PHILOSOPHIÆ NATURALIS **PRINCIPIA** MATHEMATICA

Al Caveman

December 18, 2015

Contents

Chapter 1

Preliminaries Because You Are an Idiot

NOTE: THIS CHAPTER IS PROBABLY WRONG I WILL FIX ONE DAY

There seems to be a book written on fundamentals of geometry by some guy named Hiblert something, that might have a different view. Who cares.

1.1 What is a Ruler

A stick with things on it that tell you how long other things are. This is called *The Axiom of Ruler.*

1.2 What is a Dot?

In a d-dimensional space, such as $\mathbb{R}^d$, a dot is a thing with these properties:

- Its position in the space $\mathbb{R}^d$ is defined by the vector $(x_i, x_2, \ldots, x_d)$ where for any $1 \leq i \leq d$, $x_i \in \mathbb{R}$.
- If you look at the dot from the perspective of any specific dimension (e.g. $i = 1, 2, \ldots, d$), and put a ruler on the dot, you will see that its width w_i is $0 < w_i < g$, where g is any positive real number. Do you understand? It means the tiniest possible real number that is closest to 0 but not 0. Now go fuck yourself. In other words, a dot in $\mathbb{R}^d$ can be thought as a cube with d faces.

Some say that it is impossible to find w_i such that for any positive g, $0 < w_i < g$. But I don't buy that. I think that there is, except that we cannot find it. There is a difference between something to not exist, and us not being able to find something. So I think the number exists, but we can't fine it. All we can do is define it as w_i.

Why should w_i be positive instead of just 0? Because we use dots to define lines, and volumes. If w_i for any dimension i is zero, then all lines/volumes are also 0. Dead-end. We have to make it positive, but smallest positive to permit maximum granularity.

1.3 What is a Line?

A line is formed by putting a lot of dots next to each other, such that the gap between those lines is 0.

Suppose that we demand to create a line such that if we measure its length by a ruler (using axiom of ruler), its width is x. How many dots are there? Let w be dots width, and their quantity be y such that:

$$x = wy$$

So there are y many of them by definition. So $y = \frac{x}{w}$. We also know that by definition $w = 0$.

Chapter 2

Integration

The goal here is to integrate the area under $f(x)$ from when $x = 0$ until we reach $x = x_{end}$. What follows here is bunch of steps that shows my thinking process because *shadowdaemon* asked for it (and also because I am happy now).

2.1 Areas of Lots of Extremely Tiny Rectangular Columns

So to integrate $f(x) = x^2$ we have to keep summing extremely skinny columns[1] of, each of width d.

$$\begin{aligned}(x + d - x)(x)^2 +\\ (x + 2d - (x + d))(x + d)^2 +\\ (x + 3d - (x + 2d))(x + 2d)^2 +\\ (x + 4d - (x + 3d))(x + 3d)^2 +\\ (x + 5d - (x + 4d))(x + 4d)^2 +\\ (x + 6d - (x + 5d))(x + 5d)^2 +\\ (x + 7d - (x + 6d))(x + 6d)^2 +\\ (x + 8d - (x + 7d))(x + 7d)^2 +\\ (x + 9d - (x + 8d))(x + 8d)^2 +\\ (x + 10d - (x + 9d))(x + 9d)^2 +\\ (x + 11d - (x + 10d))(x + 10d)^2 +\\ \ldots\end{aligned}$$

You see, if d is extremely tiny (near zero), then we will have to sum an infinite number of those tiny skinny areas. But for simplicity I put ... instead.

[1]**Note:** *fefelix* of *freenode/#gentoo-chat-exile* tried to look smart by attacking my rigor by saying that the term *skinny columns* is wrong and that it must be replaced by the term *infinitesimal* (facepalm moment here). He also tried to look even smarter by using the phrase *In the realm of* $\mathbb{R}$. Obviously, any half-assed mathematician knows that $\mathbb{R}$ has only a single infinitesimal which is zero. Yep, zero bitch. So the term *infinitesimal* is absolutely wrong in this context, thus even worse than *skinny columns*. The only exception is if *fefelix* wishes to live in the 1600s.

We can simplify things:

$$
\begin{aligned}
&d(x)^2+\\
&d(x+d)^2+\\
&d(x+2d)^2+\\
&d(x+3d)^2+\\
&d(x+4d)^2+\\
&d(x+5d)^2+\\
&d(x+6d)^2+\\
&d(x+7d)^2+\\
&d(x+8d)^2+\\
&d(x+9d)^2+\\
&d(x+10d)^2+\\
&\ldots
\end{aligned}
$$

Let's expand those squares:

$$
\begin{aligned}
&d(x)(x)+\\
&d(x+d)(x+d)+\\
&d(x+2d)(x+2d)+\\
&d(x+3d)(x+3d)+\\
&d(x+4d)(x+4d)+\\
&d(x+5d)(x+5d)+\\
&d(x+6d)(x+6d)+\\
&d(x+7d)(x+7d)+\\
&d(x+8d)(x+8d)+\\
&d(x+9d)(x+9d)+\\
&d(x+10d)(x+10d)+\\
&\ldots
\end{aligned}
$$

Let's multiply them square bitches:

$$
\begin{aligned}
dx^2&+\\
d(x^2 + 2dx + d^2)&+\\
d(x^2 + 4dx + 4d^2)&+\\
d(x^2 + 6dx + 9d^2)&+\\
d(x^2 + 8dx + 16d^2)&+\\
d(x^2 + 10dx + 25x^2)&+\\
d(x^2 + 12dx + 36d^2)&+\\
d(x^2 + 14dx + 49d^2)&+\\
d(x^2 + 16dx + 64d^2)&+\\
d(x^2 + 18dx + 81d^2)&+\\
d(x^2 + 20dx + 100d^2)&+\\
&\dots
\end{aligned}
$$

Let's now multiply those bitches with d so that the shit gets spread even more:

$$
\begin{aligned}
dx^2&+\\
dx^2 + 2d^2x + d^3&+\\
dx^2 + 4d^2x + 4d^3&+\\
dx^2 + 6d^2x + 9d^3&+\\
dx^2 + 8d^2x + 16d^3&+\\
dx^2 + 10d^2x + 25x^3&+\\
dx^2 + 12d^2x + 36d^3&+\\
dx^2 + 14d^2x + 49d^3&+\\
dx^2 + 16d^2x + 64d^3&+\\
dx^2 + 18d^2x + 81d^3&+\\
dx^2 + 20d^2x + 100d^3&+\\
&\dots
\end{aligned}
$$

2.2 Approximating the Area

Now this is a critical point. Below is basically saying that each row is an approximation for the area under $f(x)$ from $x = 0$ till $x = x_{end}$. So the 1^{st} one is a shit approximation where were approximate the area under that curve by only one big fat column; so d is so huge here, in fact $d = x_{end}$.

Then, the 2nd line show a slightly less shit approximation where we approximate the area under the curve by two fat ass rectangular columns. So here $d = \frac{x_{end}}{2}$.

So the approximation of the area under the curve gets more and more accurate in each line.

$$\begin{aligned}
dx^2 \\
2dx^2 + 2d^2x + d^3 \\
3dx^2 + 6d^2x + 5d^3 \\
4dx^2 + 12d^2x + 14d^3 \\
5dx^2 + 20d^2x + 30d^3 \\
6dx^2 + 30d^2x + 55d^3 \\
7dx^2 + 42d^2x + 91d^3 \\
8dx^2 + 56d^2x + 140d^3 \\
9dx^2 + 72d^2x + 204d^3 \\
10dx^2 + 90d^2x + 285d^3 \\
11dx^2 + 110d^2x + 385d^3
\end{aligned}$$

So basically, you see there is a pattern. The coefficient of the 1^{st} term is easy peasy (just incrementing from 1 to ∞). The coefficient from the 2^{nd} term is kinda interesting, it follows the equation $(i^2 + i)$ where i is the line number. Note that we start counting lines from 0. So the 1^{st} has $i = 0$ and the 2^{nd} line has $i = 1$, etc. Finally, the last term is kinda cool, it follows the pattern $\frac{(i^2+i)(2i+1)}{6}$.

Now you may ask, how did I find these patterns? Well these are well known number series. You can look them up in the On-Line Encyclopedia of Integer Sequences[2].

So, the area under the curve of $f(x)$ from $x = 0$ up to x_{end}, by any d (and its corresponding i) is:

$$dx^2 + (i^2 + i)d^2x + \frac{(i^2 + i)(2i + 1)}{6}d^3$$

Now we are almost done. We know that $x = 0$, so we can cancel a few terms:

$$\begin{aligned}
d0^2 + (i^2 + i)d^2 0 + \frac{(i^2 + i)(2i + 1)}{6}d^3 \\
\frac{(i^2 + i)(2i + 1)}{6}d^3
\end{aligned} \tag{2.1}$$

Of course, we could've canceled those terms that multiply against zero earlier, but I didn't for random reasons. I just didn't. That's the randomness of life. But it's all mathematically correct as my caveman balls tell.

You can code a simple script that you give it x_{end} and d, by which it automatically finds $i = x_{end}/d$. You will notice that as d gets smaller, you end up approaching some limit after which reduction in d does not cause any change in the estimated area under the curve.

[2] http://oeis.org/

2.3 The Precise Area Under The Bitch

Now let's find the ultimate precision in the limit as $i \to \infty$ which also means that $d \to 0$. But how about not? Cause it's too hard to solve the limit when two variables are approaching different limits.

To simplify the limits in an easier way, let's represent i in terms of d and x_{end} as follows $i = x_{end}/d$. Then the same equation would become as follows:

$$\frac{((x_{end}/d)^2 + (x_{end}/d))(2(x_{end}/d) + 1)}{6} d^3$$
$$\frac{(\frac{x_{end}^2}{d^2} + \frac{x_{end}}{d})(\frac{2x_{end}}{d} + 1)}{6} d^3$$
$$\frac{(\frac{x_{end}^2}{d^2} + \frac{x_{end}}{d})(\frac{2x_{end}}{d} + 1)}{6} d^3$$
$$\frac{\frac{1}{d}(\frac{x_{end}^2}{d} + x_{end})(\frac{2x_{end}}{d} + 1)}{6} d^3$$
$$\frac{(\frac{x_{end}^2}{d} + x_{end})(\frac{2x_{end}}{d} + 1)}{6} d^2$$
$$\frac{\frac{2x_{end}^3}{d^2} + \frac{x_{end}^2}{d} + \frac{2x_{end}^2}{d} + x_{end}}{6} d^2$$
$$\frac{\frac{2x_{end}^3 d^2}{d^2} + \frac{x_{end}^2 d^2}{d} + \frac{2x_{end}^2 d^2}{d} + x_{end} d^2}{6}$$
$$\frac{2x_{end}^3 + x_{end}^2 d + 2x_{end}^2 d + x_{end} d^2}{6}$$

Now, it's super easy. We have to find the limit of that equation as a single variable approaches 0 (we got rid of i). The equation becomes:

$$\frac{2x_{end}^3}{6}$$
$$\frac{x_{end}^3}{3}$$

That's it. Integration re-invented bitch :) — $\frac{x_{end}^2}{3}$.
Q.E. freaking DEE.

2.4 Generalizing That

Let's integrate x^c where $x, c \in \mathbb{R}$. How easy is that? Let's try.

So, back to the business of summing skinny columns:

$$\begin{aligned} d \times d^c &+ \\ d \times (2d)^c &+ \\ d \times (3d)^c &+ \\ d \times (4d)^c &+ \\ d \times (5d)^c &+ \\ \ldots \end{aligned}$$

Simplified to:

$$\begin{aligned} d \times d^c &+ \\ d \times 2^c \times d^c &+ \\ d \times 3^c \times d^c &+ \\ d \times 4^c \times d^c &+ \\ d \times 5^c \times d^c &+ \\ \ldots \end{aligned}$$

As we try to estimate that sum better we get:

$$nd \times (\frac{n(n+1)}{2})^c \times d^c$$

Chapter 3

Differentiation

Here we want to find the slope of $f(x)$ at point x. This is easy so I won't say much here.

Differentiate x^2.
$\frac{f(x+d)-f(x)}{x+d-x}$
$\frac{(x+d)^2-x^2}{x+d-x}$
$\frac{(x+d)^2-x^2}{d}$
$\frac{(x+d)(x+d)-x^2}{d}$
$\frac{x^2+xd+xd+d^2-x^2}{d}$
$\frac{x^2}{d} + \frac{2xd}{d} + \frac{d^2}{d} - \frac{x^2}{d}$
$\frac{x^2}{d} + 2x + d - \frac{x^2}{d}$
$2x + d$
Now, as $d \to 0$, it becomes $2x$. Done.

3.1 Theorems

My symbol for differentiation is diff. That's it. Now, if t, x, c are any numbers, except c is just a non-zero natural number, then:

Theorem 1. $\operatorname{diff} tx^c = ctx^{c-1}$.

Okay let's do this... let's assume that $f(x) = tx^c$ for a wise reason that I need not explain. Just trust me that I am doing the right thing.

Proof.

$$\begin{aligned}
\operatorname{diff} f(x) &= \frac{f(x+d)-f(x)}{(x+d)-x} \\
&= \frac{f(x+d)-f(x)}{d} \\
&= \frac{t(x+d)^c - tx^c}{d} \\
&= \frac{t(\sum_{n=0}^{c}\binom{c}{c-n}x^{c-n}d^n) - tx^c}{d} \\
&= \frac{(\sum_{n=0}^{c} t\binom{c}{c-n}x^{c-n}d^n) - tx^c}{d} \\
&= \frac{t\binom{c}{c-0}x^{c-0}d^0 + (\sum_{n=1}^{c} t\binom{c}{c-n}x^{c-n}d^n) - tx^c}{d} \\
&= \frac{tx^c + (\sum_{n=1}^{c} t\binom{c}{c-n}x^{c-n}d^n) - tx^c}{d} \\
&= \frac{\sum_{n=1}^{c} t\binom{c}{c-n}x^{c-n}d^n}{d} \\
&= \sum_{n=1}^{c} \frac{t\binom{c}{c-n}x^{c-n}d^n}{d} \\
&= \sum_{n=1}^{c} t\binom{c}{c-n}x^{c-n}d^{n-1} \\
&= t\binom{c}{c-1}x^{c-1}d^{1-1} + \sum_{n=2}^{c} t\binom{c}{c-n}x^{c-n}d^{n-1} \\
&= tcx^{c-1} + \sum_{n=2}^{c} t\binom{c}{c-n}x^{c-n}d^{n-1}
\end{aligned}$$

Then:

$$\lim_{d\to 0} tcx^{c-1} + \sum_{n=2}^{c} t\binom{c}{c-n}x^{c-n}d^{n-1} = tcx^{c-1}$$

OMG it's the Q.E.D. baby! □

Chapter 4

Multivariate Stuff

4.1 Functions that are sum of other linear functions

Let's say that we got linear functions $b(n)$ and $c(m)$. By definition, this means that:

$$b(n) = c_1 + w_1 n \tag{4.1}$$

$$c(m) = c_2 + w_2 m \tag{4.2}$$

Theorem 2. *If* $a(n,m) = b(n)+c(m)$*, then* $\lim_{(n,m)\to(x,y)} a(n,m) = \lim_{n\to x} b(n) + \lim_{m\to y} c(m)$

Proof. Limits of b and c:

$$\begin{aligned}\lim_{n\to x} b(n) =& \lim_{n\to x} c_1 + w_1 n \\ =& c_1 + w_1 x\end{aligned} \tag{4.3}$$

$$\begin{aligned}\lim_{m\to y} c(m) =& \lim_{n\to y} c_2 + w_2 m \\ =& c_2 + w_2 y\end{aligned} \tag{4.4}$$

Limit of a by theorem:

$$\lim_{(n,m)\to(x,y)} a(n,m) = c_1 + w_1 x + c_2 + w_2 y \tag{4.5}$$

By substitution we get expanded a:

$$a(n,m) = c_1 + w_1 n + c_2 + w_2 m \tag{4.6}$$

Limit of expanded a:

$$\begin{aligned}\lim_{(n,m)\to(x,y)} a(n,m) =& \lim_{(n,m)\to(x,y)} c_1 + w_1 n + c_2 + w_2 m \\ =& c_1 + w_1 x + c_2 + w_2 y\end{aligned} \tag{4.7}$$

Guess wat, limit of expanded a is the same as the original one. Who could've thought? □

4.2 Less Gay Stuff

Let's say that we wish to differentiate $f(x, y) = x^2/y$. What does it mean? The slope exactly at point (x, y) that touches f? And by the way, my notation for that differentiation is: $\text{diff}^{f}_{(x,y)}\, f(x, y)$ which means differentiate f with respect to (x, y).

Obviously, there are infinitely many slopes cause we are dealing with surfaces (not curves). So in my logic, you