Lecture 13: Integration by parts.

Definition. If $a < b$ then define $\int_b^a f\,d\alpha = -\int_a^b f\,d\alpha$ whenever the integral on the right exists. Then

$$\int_a^b f\,d\alpha + \int_b^c f\,d\alpha + \int_c^a f\,d\alpha = 0.$$

We are going to go slow for the next few lectures and concentrate on proving results and understanding the details of the proofs.

Theorem. *If $\int_a^b f\,d\alpha$ exists and $c \in (a,b)$ then $\int_a^c f\,d\alpha$ exists.*

Proof. For every $\varepsilon > 0$, there exists a partition $P(\varepsilon)$ of $[a,b]$ such that

$$P \supseteq P(\varepsilon) \qquad \Rightarrow \qquad \left| \sum_P f(t_j)\Delta\alpha_j - \int_a^b f\,d\alpha \right| < \varepsilon.$$

By replacing $P(\varepsilon)$ with the refinement $P(\varepsilon) \cup \{c\}$, we can assume that $c =\in P(\varepsilon)$. Set

$$\tilde{P}(\varepsilon) = P(\varepsilon) \cap [a,c], \qquad \tilde{\tilde{P}}(\varepsilon) = P(\varepsilon) \cap [c,b].$$

Now suppose that P is a partition of $[a,c]$ with $P \supset \tilde{P}(\varepsilon)$. Then $(P \cup \tilde{\tilde{P}}(\varepsilon)) \supset P(\varepsilon)$ and so

$$\left| \sum_P f(t_k)\Delta\alpha_k + \sum_{\tilde{\tilde{P}}(\varepsilon)} f(x_k)\Delta\alpha_k - \int_a^b f\,d\alpha \right| < \varepsilon.$$

This says that

$$(*) \qquad \left| \sum_P f(t_k)\Delta\alpha_k - I(\varepsilon) \right| < \varepsilon, \qquad I(\varepsilon) = \int_a^b f\,d\alpha - \sum_{\tilde{\tilde{P}}(\varepsilon)} f(x_k)\Delta\alpha_k.$$

We actually need to invoke the **axiom of choice** at this point, to see that for each $\varepsilon > 0$ we can pick some fixed $P(\varepsilon)$. Now we show that $I(\varepsilon) \to 0$ as $\varepsilon \to 0$. Indeed, if $\delta > 0$ and $\varepsilon > 0$ then choose P a partition of $[a,c]$ with $P \supset (\tilde{P}(\varepsilon) \cup \tilde{P}(\delta))$. Then we have

$$\left| \sum_P f(t_k)\Delta\alpha_k - I(\varepsilon) \right| < \varepsilon, \qquad \left| \sum_P f(t_k)\Delta\alpha_k - I(\delta) \right| < \delta.$$

Hence

$$|I(\varepsilon) - I(\delta)| \le \left| I(\varepsilon) - \sum_P f(t_k)\Delta\alpha_k \right| + \left| \sum_P f(t_k)\Delta\alpha_k - I(\delta) \right| < \varepsilon + \delta.$$

We see from this that $I(1/n)$ is a Cauchy sequence as $n \to \infty$ and hence converges to a limit I. But then

$$|I(\varepsilon) - I| \;=\; \lim_{n\to\infty} |I(\varepsilon) - I(1/n)| \;\leq\; \lim_{n\to\infty} (\varepsilon + 1/n) \;=\; \varepsilon.$$

Hence from (*) we get

$$\left| \sum_P f(t_k)\Delta\alpha_k \;\; - \;I \right| \;\leq\; \left| \sum_P f(t_k)\Delta\alpha_k \;\; - \;I(\varepsilon) \right| \;+\; |I(\varepsilon) - I| \;<\; 2\varepsilon.$$

This shows that $\int_a^c f\, d\alpha$ exists and equals I.

Theorem 7.6. *If $f \in R(\alpha)$ on $[a, b]$, then $\alpha \in R(f)$ on $[a, b]$ and*

$$\int_a^b f(x)\, d\alpha(x) \;+\; \int_a^b \alpha(x) df(x) \;=\; f(b)\alpha(b) \;-\; f(a)\alpha(a).$$

Applications of Stieltjes integral: It was introduced to determine when a sequence of numbers $\mu_0, \mu_1, \mu_2, \ldots$ are the moments of a non-decreasing function on $[0, \infty)$, that is when there exists α increasing such that

$$\int_0^\infty x^n d\alpha(x) \;=\; \mu_n, \qquad\qquad n = 0, 1, 2, \ldots$$

Physically, $d\alpha$ represents the mass distribution.

Nowadays the Stieltjes integral is most often used in probability theory. If X is a real valued random variable and $\alpha(x) = P(X \leq x)$ and $f(X)$ is a function, then the expected value of f is

$$E(f(x)) \;=\; \int f(x)\, d\alpha(x).$$

The Stieltjes integral has been used in the theory of partial differential equations for example to prove the Duhamel principal for partial differential equations, although using the Dirac delta function distribution is easier.

Suppose you can find $u(x, t)$ on $\mathbb{R}^n \times [0, \infty)$ solving the homogeneous partial differential equation

$$(*) \qquad \begin{cases} \partial_t u(x,t) \;-\; \Delta u(x,t) \;=\; 0, \quad t > 0 \\ u(x, 0) = g \end{cases}$$

Then you can solve the inhomogeneous equation

$$(**) \qquad \begin{cases} \partial_t u(x,t) \;-\; \Delta u(x,t) \;=\; f(x), \quad t > 0 \\ u(x, 0) = 0 \qquad\qquad\qquad\qquad . \end{cases}$$

The idea using Stieltjes integrals is to take the Heaviside function

$$H(t) \;=\; \begin{cases} 1 & t > 0 \\ 0 & t \le 0. \end{cases}$$

Then if f is continuous with compact support,

$$f(t) \;=\; \int f(t-\tau)\,dH(\tau) \;=\; \int H(\tau)df(t-\tau) \;=\; \int H(t-\tau)df(\tau).$$

We have thus written f as a superposition of translates of the Heaviside function. If V is a vector space and L is a linear operator sending V-valued functions on $\mathbb{R}$ to V-valued functions on $\mathbb{R}$ which is translation invariant and if u_0 solves

$$(Lu_0)(t) \;=\; H(t),$$

then

$$u(t) \;=\; \int u_0(t-\tau)\,df(\tau)$$

solves $Lu = f$.