The Fundamental Theorem of Calculus Organized

Abstract: We aim to provide a self-contained *review* of *The Fundamental Theorem of Calculus* We organize its definitions, lemmas, and theorems to form an organic whole, rendering it easier to understand and remember.

0. INTRODUCTION. Most students do not fully understand The Fundamental Theorem of Calculus (FTC). In particular, they are confused about which formulas and theorems belong to the topic and what they are really saying. This paper organizes the different parts of FTC to form an organic whole, rendering it easier to understand and remember. Indeed, writing the paper deepened my knowledge of the topic.

This treatment of the FCT best fits into a calculus class *after* the definitions of derivative and definite integral, along with their motivations and standard calculations, have already been covered.

The FTC speaks about the relationship between a function F and its derivative f. The setting is a real number interval [a,b]; any x considered belongs to [a,b]. Only functions F and f where F has a continuous first derivative and f is continuous on [a,b] are considered. *Indeed, the FTC can be false in other settings.*

Sometimes the assumptions of the setting are repeated in our theorems for completeness and emphasis. Similarly, we may state a theorem with more precise conditions. When we show the existence of an F or f we omit showing the obvious fact that it satisfies the above additional continuity properties.

An *Operator* is a function whose domain is a set of functions. It assigns a function another function. Since the FCT (Theorem 2.6**)** is a global theorem, we take a global view. F and f are objects and the derivative D is an operator which assigns the function F a new function DF (also written F'(x) or dF/dx). For example, with respect to time, F is a height function and DF is the velocity function. .

Stewart [2] is our basic reference for notation, concepts, and techniques. Mention of δ and ε comes with the assumption that $\delta > 0$ and $\varepsilon > 0$.

1. DERIVATIVE AND INDEFINITE INTEGRAL.

Definition 1.1. f is the *derivative* of F written $DF =$ f if

$$\forall x \varepsilon [a,b], \ \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = f(x)$$

The limit definition just says that as we take average rates of change (at x) over smaller and smaller changes Δx, we get closer and closer to some number which we call the *instantaneous rate of change of F* (at x). This justifies saying that the derivative f = DF is the *instantaneous rate of change of F*.

Definition 1.2. $\int f(x)dx$ denotes the i*ndefinite Integral of* f. It stands for any function F with DF = f. All such F differ by a constant so we write $\int f(x)dx$ = F(x) + C; Note that $D(F+C)$ =f.

$\int f(x)dx$ **may not exist.** If f is continuous, the indefinite integral always exists by FTC (2.6A) which provides an explicit formula for it.

C represents an arbitrary constant, the intent being that some value of C makes the above equality *exactly* true. The reason for introducing indefinite integral with its complex notation is that it simplifies writing many calculations and formulas. The below theorem, which provides a skeleton for the FTC (2.6), just restates the definition. Nevertheless, it is extremely useful.

Theorem 1.3.

A. If $\int f(x)dx$ = F(x) + C, then DF (x) = f(x) .

B. If DF (x) = f(x) , then $\int f(x)dx$ = F(x) + C.

Use 1.3A (Checking). (A) says that any integration can be checked. If you calculate $\int xe^x dx =$ xe^x - e^x + C, computing $D(xe^x - e^x) =$ xe^x verifies your answer.

Use 1.3B (Guessing). Finding $\int (x^x \ln x + x^x)dx$ looks hopeless. In desperation you may say " the only possibility is x^x." Differentiating, you find $Dx^x = x^x \ln x + x^x$. Justified by (B), you can answer $\int (x^x \ln x + x^x)dx = x^x + C$. Halpern [1] illustrates this technique and shows how to adjust if the derivative does not yield an exact match.

2. DEFINITE INTEGRAL AND DERIVATIVE.

We begin this section by defining I= $\int_a^b f(x)dx$, I *is the value of the* d*efinite integral* of f *over the interval* [a,b]. The definition justifies considering the *definite integral* as the sum of infinitesimal values over [a,b].

Definition 2.1. The sequence a = $x_0 < x_1 ..., < x_n$ = b is a *partition* P of [a,b]. The *norm* $|P|$ of P is the largest value of $\Delta x_i = x_i - x_{i-1}$ (i=1,…, n). An δ - *partition* is a partition with norm less than δ ($\Delta x_i < \delta$, i=1,…, n).

Definition 2.2. Given a partition a = $x_0 < x_1 ..., < x_n$ = b and a function f, a *Riemann approximating sum* for f is a sum $\sum_{i=1}^{n} f(c_i)\Delta x_i$ with $c_i \in [x_{i-1}, x_i]$.

Definition 2.3. $\int_a^b \mathrm{f}(\mathrm{x})\mathrm{dx}$, the *definie integral* of f over the interval [a,b], exists and has value I if

$$\forall\varepsilon\ \exists\delta \quad \sum_{i=1}^{n} f(c_i)\Delta x_i \text{ is a sum over a } \delta-\textit{partition} \ \rightarrow \left|\sum_{i=1}^{n} f(c_i)\Delta x_i - I\right| < \varepsilon$$

Informally, $\int_a^b \mathrm{f}(\mathrm{x})\mathrm{dx} = \mathrm{I}$ if we can force Riemann approximating sums to get closer and closer to I by restricting the sums to δ - partitions for smaller and smaller δ. The arbitrary choice of c_i makes definition 2.3 very strong. Its full strength is needed to prove the last part of the FCT (2.6B)) . It is remarkable that definite integrals exist for most common functions.

Lemma 2.4 (Stewart [2], Theorem 3, pp 302)**.** *If f is continuous on [a,b], then*

A. $\int_a^b \mathrm{f}(\mathrm{t})\mathrm{dt}$ *exists.*

B. For each x, $\int_a^x \mathrm{f}(\mathrm{t})\mathrm{dt}$ *is a concrete real number*

C. F(x) = $\int_a^x \mathrm{f}(\mathrm{t})\mathrm{dt}$ *defines a function of x.*

Lemma 2.4C provides a definition for functions F(x). Many important functions must be handled using the integral definition because they can't be expressed in terms of the usual elementary functions. An example is the Bell curve function:

Example 2.4. $\mathrm{F}(\mathrm{x}) = \frac{1}{\sqrt{\pi}}\int_{-x}^{x} \mathrm{e}^{-t^2}\mathrm{dt}$ is associated with the *simplest* Bell curve and computes the probability of a random variable falling in the range [-x,x].

The mean value theorem (MVT) below plays a key role in proving several important theorems of the calculus (We use it twice). We state the standard very precise version where behavior at the end points x_{i-1} and x_i is weakened. The usual interpretation is geometric, but the below alternative interpretation gives more insight into our proof of the FTC.

MVT Interpretation 2.5. For any average rate of change $\frac{\mathrm{F}(x_i) - \mathrm{F}(x_{i-1})}{x_i - x_{i-1}}$, we can find a point c_i with equal *instantaneous rate of change*, F'(c_i) = $\frac{\mathrm{F}(x_i) - \mathrm{F}(x_{i-1})}{x_i - x_{i-1}}$.

Theorem 2.5 The Mean Value Theorem. Stewart [2] (pp 216).

If f is continuous on $[x_{i-1}, x_i]$ *and* $DF = f$ *on* (x_{i-1}, x_i).

there exists at least one $c_i \in (x_{i-1}, x_i)$ *with* $\dfrac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}} = F'(c_i) = f(c_i)$.

Corollary 2.5 (Indefinite Integrals differ by a constant). (Stewart [2], Theorem 4.7**,** pp 218).

A. *If* $DH = 0$*, then* $\forall x\ H(x) = C$ *for some constant C.*
B. *If* $DF = DG$ *then* $\forall x\ G(x) = F(x) + C$ *for some constant C.*

Proof. By the MVT,
H(x) - H(x') = $(x - x') \cdot DH\,(c) = 0$ (for some x<c<x'), yielding H(x) = H(x'). ■

(2.5B) follows from defining H = F-G**.** ■

Theorem 2.6 (FTC). *For any f continuous on [a,b],*

A. *If* $F(x) = \int_a^x f(t)\,dt$*, then* $DF(x) = f(x)$

B. *If* $DF(x) = f(x)$*, then* $\int_a^x f(t)\,dt = F(x) - F(a)$.

(2.6A) is the heart of the FCT. It says that $\int_a^x f(t)\,dt + C$ is an indefinite integral, showing that $\int f(x)\,dx$ always exists for continuous f. This fills in an important gap because previously $\int f(x)\,dx$ existed only if associated with an F (DF = f). Later, (2.6A) will provide an alternate proof of (2.6B).

Proof 2.6A. To keep the main ideas clear, we sketch the proof, omitting easy arguments for the intuitive facts (for continuous f) that

(i) Min and Max of f(x) over the closed intervals $[x_0, x_0 + \Delta x]$ exist.

(ii) $\min\limits_{x \in [x_0, x_0+\Delta x]} f(x) \cdot \Delta x \le \int_{x_0}^{x_0+\Delta x} f(t)\,dt \le \max\limits_{x \in [x_0, x_0+\Delta x]} f(x) \cdot \Delta x$

(iii) $\lim\limits_{\Delta x \to 0} \min\limits_{x \in [x_0, x_0+\Delta x]} f(x) = f(x_0) = \lim\limits_{\Delta x \to 0} \max\limits_{x \in [x_0, x_0+\Delta x]} f(x)$

Assume $\Delta x > 0$ (The case $\Delta x < 0$ is similar). Calculate $DF(x_0)$, hoping to get f(x_0).

$$\lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x} = \lim_{\Delta x \to 0} \frac{\int_a^{x_0+\Delta x} f(t)\,dt - \int_a^{x_0} f(t)\,dt}{\Delta x} = \lim_{\Delta x \to 0} \frac{\int_{x_0}^{x_0+\Delta x} f(t)\,dt}{\Delta x}$$

$$\min_{x \in [x_0, x_0+\Delta x]} f(x) \le \frac{\int_{x_0}^{x_0+\Delta x} f(t)\,dt}{\Delta x} \le \max_{x \in [x_0, x_0+\Delta x]} f(x) \qquad \text{(by ii)}$$

$$f(x_0) \le \lim_{\Delta x\to 0} \frac{\int_{x_0}^{x_0+\Delta x} f(t)\,dt}{\Delta x} \le f(x_0) \quad \text{(taking limits and iii).} \quad \blacksquare$$

Discussion 2.6B.

(2.6B) is the main tool for evaluating definite integrals. It provides the final detail needed to complete the FCT picture. Basically, it provides the value of the arbitrary constant C by relating it to the lower integration bound a. (2.6A) says
F(x) + C = $\int_a^x f(t)\,dt$ and (2.6B) provides the detail that C = - F(a).

Despite seeming to fill in a minor detail, (2.6B) makes a powerful statement. It says that *the total change* F(x) – F(a) over [a, x] equals the sum of the infinitesimal changes $f(x)\,dx$ over [a, x]. This idea guides our preferred proof of (2.6B).

At first glance 2.6B seems very peculiar. It relates *instantaneous rate of change* (derivative) to the *sum of infinitesimal changes*, $\int_a^x f(t)\,dt$. Unmasking the hidden identity of f as the instantaneous rate of change of some F (DF = f) makes everything clear. Lemma 2.7 makes motivation 2.7 precise.

Motivation 2.7. f(x) dx = F'(x) dx = $\frac{dF}{dx}$ dx (= dF) is a small change in F.
Adding up the small changes in F gives the total change F(x) – F(a).

Lemma 2.7. *Consider an arbitrary partition* $a = x_0 < x_1 ..., < x_n = b$ *(* $\Delta x_i = x_i - x_{i-1}$ *)* *with F' = DF = f.*

(i) *F(b) - F(a) =*
$$F(b) - F(x_{n-1}) + F(x_{n-1}) - F(x_{n-2}) + F(x_{n-2}) - F(x_{n-3}) \cdots$$
$$+ F(x_3) - F(x_2) + F(x_2) - F(x_1) + F(x_1) - F(a)$$

(ii) *F(b) - F(a) =*
$$\sum_{i=1}^{n} F(x_{i+1}) - F(x_i) = \sum_{i=1}^{n} \frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i} \cdot (x_{i+1} - x_i) = \sum_{i=1}^{n} \frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i} \Delta x_i$$

(iii) $c_i \in [x_i, x_{i+1}]$ exist $(i = 1, \ldots, n)$ so that $F(b) - F(a) = \sum_{i=1}^{n} F'(c_i)\Delta x_i = \sum_{i=1}^{n} f(c_i)\Delta x_i$
which is a Riemann Approximating Su*m* for f.

Proof 2.7.
(i) Summing the changes in each interval $F(x_{i+1}) - F(x_i)$ reduces to F(b) - F(a) by the *telescoping* cancellations. . We wrote x_0 as a, x_n as b, and summed from the last interval to first to emphasize the cancellations.

(ii) We rewrote (i) in Σ-notation and then replaced each $F(x_{i+1}) - F(x_i)$ by its average rate of change view $\frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i} \cdot (x_{i+1} - x_i)$.

(iii) The MVT (2.5) provides $c_i \in [\, x_{i-1}, x_i \,]$ so that each $\frac{\mathrm{F}(x_i) - \mathrm{F}(x_{i-1})}{x_i - x_{i-1}} = f(c_i)$.

Proof 2.6B. It suffices to show that for any ϵ,

(2.8) $$\left| \int_a^x \mathrm{f}(\mathrm{t})\,\mathrm{dt} - (\mathrm{F}(\mathrm{b}) - \mathrm{F}(\mathrm{a})) \right| < \varepsilon .$$

Choose a δ so that for all δ-partitions

(2.9) $$\left| \int_a^x \mathrm{f}(\mathrm{t})\,\mathrm{dt} - \sum_{i=1}^{n} f(c_i)\Delta x_i \right| < \varepsilon .$$

Choose any δ-partition P satisfying (2.9). (2.7 i) yields an expression for F(b) - F(a) which by (2.7 iii) yields a Riemann approximating sum $\sum_{i=1}^{n} f(c_i)\Delta x_i =$ F(b) - F(a). Now replacing $\sum_{i=1}^{n} f(c_i)\Delta x_i$ in (2.9) by F(b) - F(a) yields (2.8) . ■

Alternate Proof 2.6B.

For F(x) = $\int_a^x \mathrm{f}(\mathrm{t})\,\mathrm{dt}$, $\int_a^x \mathrm{f}(\mathrm{t})\,\mathrm{dt}$ = F(x) – F(a) since F(a)= $\int_a^a \mathrm{f}(\mathrm{t})\,\mathrm{dt} = 0$.

By corollary 2.5 (Indefinite Integrals differ by a constant),

If DG (x) = f(x) = DF (x), then G(x) = F(x) + C for some constant C. Hence,

G(x) - G(a) = (F(x) + C) - (F(a) + C) = F(x) – F(a), yielding $\int_a^x \mathrm{f}(\mathrm{t})\,\mathrm{dt}$ = G(x) – G(a).■

Theorem 2.8 (The summary version of FTC).

Differentiation and integration are inverse operation, i.e.,

A. If you start with a function f and integrate it and then differentiate the result, you are back to the original function f. $D\int_a^x \mathrm{f}(\mathrm{t})\,\mathrm{dt}$ *(x) = f(x).*

B. If you start with a function F and differentiate it and then integrate the result, you are back to the original function F. $\int_a^x DF(\mathrm{t})\,\mathrm{dt}$ *= F(x) – F(a).*

REFERENCES

---

1. F. Halpern, Integration by guessing, (submitted).

2. J. Stewart, Calculus, 6th ed., Thomson Brooks/Cole, Belmont, 2008