Appendices to Analysis I, II, and III

Christoph Thiele

Contents

Introduction

Each section is an assignment for the seminar S2B1. Some sections build up on each other, but there are several independent entrance points. If you are interested in one section, you may want to check connections with the nearby sections. Prepare a 2-3 page summary of your topic. You should have a rough draft of the summary and thus lecture two weeks ahead of your lecture and discuss the draft with the organizers two weeks ahead of the lecture (and, if needed, again one week prior to the lecture).

1 Surreal numbers

Definition, order relation, equality relation, and simplicity theorem.

https://www.whitman.edu/documents/Academics/Mathematics/Grimm.pdf Sections 1 and 2.1. Possibly Conway's book in the references for further detail.

2 Addition and Multiplication of surreal numbers

Do addition of surreal numbers, its definition and standard properties. Do the same for multiplication, as far as time allows, but at least as far as done in https://www.whitman.edu/documents/Academics/Mathematics/Grimm.pdf (We mostly need only $1 \times 1 = 1$ to identify the surreal numbers with finite birthday with the dyadic numbers.) Use references therein (Conway) for omitted details.

3 Dyadic fundamental theorem between monotone and convex functions

Let $\mathbb{D}$ be the set of numbers $2^k n$ with $n, k \in \mathbb{Z}$ (Dyadic numbers, the surreal numbers with finite birthday.)

A function $f : \mathbb{D} \to \mathbb{R}$ is called monotone (more precisely monotone increasing, but here we will only have monotone functions that are monotone increasing), if

$$f(2^k n) \leq f(2^k(n+1))$$

for all $n, k \in \mathbb{Z}$. A monotone function $f : \mathbb{D} \to \mathbb{R}$ is called upper semicontinuous if

$$f(2^k n) = \inf_{k' \leq k} \{f(2^k n + 2^{k'})\}$$

for all $k, n \in \mathbb{Z}$. (Note that $2^k n + 2^{k'} = 2^{k'}(2^{k-k'} n + 1)$ is a dyadic number.)

A function $f : \mathbb{D} \to \mathbb{R}$ is called convex, if for all $k, n \in \mathbb{Z}$ we have

$$2f(2^k n) \leq f(2^k(n-1)) + f(2^k(n+1)) \ .$$

Theorem 1. *For each convex function $F : \mathbb{D} \to \mathbb{R}$ and each $k, n \in \mathbb{Z}$,*

$$f(2^k n) := \inf_{k' < k} 2^{-k'}(F(2^k n + 2^{k'}) - F(2^k n)) \tag{1}$$

is a real number, and the thus defined function f is upper semicontinuous monotone.

Proof. Hint for existence of f: Use convexity with $2^k n$ as left endpoint of the three term relation to show that the term on the right hand side of (1) is decreasing as $k' \to -\infty$. To obtain a lower bound, use convexity with $2^k n$ as middle and right point of the convexity relation.

Proof. Hint for monotonicity of f: Show first that for fixed k, the expression

$$F(2^k(n+1)) - F(2^k n)$$

is monotone in n. $\square$

$\square$

Theorem 2. *For each upper semicontinuous monotone function $f : \mathbb{D} \to \mathbb{R}$, and each k, n we have*

$$S(k,n) := \inf_{k'<k} 2^{k'-k} \sum_{m=1}^{2^{k-k'}} f(2^{k'}(2^{k-k'}n + m) \tag{2}$$

is a real number. We have for each $k, n \in \mathbb{Z}$

$$S(k+1,n) = S(k,2n) + S(k,2n+1) \tag{3}$$

There is a unique convex function $F : \mathbb{D} \to \mathbb{R}$ such that $F(0) = 0$ and

$$F(2^k(n+1)) - F(2^k n) = S(k,n)\ .$$

Proof. Hint: To show that $S(k,n)$ is real, show it is bounded below by $f(2^k n)$. To show (3), show that the term on the right-hand-side of (2) is decreasing if k' decreases. Existence of F uses induction on n for fixed k and induction on k. $\square$

Theorem 3. *Let $F : \mathbb{D} \to \mathbb{R}$ be convex and f the upper semicontinuous monotone function f given by Theorem 1. Applying Theorem 2 to f gives a convex function $\tilde{F}$. We then have*

$$\tilde{F}(2^k n) = F(2^k n) - F(0)$$

If time allows (probably not much), comment on the fact that upper semicontinous functions and convex functions $f : \mathbb{D} \to \mathbb{R}$ have unique extensions $\mathbb{R} \to \mathbb{R}$. One has the analogous theory as above for functions $\mathbb{R} \to \mathbb{R}$.

4 The Brenier map

We consider convex and monotone functions in higher dimensions.

A function $\phi : \mathbb{R}^n \to \mathbb{R}$ is called convex if for any $x, y \in \mathbb{R}^n$ and any $\theta \in [0,1]$ we have

$$\phi(\theta x + (1-\theta)y) \leq \theta\phi(x) + (1-\theta)\phi(y)\,.$$

let ϕ be such a convex function. A function $T : \mathbb{R}^n \to \mathbb{R}^n$ is called gradient of ϕ if for every $x, u \in \mathbb{R}^n$

$$\phi(u) \geq \phi(x) + (u-x) \cdot T(x).$$

(There may be more than one gradient)

We have the following generalisation of monotonicity in $\mathbb{R}$

Theorem 4. *If $T : \mathbb{R}^n \to \mathbb{R}^n$ is a gradient of a convex function /phi : $\mathbb{R}^n \to \mathbb{R}$, then for any $k > 0$ and any points $x_1, \ldots x_k$ the following is true. Let $y_1, \ldots y_k$ be the images of $x_1, \ldots x_k$ under D and let $u_1, \ldots u_k$ be a permutation of $y_1, \ldots, y_k$. Then*

$$\sum_{i=1}^{k} \|y_i - x_i\|^2 \leq \sum_{i=1}^{k} \|u_i - x_i\|^2 .$$

Proof. see `http://www-stat.wharton.upenn.edu/~steele/Courses/900/Library/ball-monotone-tra` `pdf` □

Theorem 5. *If mu and ν are probability measures on $\mathbb{R}^n$, ν has compact support and μ assigns no mass to any set of Hausdorff dimension $n - 1$. Then there is a convex function $\phi : \mathbb{R}^n \to \mathbb{R}$ and a gradient T of ϕ such that $\mu(T^{-1}A) = \nu(A)$ for every measurable set $A \subset \mathbb{R}^n$.*

Proof. see `http://www-stat.wharton.upenn.edu/~steele/Courses/900/Library/ball-monotone-tra` `pdf` □

A proof of Brouwer's fixed point theorem is here, https://arxiv.org/pdf/1205.4540, but it is also Ok to assume it known in the likely event that time is short.

5 Stieltjes integral

Let $C^+([0, 1])$ be the set of continuous functions on $[0, 1]$ to $\mathbb{R}_{\geq}0$ such that $f(0) = 0$ and.

Theorem 6. *Let $g : [0, 1] \cap \mathbb{D} \to \mathbb{R}$ be a monotone upper semicontinuous function. Let $f \in C^+([0, 1])$. Then*

$$J(k) := \sum_{n=1}^{2^{-k}} f(2^k(n-1))(g(2^k n) - g(2^k(n-1)) ,$$

defined for negative integers k, is increasing as $k \to -\infty$ and bounded above by

$$(g(1) - g(0)) \sup_{x \in [0,1] \cap \mathbb{D}} f(x) .$$

Proof. Hint... □

We call the limit the Stieltjes integral and write it

$$\int_0^1 f(t) dg(t)$$

For fixed g is is additive in f and takes values in the nonnegative numbers.

Theorem 7. *Let Λ be a map from $f \in C^+([0, 1])$ to $\mathbb{R}_{\geq 0}$ that is additive. Then there is a unique $\lambda \geq 0$ and a unique lower semicontinuous monotone function g with $g(0) = 0$ such that $\Lambda(f) = \lambda f(0) + \int_0^1 f(t)\, dg(t)$*

Proof. Hint: Define λ as the infimum of $\Lambda(f_k)$ where f_0 goes linearly from 1 to 0 on the interval $[0, 2^k]$ and vanishes outside this interval. Define $g(0) = 0$ and $g(1) = \Lambda(1) - \lambda_0$. Define $g(2^k n)$ for other $2^k n$ as the infimum of $\lambda(f_{k'}) - \lambda_0$, where $f_{k'}$ is constant 1 on $[0, 2^k n]$, then decays linearly to 0 until $2^k n + 2^{k'}$ and stays zero from there on. $\square$

If time allows:

By taking difference of two functions in $C^+([0,1])$, extend to $C([0,1])$, the contiuous functions from $[0,1] \to \mathbb{R}$.

By taking differences of g, extend to functions of bounded variation.

6 Variation norm

Let $1 \leq r < \infty$ Define for $s < t$ the r-variation of a function $f : [s,t] \to \mathbb{R}$ as

$$\|f\|_{V^r([s,t])} := \sup_N \sup_{s \leq a_0 < a_1 < \cdots < a_N \leq t} \left(\sum_{n=1}^{N} |f(a_n) - f(a_{n-1})|^r \right)^{\frac{1}{r}}$$

Note that for $r < r'$ we have

$$\|f\|_{V^{r'}(I)} \leq \|f\|_{V^r(I)}$$

Note also that constant functions have variation norm zero. To make the variation norm an actual norm, one can introduce a convention that the functions vanish at the initial point s.

Theorem 8 (homeomoprhism invariance)**.** *Let $f : [a,b] \to \mathbb{R}$ have finite variation and let $g : [c,d] \to [a,b]$ be a monotone increasing bijection. Then*

$$\|f\|_{V^r([a,b])} = \|f \circ g\|_{V^r([c,d])} .$$

Proof. Hint: use the definition on both sides and compare. $\square$

Theorem 9 ((super) additivity)**.** *Let $s < t < u$ and $1 \leq r < \infty$. Then we have*

$$\|f\|^r_{V^r([s,t])} + \|f\|^r_{V^r([s,t])} \leq \|f\|^r_{V^r([s,u])}$$

Proof. Hint. Take near optimal (that is approximate for arbitrary ϵ) testing sequences for the norms on the left-hand-side and combine to a testing sequence for the right hand side. $\square$

Theorem 10 (continuity)**.** *Let $1 \leq r < \infty$ Let f have finite r-variation on $[a,b]$. Then the function*

$$g(x) = \|f\|_{V^r([a,x)}$$

is continuous and monotone.

Proof. For monotonicity use suprema over increasing sets. For lower semicontinuity, approximate $g(x)$ by an instance of a sequence avoiding the point t (use continuity) and show that for every y between the last point of the sequence and x we have $g(y)$ is close to $g(x)$. For upper semi continuity, pick any y_0 to the right of x. Then pick recursively y_{n+1} between x

and y_n so that y_{n+1} is to the left of all the jumps to the right of x of a sequence approximating $g(y_n)$. Use the previous theorem to show that the portion the pieces to the right of x of the approximating sequences converges to 0 as n tends to ∞. Thus we obtain for arbitrary ϵ approximating seqeucne for $g(y_n)$ for which the last jump crossse x from left vto right. Use continuity to replace for sufficiently large n by a jump that lands exactly at x. $\square$

Theorem 11 (comparison with Hölder functions)**.** *Assume $f : [a, b] \to \mathbb{R}$ is continuous, has finite variation norm and is nowhere constant, that is it is not constant on any interval $[u, v]$ with $a \le u < v \le b$. Then the function g^r with g as in the previous theorem is a bijection from $[a, b] \to [0, \|f\|^r_{V^r([a,b])}]$ and the function $f \circ g^{-1}$ is Hölder continuous with exponent $\alpha = 1/r$.*

Proof. Hint: show g is strictly monotone increasing and invoke standard theorems to show that g is a bijection. For Hölder continuity use the additive property of an earlier theorem above. $\square$

7 Young's integral

Let $1 \le r < \infty$ Recall from the previous lecture the r-variation $\|f\|_{V^r(I)}$ of a function $f : I \to \mathbb{R}$. Let $I_0 = [0, 1]$. Recall the dyadic numbers $\mathbb{D}$.

Theorem 12. *Let $1 \le r < \infty$ Let $f : I_0 \to \mathbb{R}$ be continuous and have finite r-variation on I_0. There exists a monotone map $\gamma : \mathbb{D} \cap I_0 \to I_0$ such that $\gamma(0) = 0$ $\gamma(1) = 1$ and*

$$\|f\|_{V^r([\gamma(2^k n), 2^k(n+1))]} \le 2^{-k/r} \|f\|_{V^r(I_0)} .$$

Proof. Use additivity and continuity properties shown in previous lecture to do induction on subdivision of dyadic intervals (subdividing by the intermediate value theorem for continuous functions). $\square$

Theorem 13. *Let $1 \le r < \infty$ Let $f : I_0 \to \mathbb{R}$ be continuous and have finite r-variation on I_0. Let $r < s < \infty$ and $\epsilon > 0$ Then there exists a piecewise linear continuous function g such that*

$$\|f - g\|_{V^s(I_0)]} \le \epsilon \tag{4}$$

Proof. Hint: Use the previous theorem to subdivide I_0 into some 2^{-k} intervals J_n such that $\|f|_{V^r(J_n)]} \le \delta$ for very small δ. Choose piecewise linear g that coincides with f one the boundary points of this subdivision.

To estimate $\|f - g\|_{V^s(I_0)]}$, pick a well (factor two is enough) approximating sequence of jumps. Divide the jumps into A) those which are entirely inside some of the J_n and B) those that go across different J_n.

To estimate the first set of jumps, use the triangle inequality

$$\sum_{j \in A} |(f - g)(x_j) - (f - g)(x_{j-1})|^s$$

$$\le C \sum_{n \in A} |f(x_j) - f(x_{j-1})|^s + C \sum_{n \in A} |g(x_j) - g(x_{j-1})|^s$$

The estimates for the two pieces are similar, e.g. for f

$$\leq \sum_{j\in A} |f(x_j) - f(x_{j-1})|^r \delta^{s-r} \leq \|f\|^r_{V^r(I_0)]} \delta^{s-r}$$

To estimate the jumps across the boundaries, note that $f - g$ vanishes on the boundary points of the J_n.

Each jump across such boundary points can be estimated by the sum of two jumps, namely from the original jump points to the nearest boundary points of some J_n. These two jumps are within individual J_n and can be estimated similarly to the previous terms. □

Theorem 14. *Let $1 \leq p, q, < \infty$ with $1/p + 1/q > 1$. Let $f, g : [0,1] \to \mathbb{R}$ be continuous with $\|f\|_{V^p(I_0)} < \infty$ and $\|g\|_{V^q(I_0)} < \infty$. Assume $f(0) = 0$.*

There exists a $c > 1$ and an $\epsilon > 0$ and a map $\gamma : \mathbb{D} \cap I_0 \to I_0$ such that $\gamma(0) = 0$ $\gamma(1) = 1$ and for $k \leq 0$ and $0 < k \leq 2^{-k}$

$$\|f\|_{V^p([\gamma(2^k n), 2^k(n+1))]} \leq c 2^{\epsilon k} \|f\|_{V^p(I_0)} .$$

$$\|g\|_{V^q([\gamma(2^k n), 2^k(n+1))]} \leq c 2^{\epsilon k} \|f\|_{V^q(I_0)} .$$

$$|\gamma(2^k n) - \gamma(2^k(n+1))| \leq c 2^{\epsilon k}$$

The sequence

$$J(k) := \sum_{n=1}^{2^{-k}} f(\gamma(2^k n))(g(\gamma(2^k n)) - g(\gamma(2^k(n-1)))$$

for $k \leq 0$ has a limit J as $k \to -\infty$ that satisfies

$$|J| \leq C \|f\|_{V^p(I_0)} \|g\|_{V^r(I_0)} \|$$

(Young-Loeve estimate) for some constant C depending only on p and q.

Proof. To construct γ, proceed as in the previous theorem, taking turns between 1) subdividing f 2) subdividing g and 3) cutting the intervals in half.

To show existence of the limit, apply Hölder's inequality to the difference of two consecutive $J(k)$, use the smallness of the variation norm on the small intervals to obtain a geometric decay, so that the series of differences is summable. □

.

Theorem 15. *If g can be approximated in $V^r(I_0)$ norm by piecewise linear functions, then the limit in the last theorem is independent of the choice of γ.*

Proof. Hint: If g is piecewise differentiable, the limit is

$$\int_{I_0} f(t) g'(t)\, dt$$

in the sense of a Riemann integral. In general, approximate by differentiable functions and use the Young-Loeve estimate in the previous theorem. □

We call this limit the Young integral

$$\int_{I_0} f dg$$

8 A rough integral

We wish to extend the theory of Young's integral to r-variation with $2 \leq r < 3$. (An important example is Brownian motion, which almost surely on a bounded interval has finite r- variation with $r > 2$.)

We start with a simple example. We replace the Riemann sum for $J(k)$ by a trapezoidal sum

$$J(k) := \sum_{n=1}^{2^{-k}} \frac{1}{2}(f(\gamma(2^k(n-1))) + f(\gamma(2^k n)))(g(\gamma(2^k n)) - g(\gamma(2^k(n-1)))$$

and look at the special case $f = g$. In this case we obtain by the binomial formula

$$J(k) := \sum_{n=1}^{2^{-k}} \frac{1}{2}(g(\gamma(2^k n))^2 - g(\gamma(2^k(n-1)))^2),$$

which telescopes into

$$\frac{1}{2}(g(1)^2 - g(0)^2).$$

This is independent of k and thus $J(k)$ converges trivially to this same expression. We dont even need any regularity assumption on f other than being able to evaluate at every point, which is e.g. possible for every continuous function.

This example can be generalized, keeping the trapezoidal rule but replacing f by $F(g)$ for some fairly smooth F.

Theorem 16. *Let $F : \mathbb{R} \to \mathbb{R}$ be twice continuously differentiable. let $1 \leq r < 3$ and $g : I_0 \to \mathbb{R}$ be continuous with finite r variation. Let γ be a map $\mathbb{D} \cap I_o \to I_0$ as in the definition of the Young integral. Then*

$$J(k) := \sum_{n=1}^{2^{-k}} \frac{1}{2}(F(g(\gamma(2^k(n-1)))) + F(g(f(\gamma(2^k n))))(g(\gamma(2^k n)) - g(\gamma(2^k(n-1)))$$

converges for k to $-\infty$.

Proof. Hint: Define $K(s,t) := \int_s^t (g(u) - g(s))dg(u)$ using the example and trivial telescoping of the constant integral. Prove

$$K(s,u) - K(s,t) - K(t,u) = (g(t) - g(s))(g(u) - g(t))$$

and

$$|K(s,t)| \leq \|g\|^2_{V^r([s,t])}$$

Define the sequence (motivated by Taylor expansion) in analogy to the previous lecture

$$J(k) :=$$

$$\sum_{n=1}^{2^{-k}} F(g(\gamma(2^k n)))(g(\gamma(2^k n)) - g(\gamma(2^k(n-1))) + F'(g(\gamma(2^k n)))K(\gamma(2^k n), \gamma(2^k(n-1)))$$

Show that the series converges similarly as in the existence for the Young integral. □

Note that in case $r < 2$, one can integrate using the Young integral and that F is Lipschitz so that $F \circ g$ is also in V^r. The passage beyond $r = 2$ is important for example in probability, as the Brownian path is in V^r for $r > 2$). One can adapt the above theorem to the case of general f. One demands knowledge of $\int f dg$ and deduce $\int F(f) dg$. (Such knowledge can for example be given in probability, when f and g are independent Brownian paths, then one can say something about the integral in the almost surely sense.) Let $1 \leq r < 3$. Let $f, g : [I_0] \to \mathbb{R}$ have finite r-variation. Let $K : I_0 \times I_0 \to \mathbb{R}$ satisfy for $s < t < u$

$$K(s,u) - K(s,t) - K(t,u) = (f(t) - f(s))(g(u) - g(t))$$

and the variation type bound

$$\sup_{x \in I} |G(0,t)| + \sup_N \sup_{s \leq a_0 < a_1 < \cdots < a_N \leq t} \left(\sum_{n=1}^{N} |G(a_{n-1}, a_n)|^{\frac{r}{2}} \right)^{\frac{2}{r}} < \infty$$

Then one can develop an integration along the above lines, demanding

$$\int_s^t f dg := K(s,t)$$

and writing a formula for $J(k)$ that expands $F(f)$ in Taylor series and inserts the knowledge of K for the linear terms. The proof goes along the lines of the above theorem. See for further material e.g. `https://www.hairer.org/notes/RoughPaths.pdf`

9 Hausdorff Momentensatz

Define for a sequence $m = (m_n)_{n \in \mathbb{N}}$ the (negative) discrete derivative

$$\Delta m(n) = m(n) - m(n+1) \ .$$

Define recursively powers of the discrete derivative by $\Delta^0 m = m$ and

$$\Delta^{k+1} m = \Delta(\Delta^k m) \, .$$

Theorem 17. *Let $m = (m_n)_{n \in \mathbb{N}}$ be a sequence of real numbers. There exists a nonnegative measure μ on $[0,1]$ with $m_n = \int_0^1 x^n d\mu$ if and only if $\Delta^k m(n) \geq 0$ for all $k, n \in \mathbb{N}$.*

Reference: Hausdorff's paper

"Momentprobleme für ein endliches Intervall" pages 220-225 `https://link.springer.com/article/10.1007/BF01175684`

10 Absolutely monotone functions

Theorem 18. *Assume $f : [0,T] \to \mathbb{R}$ is infinitely often continuously differentiable and the n-th derivative ($f^{(0)} = f$) satisfies*

$$f^{(n)}(x) \geq 0$$

for all $x \in [0,T]$ and all $n \geq 0$. Then f equals its Taylor series about 0, which has radius of convergence at least T.

Proof. Hint: Using Taylor expansion with remainder term, and nonnegativity of all derivatives, show that each Taylor polynom remains below the function to the right of the expanision point. Use this to obtain upper bounds on all derivatives. use thes upper bounds to show that the remainder term of the Taylor expansion tends to zero. □

Note that if f has only nonnegative derivatives, then the derivatives of g defind by $g(t) = f(-t)$ have alternating signs.

Theorem 19 (Bernstein-Widder)**.** *let $f : [0, \infty] \to \mathbb{R}$ be infinitely often continuously differentiable and assume its derivatives have alternating signs. Then there is a nonnegative finite measure μ on $[0, \infty]$ such that*

$$f(x) = \int_0^\infty e^{-tx} d\mu(t)$$

Proof. This is reduced to the Hausdorff Momentensatz. see: Rene Schilling, Renming Song and Zoran Vondraček (2010). Bernstein functions. De Gruyter. □

11 Hilbert spaces through parallelogram law

Theorem 20. *A (real or complex) Banach space is induced by inner product and thus a Hilbert space if and only if the norm satisfies the parallelogram law.*

Proof. This is the Jordan-von Neumann theorem. The hard part is to show the necessary properties of the inner product given through polrization from the parallelogram law.

See https://matthewhr.wordpress.com/wp-content/uploads/2012/09/jordan-von-neumann-theorem.pdf of the original reference to Jordan and von Neumann

Annals of Math 1935. □

If time allows, discuss type and cotype as generalizations of the parallelogram law.

12 Delta calculus 1

Prove the following theorems.

Theorem 21. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a continuously differentiable function such that for all $x \in \mathbb{R}^n$ with $f(x) = 0$ we have $\mathrm{rank} \nabla \mathrm{f(x)} = \mathrm{m}$. Let $g : \mathbb{R}^n \to \mathbb{R}$ be continuous with compact support.*

Then the limit

$$\lim_{\epsilon \to 0} \int_{\mathbb{R}^n} g(x) \epsilon^{-m} e^{-\epsilon^{-2}\pi |f(x)|^2} \, dx \tag{5}$$

exists.

We denote the limit as

$$\int_{\mathbb{R}^n} g(x) \delta(f(x)) \, dx$$

or, to explicitly notate the dimension of the range of f,

$$\int_{\mathbb{R}^n} g(x) \delta_m(f(x)) \, dx$$

Theorem 22. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a continuously differentiable function such that for all $x \in \mathbb{R}^n$ with $f(x) = 0$ we have* $\mathrm{rank}\nabla\mathrm{f}(\mathrm{x}) = \mathrm{m}$. *Let $g_1, g_2 : \mathbb{R}^n \to \mathbb{R}$ be two continuous functions with compact support such that for all x with $f(x) = 0$ we have $g_1(x) = g_2(x)$. Then*

$$\int_{\mathbb{R}^n} g_1(x)\delta(f(x))\,dx = \int_{\mathbb{R}^n} g_2(x)\delta(f(x))\,dx$$

13 Delta calculus 2

Prove the following theorems.

For a continuously differentiable $f : \mathbb{R}^n \to \mathbb{R}^m$, recall that the matrix $\nabla f \nabla f^T$ is an $m \times m$ matrix. Define $Jf = \sqrt{\det \nabla f \nabla f^T}$

Theorem 23. *Let $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}^m$ becontinuously differentiable functions with the same null set such that for all x in this null set we have* $\mathrm{rank}\nabla\mathrm{f}_1(\mathrm{x}) = \mathrm{rank}\nabla\mathrm{f}_2(\mathrm{x}) = \mathrm{m}$. *Let $g : \mathbb{R}^n \to \mathbb{R}$ be continuous with compact support. Then*

$$\int_{\mathbb{R}^n} g(x)Jf_1(x)\delta(f_1(x))\,dx = \int_{\mathbb{R}^n} g(x)Jf_2(x)\delta(f_2(x))\,dx$$

Under the assumptions of the definition of the delta integral, we call the integral

$$\int_{\mathbb{R}^n} g(x)Jf(x)\delta(f(x))\,dx$$

the surface integral of g over the null set of f.

14 At least two completely different solutions to the Basel problem.

Such as from the channel one blue three brown via sums on circles, or via Wallis product and `https://scholar.google.com/scholar?hl=de&as_sdt=0%2C5&q=w%C3%A4stlund+elementary&btnG=` or via Plancherel for the function on $[-\pi, \pi]$

$$f(x)\sum_{n=1}^{N} \frac{1}{-n}e^{inx} + \frac{1}{in}e^{-inx}$$

(The derivative f' has an explicit formula by geometric series, and one can approximate the odd function f away from 0 by integrating from pi)