric ball $B_s(x)$, $s \leq 1$.

The energy $\mathcal{E}$ is minus the log of the volume ratio appearing in the Bishop-Gromov inequality.

Monotonicity of relative defects is (equivalent to) the Bishop-Gromov inequality.

The minimizers with special structure are metric cones.

Coercivity of relative defects is the "almost volume annulus implies almost metric annulus theorem" (Cheeger-Colding) [Ann. of Math., 1996].

Although at *every* point there is a bound on the number of "ϵ-bad scales", in general, the particular collection of such scales will vary from point to point.

This leads to a *key* decomposition into subsets whose points all share the same collection of good and bad scales.

The quantitative differentiation theorem provides a crucial bound on the number of nonempty sets in the decompostion.

The fact that the minimizers with special structure are *cones* plays a crucial role in the applications of the quantitative differentiation theorem.

Remark. All aspects of the preceeding discussion have counterparts in the context of harmonic maps, minimal submanifolds and other equations; (Cheeger-Naber) [to appear].