2.3. Stieltjes Integral

On the contents of the lecture. The Stieltjes relativization of the integral makes the integral flexible. We learn the main transformations of integrals. They allow us to evaluate a lot of integrals.

Basic rules. A *parametric curve* is a mapping of an interval into the plane. In cartesian coordinates a parametric curve can be presented as a pair of functions $x(t), y(t)$. The first function $x(t)$ represents the value of abscises at the moment t, and the second $y(t)$ is the ordinate at the same moment. We define the integral $\int_a^b f(t)\,dg(t)$ for a nonnegative function f, called the *integrand*, and with respect to a nondecreasing *continuous* function g, called the *differand*, as the area below the curve $f(t), g(t) \mid t \in [a,b]$.

A monotone function f is called *continuous* over the interval $[a,b]$ if it takes all intermediate values, that is, the image $f[a,b]$ of $[a,b]$ coincides with $[f(a), f(b)]$. If it is not continuous for some $y \in [f(a), f(b)] \setminus f[a,b]$, there is a point $x(y) \in [a,b]$ with the following property: $f(x) < y$ if $x < x(y)$ and $f(x) > y$ if $x > x(y)$. Let us define a *generalized preimage* $f^{[-1]}(y)$ of a point $y \in [f(a), f(b)]$ either as its usual preimage $f^{-1}(y)$ if it is not empty, or as $x(y)$ in the opposite case.

Now the curvilinear trapezium below the curve $f(t), g(t)$ over $[a,b]$ is defined as $\{(x,y) \mid 0 \leq y \leq g(f^{[-1]}(x))\}$.

The basic rules for relative integrals transform into:

Rule of constant	$\int_a^b f(t)\,dg(t) = c(g(b) - g(a))$, if $f(t) = c$	for $t \in (a,b)$,
Rule of inequality	$\int_a^b f_1(t)\,dg(t) \leq \int_a^b f_2(t)\,dg(t)$, if $f_1(t) \leq f_2(t)$	for $t \in (a,b)$,
Rule of partition	$\int_a^c f(t)\,dg(t) = \int_a^b f(t)\,dg(t) + \int_b^c f(t)\,dg(t)$	for $b \in (a,c)$.

Addition theorem. The proofs of other properties of the integral are based on piecewise constant functions. For any number x, let us define its ε-integral part as $\varepsilon[x/\varepsilon]$. Immediately from the definition one gets:

LEMMA 2.3.1. *For any monotone nonnegative function f on the interval $[a,b]$ and for any $\varepsilon > 0$, the function $[f]_\varepsilon$ is piecewise constant such that $[f(x)]_\varepsilon \leq f(x) \leq [f(x)]_\varepsilon + \varepsilon$ for all x.*

THEOREM 2.3.2 (on multiplication). *For any nonnegative monotone f, and any continuous nondecreasing g and any positive constant c one has*

$$\int_a^b cf(x)\,dg(x) = c\int_a^b f(x)\,dg(x) = \int_a^b f(x)\,dcg(x). \tag{2.3.1}$$

PROOF. For the piecewise constant $f_\varepsilon = [f]_\varepsilon$, the proof is by a direct calculation. Hence

$$\int_a^b cf_\varepsilon(x)\,dg(x) = c\int_a^b f_\varepsilon(x)\,dg(x) = \int_a^b f_\varepsilon(x)\,dcg(x) = I_\varepsilon. \tag{2.3.2}$$

Now let us estimate the differences between integrals from (2.3.1) and their approximations from (2.3.2). For example, for the right-hand side integrals one has:

$$\int_a^b f\,dcg - \int_a^b f_\varepsilon\,dcg = \int_a^b (f - f_\varepsilon)\,dcg \leq \int_a^b \varepsilon\,dcg = \varepsilon(cg(b) - cg(a)). \tag{2.3.3}$$

Hence $\int_a^b f\,dcg = I_\varepsilon + \varepsilon_1$, where $\varepsilon_1 \leq c\varepsilon(g(b) - g(a))$. The same argument proves $c\int_a^b f\,dg = I_\varepsilon + \varepsilon_2$ and $\int_a^b cf\,dg = I_\varepsilon + \varepsilon_3$, where $\varepsilon_2, \varepsilon_3 \leq c\varepsilon(g(b) - g(a))$. Then the pairwise differences between the integrals of (2.3.1) do not exceed $2c\varepsilon(g(b) - g(a))$. Consequently they are less than any positive number, that is, they are zero. $\square$

THEOREM 2.3.3 (Addition Theorem). *Let f_1, f_2 be nonnegative monotone functions and g_1, g_2 be nondecreasing continuous functions over $[a, b]$, then*

$$\int_a^b (f_1(t) + f_2(t))\,dg_1(t) = \int_a^b f_1(t)\,dg_1(t) + \int_a^b f_2(t)\,dg_1(t), \tag{2.3.4}$$

$$\int_a^b f_1(t)\,d(g_1(t) + g_2(t)) = \int_a^b f_1(t)\,dg_1(t) + \int_a^b f_1(t)\,dg_2(t). \tag{2.3.5}$$

PROOF. For piecewise constant integrands both the equalities follow from the Rule of Constant and the Rule of Partition. To prove (2.3.4) replace f_1 and f_2 in both parts by $[f_1]_\varepsilon$ and $[f_2]_\varepsilon$. We get equality and denote by I_ε the common value of both sides of this equality. Then by (2.3.3) both integrals on the right-hand side differ from they approximation at most by $\varepsilon(g_1(b) - g_1(a))$, therefore the right-hand side of (2.3.4) differs from I_ε at most by $2\varepsilon(g_1(b) - g_1(a))$. The same is true for the left-hand side of (2.3.4). This follows immediately from (2.3.3) in case $f = f_1 + f_2$, $f_\varepsilon = [f_1]_\varepsilon + [f_2]_\varepsilon$ and $g = g_1$. Consequently, the difference between left-hand and right-hand sides of (2.3.4) does not exceed $4\varepsilon(g_1(b) - g_1(a))$. As ε can be chosen arbitrarily small this difference has to be zero.

The proof of (2.3.5) is even simpler. Denote by I_ε the common value of both parts of (2.3.5) where f_1 is changed by $[f_1]_\varepsilon$. By (2.3.3) one can estimate the differences between the integrals of (2.3.5) and their approximations as being $\leq \varepsilon(g_1(b) + g_2(b) - g_1(a) - g_2(a))$ for the left-hand side, and as $\leq \varepsilon(g_1(b) - g_1(a))$ and $\leq \varepsilon(g_2(b) - g_2(a))$ for the corresponding integrals of the right-hand side of (2.3.5). So both sides differ from I_ε by at most $\leq \varepsilon(g_1(b) - g_1(a) + g_2(b) - g_2(a))$. Hence the difference vanishes. $\square$

Differential forms. An expression of the type $f_1dg_1 + f_2dg_2 + \cdots + f_ndg_n$ is called a *differential form*. One can add differential forms and multiply them by functions. The integral of a differential form $\int_a^b(f_1\,dg_1 + f_2dg_2 + \cdots + f_ndg_n)$ is defined as the sum of the integrals $\sum_{k=1}^n \int_a^b f_k\,dg_k$. Two differential forms are called equivalent on the interval $[a, b]$ if their integrals are equal for all subintervals of $[a, b]$. For the sake of brevity we denote the differential form $f_1dg_1 + f_2dg_2 + \cdots + f_ndg_n$ by FdG, where $F = \{f_1, \dots, f_n\}$ is a collection of integrands and $G = \{g_1, \dots, g_n\}$ is a collection of differands.

THEOREM 2.3.4 (on multiplication). *Let FdG and $F'dG'$ be two differential forms, with positive increasing integrands and continuous increasing differands, which are equivalent on $[a, b]$. Then their products by any increasing function f on $[a, b]$ are equivalent on $[a, b]$ too.*

PROOF. If f is constant then the statement follows from the multiplication formula. If f is piecewise constant, then divide $[a, b]$ into intervals where it is constant and prove the equality for parts and after collect the results by the Partition Rule. In the general case, $0 \leq \int_a^b fF\,dG - \int_a^b [f]_\varepsilon F\,dG \leq \int_a^b \varepsilon F\,dG = \varepsilon \int_a^b F\,dG$. Since $\int_a^b [f]_\varepsilon F'\,dG' = \int_a^b [f]_\varepsilon F\,dG$, one concludes that $\left|\int_a^b fF'\,dG' - \int_a^b fF\,dG\right| \leq$

$\varepsilon \int_a^b F\, dG + \varepsilon \int_a^b F'\, dG'$. The right-hand side of this inequality can be made arbitrarily small. Hence the left-hand side is 0. □

Integration by parts.

THEOREM 2.3.5. *If f and g are continuous nondecreasing nonnegative functions on $[a, b]$ then $d(fg)$ is equivalent to $fdg + gdf$.*

PROOF. Consider $[c, d] \subset [a, b]$. The integral $\int_c^d f\, dg$ represents the area below the curve $(f(t), g(t))_{t\in[c,d]}$. And the integral $\int_c^d g\, df$ represents the area on the left of the same curve. Its union is equal to $[0, f(d)] \times [0, g(d)] \setminus [0, f(c)] \times [0, g(c)]$. The area of this union is equal to $(f(d)g(d) - f(c)g(c) = \int_c^d dfg$. On the other hand the area of this union is the sum of the areas of curvilinear trapezia representing the integrals $\int_c^d f\, dg$ and $\int_c^d g\, df$. □

Change of variable. Consider a Stieltjes integral $\int_a^b f(\tau)\, dg(\tau)$ and suppose there is a continuous nondecreasing mapping $\tau\colon [t_0, t_1] \to [a, b]$, such that $\tau(t_0) = a$ and $\tau(t_1) = b$. The composition $g(\tau(t))$ is a continuous nondecreasing function and the curve $\{(f(\tau(t), g(\tau(t))) \mid t \in [t_0, t_1]\}$ just coincides with the curve $\{(f(\tau), g(\tau)) \mid \tau \in [a, b]$. Hence, the following equality holds; it is known as the *Change of Variable* formula:

$$\int_{t_0}^{t_1} f(\tau(t))\, dg(\tau(t)) = \int_{\tau(t_0)}^{\tau(t_1)} f(\tau)\, dg(\tau).$$

For differentials this means that the equality $F(x)dG(x) = F'(x)dG'(x)$ conserves if one substitutes instead of an independent variable x a function.

Differential Transformations.

Case dx^n. Integration by parts for $f(t) = g(t) = t$ gives $dt^2 = tdt + tdt$. Hence $tdt = d\frac{t^2}{2}$. If we already know that $dx^n = ndx^{n-1}$, then $dx^{n+1} = d(xx^n) = xdx^n + x^n dx = nxx^{n-1}dx + x^n dx = (n+1)x^n dx$. This proves the Fermat Theorem for natural n.

Case $d\sqrt[n]{x}$. To evaluate $d\sqrt[n]{x}$ substitute $x = y^n$ into the equality $dy^n = ny^{n-1}dy$. One gets $dx = \frac{nx}{\sqrt[n]{x}} d\sqrt[n]{x}$, hence $d\sqrt[n]{x} = \frac{\sqrt[n]{x}}{nx} dx$.

Case $\ln x dx$. We know $d \ln x = \frac{1}{x} dx$. Integration by parts gives $\ln x dx = d(x \ln x) - xd \ln x = d(x \ln x) - dx = d(x \ln x - x)$.

Problems.

1. Evaluate $dx^{2/3}$.
2. Evaluate dx^{-1}.
3. Evaluate $x \ln x\, dx$.
4. Evaluate $d \ln^2 x$.
5. Evaluate $\ln^2 x\, dx$.
6. Evaluate de^x.
7. Investigate the convergence of $\sum_{k=2}^{\infty} \frac{1}{k \ln k}$.