MVT implies FTC

I want to deduce a form of the Fundamental Theorem of Calculus (FTC) from the Mean Value Theorem (MVT) for several reasons.

- To *use* the MVT, and not exhibit it as an isolated curiosity. The latter is its fate in many calculus courses. The MVT is one of the central observations of calculus, and it and its immediate consequences are important in many applications.
- I want to review the FTC, a wonderful and unexpected connection between limits of sums and limits of slopes. Students and others who use calculus can easily forget the power of the FTC. Confusion of definitions and results occurs (is the integral the difference in values of two antiderivatives or is it some complicated idea involving dissections of areas into slim rectangles?), often encouraged by a choice of nomenclature (**in**definite integral versus definite integral) which in my opinion is rather confusing..
- The discussion below gives a numerical estimate of the difference between certain Riemann sums and their limits. One elementary numerical technique for approximating definite integrals is verified, and (in principal!) given enough time and computer help, many definite integrals could be approximated with a given accuracy. Many functions are introduced as values of definite integrals and the computational implications of such definitions should be appreciated.

Here's a statement of the MVT:

Mean Value Theorem. *Suppose F is a function defined and differentiable on the interval $[A, B]$. Then there is at least one number C in the interval so that:*

$$\frac{F(B) - F(A)}{B - A} = F'(C)$$

In most texts this is usually followed by a series of obvious examples intended to reassure the reader. These examples don't show how from the way the MVT is used in applications of mathematics. Proving the MVT is not difficult, and is usually accompanied by a nice picture. See section 4.2 of the fifth edition of Stewart's *Calculus (Early Transcendentals)*.

Now suppose that f is a differentiable function defined on the unit interval, $[0, 1]$. First chop up the interval into n equal pieces, where n is supposed to be a very large integer. The j^{th} subinterval has endpoints $\left[\frac{j-1}{n}, \frac{j}{n}\right]$. You can check this assertion, as I always do, by looking at the extreme values of j: the first subinterval, where $j = 1$, has endpoints $\left[0, \frac{1}{n}\right]$ because $j - 1 = 0$, and the last subinterval, where $j = n$, has endpoints $\left[\frac{n-1}{n}, 1\right]$ because $\frac{n}{n} = 1$.

Apply the MVT in each subinterval. So in the j^{th} such subinterval, we know that there is at least one number c_j so that

$$\frac{f(\frac{j}{n}) - f(\frac{j-1}{n})}{\frac{1}{n}} = f'(c_j)$$

The $B - A$ in the bottom of the right-hand side has become $\frac{1}{n}$, the length of the subinterval.

We don't know very much about the c_j in the above equation. Its existence is guaranteed by the MVT, but all we know is that it is somewhere in the j^{th} subinterval. What if we wanted to "change" c_j? What if we wanted to write $f'(d_j)$ in the equation, where d_j is some number in the same subinterval, but a number which *we* specify or pick or choose or whatever. The equation need not be true then. If it were still true, we'd either have been very lucky or we'd have a constant function for f', which wouldn't be terribly interesting! So what sort of *error* could we commit if we changed $f'(c_j)$ to $f'(d_j)$? Here is a central theme of a large chunk of mathematics: since what we'd like to do is not true precisely, can we estimate much they *fail* to be true? We want to "control"

$$f'(c_j) - f'(d_j)$$

Now some inspiration is necessary. We can control or estimate this by applying the MVT to the function f' on the interval $[c_j, d_j]$. Since f'' is the derivative of f' we know:

$$\frac{f'(c_j) - f'(d_j)}{c_j - d_j} = f''(e_j)$$

But multiply and add to get rid of both the division sign and the subtraction sign (always good ideas!). The equation becomes:

$$f'(c_j) = f'(d_j) + f''(e_j) \cdot (c_j - d_j)$$

How can we estimate the second term on the right-hand side of this equation? I'll be interested in the magnitude of the error so I'll estimate the sizes (with no signs – the absolute values) of everything.

We know that $|c_j - d_j|$ must be less than $\frac{1}{n}$, because they are both in the same subinterval of length $\frac{1}{n}$. Also suppose I know a number M_2 which is an overestimate of the function $|f''(x)|$ can be on the whole interval $[0,1]$. I get M_2 is any way. In practice this is not a problem and I only need an overestimate, not an exact value. You'll see why.

So I know that

$$f'(c_j) = f'(d_j) + \mathbf{FUZZ}$$

where

$$|\,\mathbf{FUZZ}\,| \le M_2 \cdot \left(\frac{1}{n}\right)$$

Take this and replace the $f'(c_j)$ in the equation on the top of this page:

$$\frac{f\left(\frac{j}{n}\right) - f\left(\frac{j-1}{n}\right)}{\frac{1}{n}} = f'(d_j) + \mathbf{FUZZ}$$

Get rid of the quotient by multiplying. The equation becomes:

$$f\left(\frac{j}{n}\right) - f\left(\frac{j-1}{n}\right) = f'(d_j)\left(\frac{1}{n}\right) + \mathbf{NEW\ FUZZ}$$

where

$$|\,\textbf{NEW FUZZ}\,| \le M_2 \cdot \left(\frac{1}{n^2}\right)$$

since $\left(\frac{1}{n}\right) \cdot \left(\frac{1}{n}\right) = \frac{1}{n^2}$.

There are n equations like that one, one equation for each subinterval. Line them up:

$$\begin{aligned}
f\left(\frac{1}{n}\right) - \quad f(0) \quad &= f'(d_j)\left(\frac{1}{n}\right) + \textbf{NEW FUZZ} \\
&\vdots \\
f\left(\frac{j}{n}\right) - f\left(\frac{j-1}{n}\right) &= f'(d_j)\left(\frac{1}{n}\right) + \textbf{NEW FUZZ} \\
&\vdots \\
f(1) \; - f\left(\frac{n-1}{n}\right) &= f'(d_j)\left(\frac{1}{n}\right) + \textbf{NEW FUZZ}
\end{aligned}$$

I am *not* asserting that the various **NEW FUZZ** terms are the same. Now let's add these equations. The left-hand sides will collapse or "telescope" because of the sign patterns. On the right-hand sides there will be a sum of n **NEW FUZZ**'s. And if we assume the worst (which we should if we're trying to do an honest error analysis), the errors might all reinforce. But how big can the resulting error be? Each piece of the error is bounded by $M_2 \cdot \left(\frac{1}{n^2}\right)$, and n of these (now the powers cancel) give us an error bound of $n \cdot M_2 \cdot \left(\frac{1}{n^2}\right) = M_2 \cdot \left(\frac{1}{n}\right)$. The other "stuff" on the right-hand sides can be written using Σ notation, the mathematical abbreviation for sums. Appendix E of the text discusses such notation and Section 5.1 applies this notation to problems similar to those we consider here.

$$f(1) - f(0) = \sum_{j=1}^{n} f'(d_j)\left(\frac{1}{n}\right) + \textbf{FINAL FUZZ}$$

One more notational change: I'll write g in place of f'. Then we get:

$$\boxed{f(1) - f(0) = \sum_{j=1}^{n} g(d_j)\left(\frac{1}{n}\right) + \textbf{FINAL FUZZ}}$$

where $|\,\textbf{FINAL FUZZ}\,| \le \frac{\text{constant}}{n}$. I'm not particularly interested in the "constant" right now but we do know where it came from.

What I like very much about this equation is that the d_j's are not supplied to us by the MVT, but *we* get to choose or specify any numbers we like subject to the condition that they're inside the appropriate j^{th} subinterval of $[0, 1]$

The term $\sum_{j=1}^{n} g(d_j)\left(\frac{1}{n}\right)$ is called a **Riemann sum** for g on $[0, 1]$. Such sums arise in computing areas and volumes. They also occur in the analysis of a wide variety of "processes" where samples are taken over a chopped-up duration interval, and the samples are multiplied by the length of the duration subinterval. Think of water flowing down a small stream from time $t = 0$ to time $t = 1$. Chop up this interval into small subintervals of duration $\triangle t_j$. In each little "chunk" of time, make one measurement, $g(d_j)$, of water flow (in gallons per second?). Then $\sum_{j=1}^{n} g(d_j)\triangle t_j$ approximates the total water flow.

What does the boxed equation above declare? Please see the following consequences.

- ⋆ The Riemann sums for the function g, *no matter what the sample points*, must tend to a unique limit as $n \to \infty$. That is, $\lim_{n\to\infty} \sum_{j=1}^{n} g(d_j)\left(\frac{1}{n}\right)$ exists and is always the same, for any collection of sample points.
- ⋆ The "rate" at which the sums tend to this limit is measured or bounded by a computable error term for the functions considered here. The bound for the error is at most of order n^{-1} (that is, the error is $\leq \frac{\text{constant}}{n}$). Other algorithms with faster rates of convergence might be preferred in practice. The constant here is related to the first derivative of g, which is how we renamed f'.
- ⋆ If we know a function f whose derivative is g then the limit *is* $f(1) - f(0)$. This is a wonderful way to compute the limit, and we will investigate some intricate strategies for discovering such functions (antiderivatives). Searching for such antiderivatives may be tedious or difficult, and it may even be pointless because we can always compute accurate approximations.

These observations are most of the **Fundamental Theorem of Calculus**. You can see that the FTC has many facets. Many students believe that the FTC is exactly the equation

$$\frac{d}{dx}\int_a^x g(t)dt = g(x)$$

which is related to what we've discussed: the limit of the Riemann sums for g creates a function which is an antiderivative for g.