LECTURE 10, 4/20/2022

AREA UNDER G_f. THE NEWTON INTEGRAL. INTEGRATION BY PARTS AND BY SUBSTITUTION

• *What are antiderivatives good for?* For computing areas A_f of domains D_f under graphs G_f of functions $f\colon I \to \mathbb{R}$ defined on nontrivial intervals $I \subset \mathbb{R}$. Recall that

$$G_f = \{(x,\ f(x)) \mid x \in I\} \subset \mathbb{R}^2$$

and that $I(c,d) \subset \mathbb{R}$ denotes the closed interval with the endpoints $c, d \in \mathbb{R}$. We define the *domain under* G_f as the plane set

$$D_f := \{(x,\ y) \mid x \in I \wedge y \in I(0,\ f(x))\} \subset \mathbb{R}^2$$

(so $G_f \subset D_f$). But what exactly is the plane *area* $A_f \in \mathbb{R}$ of D_f? Two remarks are in order. First, A_f will be a *signed area*, the parts of D_f below the x-axis will contribute to A_f negatively and those above the x-axis positively. Second, A_f has not yet been defined in our lectures and for us it does not yet exist as a rigorous mathematical object. We bring it in existence only by a precise definition.

This perspective on A_f differs from that of physicists. For continuous f they may measure A_f (say when $f \ge 0$) as follows. They draw D_f on a sheet of paper and draw on it also one $1\,\mathrm{cm} \times 1\,\mathrm{cm}$ square S. They cut by scissors both D_f and S from the sheet and weight them. They get the area

$$A_f \approx \frac{\mathrm{weight}(D_f)}{\mathrm{weight}(S)}\,\mathrm{cm}^2\ .$$

The only thing we mathematicians can say to this is to remark that the "D_f" cut from the paper and the plane set $D_f \subset \mathbb{R}^2$ are two completely different things. For mathematicians A_f is the area of $D_f \subset \mathbb{R}^2$ and it does not exist until they define it, and then A_f is what they have defined it to be, which can be done is several ways.[1] We give two definitions of A_f in Definition 5 and a third one in Definition 6.

• *Riemann sums and telescoping* PF *sums for A_f.* Still, we want to approximate or estimate A_f somehow, whatever it is or will be. We consider two setups, with functions $f\colon I \to \mathbb{R}$ where I is an interval. The first one, in this passage, is of

continuous functions $f\colon [a,\, b] \to \mathbb{R}$, for real numbers $a < b$.

We select a *partition* $P = (a_0, a_1, \dots, a_k)$ of $[a, b]$, $k \in \mathbb{N}$ and $a = a_0 < a_1 < \cdots < a_k = b$, and define the corresponding *Riemann sum* as

$$R(P,\, \overline{t},\, f) := \sum_{i=1}^{k} (a_i - a_{i-1}) \cdot f(t_i) \ ,$$

where $\overline{t} = (t_1, \dots, t_k)$ with $t_i \in [a_{i-1}, a_i]$ are any k *test points of P*. This definition applies to any function $f\colon [a, b] \to \mathbb{R}$, not only to continuous ones. Note that $R(P, \overline{t}, f)$ is the signed area of the *bar graph* $B_f \subset \mathbb{R}^2$ consisting of k bars (rectangles),

$$B_f := \bigcup_{i=1}^{k} [a_{i-1},\, a_i] \times I(0,\, f(t_i)) \ .$$

Bars under the x-axis (i.e., with $f(t_i) < 0$) contribute negative areas. We define the *norm of P* as

$$\Delta(P) := \max(\{a_i - a_{i-1} \mid i = 1, 2, \dots, k\}) \ .$$

[1] This is at odds with teaching of elementary and high school mathematics on areas of plane figures, which they view as physical quantities. In high school this is appropriate and OK but we are on University now.

The next proposition shows that all partitions with small norm and arbitrary test points yield similar Riemann sums.

Proposition 1 (on Riemann sums) *Let $a, b \in \mathbb{R}$ with $a < b$ and $f\colon [a,b] \to \mathbb{R}$ be a continuous function. Then $\forall \varepsilon \, \exists \delta$ such that if P and Q are partitions of $[a,b]$ with both norms*

$$\Delta(P),\ \Delta(Q) < \delta$$

and $\bar{t}$ and $\overline{u}$ are two tuples of test points of P and Q, respectively, then

$$|R(P,\, \bar{t},\, f) - R(Q,\, \overline{u},\, f)| < \varepsilon\ .$$

Proof. Let a, b and f be as stated, and let an ε be given. By Theorem 15 in the last lecture we know that f is uniformly continuous and therefore there is a δ such that for any $c, d \in [a,b]$, $|c - d| < \delta \Rightarrow |f(c) - f(d)| < \varepsilon/2(b-a)$. Now suppose that $P = (a_0, a_1, \ldots, a_k)$ is a partition of $[a,b]$ with test points $\bar{t}$, that $Q = (b_0, b_1, \ldots, b_l)$ is a partition of $[a,b]$ with test points $\overline{u}$, and that both $\Delta(P), \Delta(Q) < \delta$. We assume additionally that $P \subset Q$, i.e., that $a_0 = b_{i_0} = a$, $a_1 = b_{i_1}$, $\ldots$, $a_k = b_{i_k} = b$ for some indices $i_0 = 0 < i_1 < \cdots < i_k = l$. Later we reduce general partitions P

and Q to this case. We have that

$$\begin{aligned}
&|R(P,\,\bar{t},\,f) - R(Q,\,\overline{u},\,f)| \\
&= \Big|\sum_{i=1}^{k}(a_i - a_{i-1})\cdot f(t_i) - \sum_{i=1}^{l}(b_i - b_{i-1})\cdot f(u_i)\Big| \\
&= \Big|\sum_{r=1}^{k}\sum_{j=i_{r-1}+1}^{i_r}(b_j - b_{j-1})\cdot(f(t_r) - f(u_j))\Big| \\
&\overset{|t_r - u_j|<\delta \text{ and } \Delta \text{ ineq.}}{<} \sum_{r=1}^{k}\sum_{j=i_{r-1}+1}^{i_r}(b_j - b_{j-1})\cdot\varepsilon/2(b-a) \\
&= (b-a)\cdot\varepsilon/2(b-a) = \varepsilon/2\ .
\end{aligned}$$

If P and Q are general partitions of $[a,b]$ with respective test points $\bar{t}$ and $\overline{u}$ and with $\Delta(P), \Delta(Q) < \delta$, we set $R := P \cup Q$ (then also $\Delta(R) < \delta$) and take arbitrary test points $\overline{v}$ of R. Since $P \subset R$ and $Q \subset R$, we get by the previous case that

$$\begin{aligned}
&|R(P,\,\bar{t},\,f) - R(Q,\,\overline{u},\,f)| \leq \\
&\leq |R(P,\,\bar{t},\,f) - R(R,\,\overline{v},\,f)| + |R(R,\,\overline{v},\,f) - R(Q,\,\overline{u},\,f)| \\
&< \varepsilon/2 + \varepsilon/2 = \varepsilon\ .
\end{aligned}$$

□

Since for small $\Delta(P)$ the bar graph B_f closely approximates the domain D_f, one can expect that $R(P, \bar{t}, f) \to A_f$ as $\Delta(P) \to 0$. We define this limit formally.

Definition 2 (limits of Riemann sums) *Let $a, b, L \in \mathbb{R}$, $a < b$ and $f \colon [a, b] \to \mathbb{R}$ be a function, not necessarily continuous. If for any sequences (P_n) of partitions P_n of $[a, b]$ and $(\overline{t(n)})$ of tuples $\overline{t(n)}$ of test points of P_n it is true that*

$$\lim \Delta(P_n) = 0 \Rightarrow \lim R(P_n, \overline{t(n)}, f) = L ,$$

we write $\lim_{\Delta(P)\to 0} R(P, \bar{t}, f) = L$ and say that the Riemann sums of f have the limit L.

These limits are unique by definition and below we easily deduce from Proposition 1 that for continuous functions they always exist.

Corollary 3 (limits of R. sums exist) *For every continuous function $f \colon [a, b] \to \mathbb{R}$, $a, b \in \mathbb{R}$ with $a < b$, the (finite) limit*

$$\lim_{\Delta(P)\to 0} R(P, \bar{t}, f) \in \mathbb{R}$$

exists.

Proof. Let f, a and b be as stated, and let (P_n) be an arbitrary sequence of partitions of the interval $[a, b]$ with respective test points $\overline{t(n)}$ and such that $\lim \Delta(P_n) = 0$. By Proposition 1 the sequence $(R(P_n, \overline{t(n)}, f))$ is Cauchy and therefore it has a limit $L \in \mathbb{R}$. If (Q_n) and $\overline{u(n)}$ is another sequence of partitions of $[a, b]$ with respective test points $\overline{u(n)}$ and with $\lim \Delta(Q_n) = 0$, then by Proposition 1,

$$\lim_{n\to\infty} \big(R(P_n, \overline{t(n)}, f) - R(Q_n, \overline{u(n)}, f)\big) = 0 .$$

Therefore also $\lim R(Q_n, \overline{u(n)}, f) = L$. □

However, in this lecture we are more interested in Newton's approach to the areas A_f. We express the summands $(a_i - a_{i-1}) \cdot f(t_i)$ in Riemann sums in terms of any PF F of the continuous f as follows; we know that F exists by the last theorem in the previous lecture. Let $P = (a_0, a_1, \ldots, a_k)$ be any partition of $[a, b]$. We use Lagrange's mean value theorem for F and every interval $[a_{i-1}, a_i]$:

$$\frac{F(a_i) - F(a_{i-1})}{a_i - a_{i-1}} = F'(c_i) = f(c_i)$$

for some point $c_i \in (a_{i-1}, a_i)$. Thus

$$\begin{aligned} F(b) - F(a) &= \sum_{i=1}^{k}(F(a_i) - F(a_{i-1})) = \sum_{i=1}^{k}(a_i - a_{i-1}) \cdot f(c_i) \\ &= R(P,\, \overline{c},\, f)\ , \end{aligned}$$

with the test points $\overline{c} = (c_1, \ldots, c_k)$ of P. In view of Proposition 1 we get the following equality.

Corollary 4 (Riemann = Newton) *Let $a < b$ be real numbers, let $f\colon [a, b] \to \mathbb{R}$ be a continuous function and let $F\colon [a, b] \to \mathbb{R}$ be a primitive of f. Then*

$$\lim_{\Delta(P)\to 0} R(P,\, \overline{t},\, f) = F(b) - F(a)\ .$$

Proof. Let a, b, f and F be as stated, and let (P_n) be any sequence of partitions of $[a, b]$ with test points $\overline{t(n)}$ and such that $\lim \Delta(P_n) = 0$. We know by the above argument that there exist test points $\overline{c(n)}$ of P_n such that, for every n.

$$F(b) - F(a) = R(P_n,\, \overline{c(n)},\, f)\ .$$

Hence, by the arithmetic of limits of sequences,

$$\begin{aligned}
&\lim R(P_n,\, \overline{t(n)},\, f) \\
&= \underbrace{\lim \big(R(P_n,\, \overline{t(n)},\, f) - R(P_n,\, \overline{c(n)},\, f)\big)}_{= 0 \text{ by Prop. 1}} \\
&+ \lim \underbrace{R(P_n,\, \overline{c(n)},\, f)}_{=F(b)-F(a)} \\
&= 0 + F(b) - F(a) = F(b) - F(a)\ .
\end{aligned}$$

Thus we get the stated limit. □

Now we can give two definitions of the area A_f of the domain D_f under G_f for any continuous function $f\colon [a,b] \to \mathbb{R}$. By the previous corollary they give for A_f the same value.

Definition 5 (area under graph) *Let $f\colon [a,b] \to \mathbb{R}$, for real numbers $a < b$, be a continuous function and $D_f \subset \mathbb{R}^2$ be the domain under its graph G_f, as defined earlier. One can define the area $A_f \in \mathbb{R}$ of D_f in two ways.*

1. *(I. Newton) Set $A_f := F(b) - F(a)$ for any antiderivative $F\colon [a,b] \to \mathbb{R}$ of f.*
2. *(B. Riemann) Set $A_f := \lim_{\Delta(P)\to 0} R(P, \overline{t}, f)$ (see Definition 2).*

At first look these two definitions appear very differently, but we know well from Corollary 4 that A_f is the same in both. The former is considerably simpler than the latter, but the latter works in certain cases when the former does not work. Later we will see that the scopes of both definitions are in fact incomparable.

For example, if $f(x) = x^2\colon [-1,1] \to \mathbb{R}$ then $F(x) = x^3/3$ is a primitive of f on $[-1,1]$. By Newton's definition the area of the domain $D_f = \{(x,y) \mid -1 \le x \le 1 \wedge 0 \le y \le x^2\}$ equals

$$A_f = F(1) - F(-1) = \frac{1^3}{3} - \frac{(-1)^3}{3} = \frac{2}{3}\ .$$

• *The Newton integral.* Now we consider the second setup of functions $f\colon I \to \mathbb{R}$, namely of functions

$$f\colon (a,\,b) \to \mathbb{R}, \text{ for real } a < b, \text{ that have a primitive function } F\ .$$

Definition 6 (Newton integral) *Let $a, b \in \mathbb{R}$ with $a < b$ and $F, f\colon (a,b) \to \mathbb{R}$ be functions such that F is a primitive of f. We define the Newton integral of f over the interval (a,b) as the difference*

$$(\mathrm{N})\int_a^b f = F(b) - F(a) := \lim_{x\to b} F(x) - \lim_{x\to a} F(x)\ ,$$

if the last two limits exist and are finite. Then we define the area A_f of the domain D_f under G_f as

$$A_f := (\mathrm{N})\int_a^b f\ .$$

By now it is clear that above we need not use one-sided limits. Since any two primitives F_1 and F_2 of f differ only by a constant shift, $F_1 = F_2 + c$, the value of $(\mathrm{N})\int_a^b f$, if it exists, is independent of the choice of F. We explain why this latter setup for A_f with functions $f\colon (a,b) \to \mathbb{R}$ is strictly more general than the former one with continuous $f\colon [a,b] \to \mathbb{R}$. If $f\colon [a,b] \to \mathbb{R}$ is continuous, by the last theorem in Lecture 9 f has a primitive function $F\colon [a,b] \to \mathbb{R}$. This F is continuous and therefore

$$\lim_{x \to a} F(x) = F(a) \ \text{ and } \ \lim_{x \to b} F(x) = F(b)\ .$$

So the area A_f of D_f in the former setup (the first definition in Definition 5) now equals to its third definition above:

$$A_f = F(b) - F(a) = (\mathrm{N}) \int_a^b f\ .$$

But the situation in Definition 6 is strictly more general than the former setup because if a function $f\colon (a,b) \to \mathbb{R}$ is such that it has a primitive $F\colon (a,b) \to \mathbb{R}$ then f need not be continuous. Even if f is continuous and F is extended to $F\colon [a,b] \to \mathbb{R}$ by limits at a and b, then the derivatives $F'(a)$ and $F'(b)$ need not exist and f cannot be extended to a and b. As a final remark on the Newton integral we say that also the name of G. W. Leibniz should be attached to it but we want to keep the term short.

If for a function $f\colon (a,b) \to \mathbb{R}$, where $a < b$ are real numbers, the Newton integral $(\mathrm{N}) \int_a^b f$ exists, we say that the function f is *Newton-integrable (on (a,b))* and write that

$$f \in \mathrm{N}(a,b)\ .$$

It is easy too see that if $f \in \mathrm{N}(a,b)$ then $f \in \mathrm{N}(c,d)$ for any numbers $c < d$ in the interval (a,b). We prove monotonicity of the Newton integral.

Proposition 7 (monotonicity of the (N) $\int$**)** *If* $f, g \in \mathrm{N}(a,b)$ *and* $f \le g$ *on* (a,b) *then*

$$(\mathrm{N})\int_a^b f \le (\mathrm{N})\int_a^b g\ .$$

Proof. Let F and G be the respective primitives of f and g on (a,b). We take any numbers $c < d$ in (a,b) and use the Lagrange mean value theorem for the function $F - G$ and interval $[c,d]$. We get that for some point $e \in (c,d)$,

$$\begin{aligned}(F(d) - G(d)) - (F(c) - G(c)) &= (F - G)'(e) \cdot (d - c) \\ &= (F'(e) - G'(e)) \cdot (d - c) \\ &= (f(e) - g(e)) \cdot (d - c) \le 0\ .\end{aligned}$$

Hence $F(d) - F(c) \le G(d) - G(c)$. This inequality is preserved under the limit transitions $c \to a$ and $d \to b$ and we get the stated inequality between both Newton integrals. □

We give two examples of Newton integrals:

$$(\mathrm{N})\int_0^1 \sqrt{x} = \frac{2 \cdot 1^{3/2}}{3} - \frac{2 \cdot 0^{3/2}}{3} = \frac{2}{3}$$

but

$$(\mathrm{N})\int_0^1 \frac{1}{x} = \log 1 - \log 0 = 0 - (-\infty) = ?$$

does not exist because the limit of the primitive $\log x$ at 0 is not finite.

• *Proof of the second case of l'Hospital's rule.* As an application of the Newton integral we prove the remaining case of l'Hospital's rule for $\lim_{x\to A} g(x) = \pm\infty$ (Condition 2 in Theorem 7 in Lecture 8). We prove first an asymptotics for Newton integrals.

Proposition 8 (asymptotics of (N) $\int$) *We assume that $f, g \in \mathrm{N}(a,b)$, $g > 0$ on (a,b), that $f(x) = o(g(x))$ $(x \to a)$ and that $\lim_{x\to a} (\mathrm{N}) \int_x^b g = +\infty$. Then*

$$(\mathrm{N}) \int_x^b f = o\left((\mathrm{N}) \int_x^b g\right) \quad (x \to a)\ .$$

Proof. Let an ε be given. By the assumption of the first o there exists a $\delta \le b - a$ such that $x \in (a, a+\delta) \Rightarrow |f(x)| < \frac{\varepsilon}{2} \cdot g(x)$. By the assumption of the limit $+\infty$ there exists a $\theta < \delta$ such that $x \in (a, a+\theta) \Rightarrow |(\mathrm{N}) \int_{a+\delta}^b f| < \frac{\varepsilon}{2} \cdot (\mathrm{N}) \int_x^b g$. Thus if $x \in (a, a+\theta)$ then

$$\begin{aligned} \left|(\mathrm{N}) \int_x^b f\right| &= \left|(\mathrm{N}) \int_x^{a+\delta} f + (\mathrm{N}) \int_{a+\delta}^b f\right| \\ &\overset{\Delta \text{ ineq.}}{\le} \left|(\mathrm{N}) \int_x^{a+\delta} f\right| + \left|(\mathrm{N}) \int_{a+\delta}^b f\right| \\ &\overset{\text{both } \Rightarrow\text{s and Prop. 7}}{<} \frac{\varepsilon}{2} \cdot (\mathrm{N}) \int_x^{b \text{ or } a+\delta} g + \frac{\varepsilon}{2} \cdot (\mathrm{N}) \int_x^b g \\ &= \varepsilon \cdot (\mathrm{N}) \int_x^b g\ . \end{aligned}$$

□

Theorem 9 (l'Hospital's rule, Condition 2) *Let* $A \in \mathbb{R}$. *Let for a* δ *functions* $f, g\colon P^+(A,\delta) \to \mathbb{R}$ *have on* $P^+(A,\delta)$ *finite derivatives,* $g' \neq 0$ *on* $P^+(A,\delta)$*, and let* $\lim_{x\to A} g(x) = \pm\infty$. *Then*

$$\lim_{x\to A} \frac{f(x)}{g(x)} = \lim_{x\to A} \frac{f'(x)}{g'(x)}$$

if the last limit exists. This theorem also holds for left neighborhoods $P^-(A,\delta)$*, ordinary neighborhoods* $P(A,\delta)$ *and for* $A = \pm\infty$.

Proof. Let A, δ, f and g be as stated and let $A \in \mathbb{R}$. We assume that $\lim_{x\to A} g(x) = +\infty$ and $g > 0$ on $(A, A+\delta)$, the case with the limit $-\infty$ is treated similarly. Let $\lim_{x\to A} f'(x)/g'(x) =: L \in \mathbb{R}^*$. We assume first that $L = 0$, i.e., $f'(x) = o(g'(x))$ $(x \to A)$. We fix a $\theta < \delta$ and get by the previous theorem that

$$(\mathrm{N}) \int_x^\theta f' = o\left((\mathrm{N}) \int_x^\theta g' \right) \quad (x \to A)\ ,$$

which gives that $f(x) = f(\theta) - o(1)(g(\theta) - g(x))$. Thus $f(x)/g(x) = f(\theta)/g(x) + o(1)(1 - g(\theta)/g(x)) = o(1) + o(1)(1 - o(1)) = o(1)$ and $\lim_{x\to A} f(x)/g(x) = 0 = L$.

Let $L \in \mathbb{R}$. But then with $h(x) := f(x) - Lg(x)$ we have that $\lim_{x\to A} h'(x)/g'(x) = 0$ and therefore, by the just proved case,

$$0 = \lim_{x\to A} \frac{h(x)}{g(x)} = \lim_{x\to A} \frac{f(x)}{g(x)} - L$$

and $\lim_{x\to A} f(x)/g(x) = L$. If $L = +\infty$ then $\lim_{x\to A} g'(x)/f'(x) = 0^+$. Thus by the previous case $\lim_{x\to A} g(x)/f(x) = 0^+$ and we get that $\lim_{x\to A} f(x)/g(x) = +\infty$.

For the left deleted neighborhoods $P^-(A,\delta)$ and for two-sided neighborhoods $P(A,\delta)$ the proofs are similar, and for $A=\pm\infty$ we use the substitution $x:=1/y$ as in the $\frac{0}{0}$ case. $\square$

We took the previous proof of the $\frac{\infty}{\infty}$ case of l'Hospital's rule from pp. 206–7 of the textbook I. I. Ljaško, V. F. Emel'janov and A. K. Bojarčuk, *Osnovy klassičeskogo i sovremennogo Matematičeskogo Analiza* (Kiev, 1988).

• *The Stirling formula.* One can prove the Stirling asymptotic formula

$$1\cdot 2\cdot\ldots\cdot n=n!\sim\sqrt{2\pi n}\left(\frac{n}{\mathrm{e}}\right)^n\quad(n\to\infty)$$

by using only the Newton integral (but it is not too simple), for details see MK, *The Newton integral and the Stirling formula*, `https://arxiv.org/abs/1907.02553`.

In the next three passages we give three (or four or five) results by which primitives can be computed or can be shown not to exist. Unlike computing the derivative, which for any function given by a formula is (using the rules given earlier) fairly straightforward, computing an antiderivative, when it exists, may be a complex task. For more information see the article `https://en.wikipedia.org/wiki/Risch_algorithm` on the *Risch algorithm*.

• *The Darboux property.* A function $f\colon I\to\mathbb{R}$, defined on an interval $I\subset\mathbb{R}$, has the *Darboux property* (or is *Darboux*) if it attains every intermediate value: if $a<b$ are in I and c is such that $f(a)<c<f(b)$ or $f(a)>c>f(b)$ then $c=f(d)$ for some $d\in(a,b)$. We proved earlier (Theorem 8 in Lecture 6) that continuous functions are Darboux. Now we extend it to a larger class of functions.

Theorem 10 (derivatives are Darboux) *Any function $f\colon I \to \mathbb{R}$, defined on an interval $I \subset \mathbb{R}$, with a primitive function has the Darboux property.*

Proof. We assume that $f\colon [a,b] \to \mathbb{R}$, where $a < b$ are real numbers, has a primitive function $F\colon [a,b] \to \mathbb{R}$ and that $f(a) < c < f(b)$, the case $f(a) > c > f(b)$ is treated similarly. We consider the function

$$G(x) := F(x) - cx\colon [a,\, b] \to \mathbb{R}\ .$$

It has on $[a,b]$ the finite derivative $G'(x) = F'(x) - c = f(x) - c$. In particular, G is continuous. By an earlier theorem (Theorem 13 in Lecture 6), G attains at some $d \in [a,b]$ its minimum value. From

$$G'(a) = f(a) - c < 0 \ \text{ and } \ G'(b) = f(b) - c > 0$$

it follows (by Proposition 5 in Lecture 8) that $d \in (a,b)$. By another earlier theorem (Theorem 4 in Lecture 7), $f(d) - c = G'(d) = 0$, so that $f(d) = c$. $\square$

Since every continuous function has a primitive function and since there exist non-continuous functions which have primitives (Proposition 18 in Lecture 7), the previous class of functions with the Darboux property is strictly larger than the class of continuous functions. The theorem is usually used in reverse: if a function does not have the Darboux property then it has no primitive function. For example, the signum function $\mathrm{sgn}(x)$ is not Darboux on any nontrivial interval $I \ni 0$ and therefore it does not have primitive there.

Recall that the notation

$$F = \int f$$

for two functions $F, f \colon I \to \mathbb{R}$ means that F is a primitive of f. A simple but useful result says that taking antiderivatives is a linear operation.

Proposition 11 (linearity of $\int$) *Suppose that $f, g \colon I \to \mathbb{R}$ are functions defined on a nontrivial interval $I \subset \mathbb{R}$ and that $a, b \in \mathbb{R}$. Then*

$$\int (af + bg) = a \int f + b \int g \ ,$$

meaning that if F, resp. G, is an antiderivative of f, resp. g, then $aF + bG$ is an antiderivative of $af + bg$.

Proof. Clearly, if f, g, I, a, b, F and G are as stated, then by linearity of differentiation

$$(aF + bG)' = aF' + bG' = af + bg \ .$$

□

Thus we see at once that, for example, $\int (2 \sin x + x) = 2 \int \sin x + \int x = -2 \cos x + x^2/2$.

• *Integration by parts.* This may look as a mere technical result on primitives, but the impact of this formula is is actually quite wide. For example, irrationality of certain real numbers can be proven by integration by parts (`https://kam.mff.cuni.cz/~klazar/irratByparts.pdf`, when I write it down).

Theorem 12 (integration by parts) *Suppose that $I \subset \mathbb{R}$ is a nontrivial interval and that $f, g, F, G\colon I \to \mathbb{R}$ are functions such that F is a primitive of f and G is a primitive of g. Then*

$$\int fG = FG - \int Fg \;,$$

meaning that if H is a primitive of Fg then $FG - H$ is a primitive of fG.

Proof. This is an immediate consequence of the Leibniz formula and linearity of differentiation:

$$(FG - H)' = F'G + FG' - H' = fG + Fg - Fg = fG \;.$$

□

One can write the integration by parts formula also as

$$\int F'G = FG - \int FG' \;.$$

If a primitive of FG' is known, the formula gives a primitive for $F'G$. Note how the prime moves from F to G. For example,

$$\begin{aligned} \int \log x \;&=\; \int x' \log x = x \log x - \int x(\log x)' \\ &=\; x \log x - \int \frac{x}{x} = x \log x - x \;. \end{aligned}$$

Or

$$\begin{aligned}\int x\sin x &= \int x(-\cos x)' = -x\cos x + \int x'\cos x \\ &= -x\cos x + \sin x\ .\end{aligned}$$

It is usually easy to check the result by taking the derivative.

• *Integration by substitution.* This is another useful technique for computing primitives. The formula has two forms.

Theorem 13 (integration by substitution) *If $I, J \subset \mathbb{R}$ are nontrivial intervals, $g\colon I \to J$, $f\colon J \to \mathbb{R}$ and g has on I finite g', then the following hold.*

1. *If $F = \int f$ on J then*

$$F(g) = \int f(g)\cdot g' \text{ on } I\ .$$

2. *If g is onto and $g' \neq 0$ on I then one has the implication*

$$G = \int f(g)\cdot g' \text{ on } I \Rightarrow G(g^{-1}) = \int f \text{ on } J.$$

Proof. 1. The formula for derivatives of composite functions gives that $(F(g))' = F'(g)\cdot g' = f(g)\cdot g'$.

2. Since g' is Darboux (Theorem 10), either $g' > 0$ or $g' < 0$ on I. Therefore g either increases or decreases. Thus we have the continuous inverse $g^{-1}\colon J \to I$ because g is continuous on an interval. The formulas for derivatives of composite functions and

derivatives of inverse functions give that

$$\begin{aligned}\big(G(g^{-1})\big)' &= G'(g^{-1})\cdot(g^{-1})' \\ &= f(\cancel{g(g^{-1})})\cdot \cancel{g'(g^{-1})}\cdot \frac{1}{\cancel{g'(g^{-1})}} = f\ .\end{aligned}$$

□

We give two examples.

Example 1. If $F=\int f$ on I and $a,b\in\mathbb{R}$ with $a\neq 0$ then the first formula gives that

$$\frac{F(ax+b)}{a}=\int f(ax+b)\ \text{ on } J:=(I-b)/a\ .$$

Example 2. What is $\int f := \int \sqrt{1-t^2}$ on $J=(-1,1)$? We plug in for t the function $g(x):=\sin x\colon\ I:=(-\pi/2,\pi/2)\to J$. We get by integration by parts that

$$\begin{aligned}\int f(g)\cdot g' &= \int \cos^2 x = \int (\sin x)'\cos x \\ &= \sin x\cdot\cos x - \int \sin x(\cos x)' \\ &= \sin x\cdot \cos x + \int (1-\cos^2 x) \\ &= \sin x\cdot\cos x + x - \int \cos^2 x\end{aligned}$$

and therefore

$$\int f(g)\cdot g' = \int \cos^2 x = \frac{\sin x\cdot\cos x + x}{2} =: G(x)\ .$$

Thus by the second formula and since $\cos x=\sqrt{1-\sin^2 x}$ on I,

$$\int f = \int \sqrt{1-t^2} = G(g^{-1}) = \frac{t\sqrt{1-t^2}+\arcsin t}{2}\ .$$

It is easy to check this formula by differentiation.

THANK YOU FOR YOUR ATTENTION!