Integration: Step by Step[1]

The aim of this note is to summarize the basic theory of Riemann integration on $\mathbf{R}^n$ from the point of view of step functions. In addition to definitions and statements of main results, I give a few (though not nearly all) illustrative proofs along the way.

0.1. **Rectangles, partitions and step functions.**

Definition 0.1. *A* rectangle *is a compact set* $R \subset \mathbf{R}^n$ *of the form* $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ *where each closed interval* $[a_j, b_j] \subset \mathbf{R}$ *has finite but non-zero length. That is,*

$$R = \{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_j \in I_j\}.$$

The volume *of* R *is the product* $\mathrm{Vol}\, R := (b_1 - a_1) \cdot \cdots \cdot (b_n - a_n)$ *of the lengths of the intervals* $[a_j, b_j]$.

Note that as is typical in this subject, we use the word 'rectangle' for $R \subset \mathbf{R}^n$ regardless of whether $n = 1$ (in which case R is simply a closed interval), $n = 2$ (in which case R is a rectangle in the usual sense) or $n = 39$ (dimension-specific naming conventions for rectangles stop with $n = 28$...[2]). In what follows, we will often want to decompose a given rectangle R into a grid of smaller rectangles.

One can always subdivide a given rectangle R into a collection of smaller rectangles Q. We will be interested in doing this so that the smaller rectangles are arranged in a 'grid' pattern. Specifically, when $R = [a, b] \subset \mathbf{R}$ is an interval, a partition of R is a collection $\mathcal{I}$ of subintervals $[a_{j-1}, a_j]$ corresponding to an increasing sequence $a = a_0 < a_1 < \cdots < a_{n-1} < a_n = b$ such that $\mathcal{I} = \{[a_{j-1}, a_j] : 1 \leq j \leq n\}$. Alternatively (but equivalently), $\mathcal{I}$ is a collection of subintervals $I \subset [a, b]$ such that

- $\bigcup_{I \in \mathcal{I}} = [a, b]$; and
- different intervals $I_1, I_2 \in \mathcal{I}$ satisfy $\mathrm{int} I_1 \cap \mathrm{int} I_2 = \emptyset$.

Now if $R = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbf{R}^n$ is a higher dimensional rectangle, then a partition of R is again a collection $\mathcal{P}$ of non-overlapping subrectangles $Q \subset R$ that cover R. But we further require that these rectangles 'form a grid', i.e. there are partitions $\mathcal{I}_j$ of $[a_j, b_j]$ such that $Q \in \mathcal{P}$ if and only if $Q = I_1 \times \cdots \times I_n$ for intervals $I_1 \in \mathcal{I}_1, \ldots I_n \in \mathcal{I}_n$.

The point behind partitions is that we can make them 'smaller and smaller.' More precisely,

Definition 0.2. *Let* $\mathcal{P}, \mathcal{P}'$ *be partitions of the same rectangle* R*. We say that* $\mathcal{P}'$ refines $\mathcal{P}$ *if for any rectangle* $Q' \in \mathcal{P}'$*, there exists a rectangle* $Q \in \mathcal{P}$ *such that* $Q' \subset Q$.

The next two propositions are basic facts about partitions that we will use a lot. Both are 'obvious' from an intuitive standpoint. Nevertheless, it is somewhat tricky to give an uncluttered formal justification.

Proposition 0.3. *Let* $\mathcal{P}$ *be a partition of a rectangle* $R \subset \mathbf{R}^n$*. Then*

$$\sum_{Q \in \mathcal{P}} \mathrm{Vol}\, Q = \mathrm{Vol}\, R.$$

If, moreover, $\mathcal{P}'$ *is another partition of* R*, then there is a third partition* $\mathcal{P}''$ *of* R *that refines both* $\mathcal{P}$ *and* $\mathcal{P}'$.

[1] Pun intended.

[2] You understand I'm kidding, right?

To any subset $S \subset \mathbf{R}^n$, one can associate an *indicator function* $\mathbf{1}_S : \mathbf{R}^n \to \mathbf{R}$ given by $\mathbf{1}_S(x) = 1$ for $x \in S$ and $\mathbf{1}_S(x) = 0$ otherwise. More generally, if $R \subset \mathbf{R}^n$ is a rectangle with a partition $\mathcal{P}$ we will call a function $f : R \to \mathbf{R}$ a *step function subordinate to* $\mathcal{P}$ if f is constant on each rectangle $Q \in \mathcal{P}$. That is, there are constants $f_Q \in \mathbf{R}$ such that f is a linear combination

$$f = \sum_{Q \in \mathbf{P}} f_Q \mathbf{1}_Q$$

of indicator functions. Note that, strictly speaking, when we say 'piecewise constant' we are disregarding overlapping edges between different rectangles. This little transgression makes no difference, so we will not repent of it in what follows. On 'base times height' grounds, it is reasonable that for rectangles $Q \subset \mathbf{R}^n$ we should have $\int_R \mathbf{1}_Q \, dV = \operatorname{Vol} Q$. So we will just define things that way. More generally,

Definition 0.4. *Let $f : R \to \mathbf{R}$ be a step function on a rectangle $R \subset \mathbf{R}^n$ subordinate to a partition $\mathcal{P}$ of R. Then the* integral of f *over R is the quantity*

$$\int_R f \, dV = \sum_{Q \in \mathcal{P}} f_Q \operatorname{Vol} Q.$$

An important thing to note is that if $f : R \to \mathbf{R}$ is a step function subordinate to $\mathcal{P}$ and $\mathcal{P}'$ refines $\mathcal{P}$, then f is also subordinate to $\mathcal{P}'$. This opens up the possibility of evaluating the integral of f using either partition. It is therefore important to know that our definition of integral for a step function is independent of the underlying partition.

Proposition 0.5. *Let $f : R \to \mathbf{R}$ be a step function on a rectangle $R \subset \mathbf{R}^n$. If f is subordinate to two partitions $\mathcal{P}$ and $\mathcal{P}'$ of R, then*

$$\sum_{Q \in \mathcal{P}} f_Q \operatorname{Vol} Q = \sum_{Q' \in \mathcal{P}'} f_{Q'} \operatorname{Vol} Q'$$

I.e. the value of $\int_R f \, dV$ does not depend on which partition one chooses to compute it.

Finally, if $f, g : R \to \mathbf{R}$ are step functions on the same rectangle, then by choosing a common refinement of the two partitions on which f and g are based, we can assume that f and g are subordinate to the same partition $\mathcal{P}$ of R. Hence, for example,

$$f + g = \sum_{Q \in \mathcal{P}} (f_Q + g_Q) \mathbf{1}_Q$$

is also a step function. More generally, we have

Proposition 0.6. *Let $f, g : R \to \mathbf{R}$ be step functions on a rectangle $R \subset \mathbf{R}^n$. Then $f + g$, $f - g$, fg, $\max\{f, g\}$, and $\min\{f, g\}$ are all step functions on R. Moreover,*

- *if $f \leq g$, then $\int_R f \, dV \leq \int_R g \, dV$;*
- *$\int_R cf \, dV = c \int_R f \, dV$ for all $c \in \mathbf{R}$; and*
- *$\int_R (f + g) \, dV = \int_R f \, dV + \int_R g \, dV$.*

In the next section we will go on to define the integral of more general functions by approximating them from above and below with step functions.

0.2. **Riemann integrable functions.**

Definition 0.7. *Let $R \subset \mathbf{R}^n$ be a rectangle and $f : R \to \mathbf{R}$ be a bounded function. The* upper integral of f on R *is the quantity*

$$\overline{\int_R} f\,dV := \inf\left\{\int_R h\,dV : h \geq f \textit{ is a step function on } R\right\}.$$

The lower integral of f on R *is the quantity*

$$\underline{\int_R} f\,dV := \sup\left\{\int_R g\,dV : g \leq f \textit{ is a step function on } R\right\}.$$

If the upper and lower integrals are the same, then we say that f is Riemann integrable *on R. Either quantity is then called the* Riemann integral *of f on R and is denoted simply*

$$\int_R f\,dV$$

Typically, we will write simply 'integrable' in place of 'Riemann integrable' here. While there are other notions of integrability (most notably 'Lebesgue integrability'), we will not consider them here.

The lower integral of a given function is never greater than the upper integral. Hence we have the following convenient criterion for Riemann integrability.

Proposition 0.8. *A bounded function $f : R \to \mathbf{R}$ is Riemann integrable on a rectangle $R \subset \mathbf{R}^n$ if for any $\epsilon > 0$ there exist step functions $g \leq f \leq h$ on R such that*

$$\int_R (h - g)\,dV < \epsilon.$$

Integration interacts well with basic arithmetic, etc of functions.

Theorem 0.9. *Let $f, g : R \to \mathbf{R}$ be integrable functions on a rectangle $R \subset \mathbf{R}^n$. Then $f+g$, $f-g$, fg, $\max\{f,g\}$, $\min\{f,g\}$ are also integrable on R. Moreover,*

- *$f \geq g$ implies $\int_R f\,dV \geq \int_R g\,dV$;*
- *$\int_R (f+g)\,dV = \int_R f\,dV + \int_R g\,dV$;*
- *for any $c \in \mathbf{R}$, we have $\int_R cf\,dV = c\int_R f\,dV$.*

Proof. I discuss only the third item here and leave the other items as exercises. When $c = 0$, the third assertion just says that $\int_R 0\,dV = 0$, which requires little justification. Suppose instead that $c > 0$ (the case $c < 0$ is similar and left to the reader). Since f is integrable, Proposition 0.8 tells us that there are step functions $f_1 \leq f \leq f_2$ such that

$$\int_R (f_2 - f_1)\,dV < \frac{\epsilon}{c}$$

Then by Proposition 0.6 cf_1 and cf_2 are step functions satisfying $cf_1 \leq cf \leq cf_2$ and

$$\int_R (cf_1 - cf_2)\,dV = c\int_R (f_2 - f_1)\,dV < c \cdot \frac{\epsilon}{c} = \epsilon.$$

Hence cf is integrable. We also have by definition of the Riemann integral that

$$\int_R cf_1\,dV < \int_R cf\,dV < \int_R cf_2\,dV.$$

Likewise,

$$c\int_R f_1\,dV < c\int_R f\,dV < c\int_R f_2\,dV.$$

Note that by our choice of f_1 and f_2, the left and right sides of the previous two inequalities differ by less than ϵ. Moreover, Proposition 0.6 tells us that the left sides are equal and the right sides are equal. It follows that

$$\left|c\int_R f\,dV - \int_R cf\,dV\right| < \epsilon.$$

Since $\epsilon > 0$ is arbitrary, I conclude that $c\int_R f\,dV = \int_R cf\,dV$, as desired. □

The following theorem gives us our most important set of integrable functions.

Theorem 0.10. *Any continuous function $f : R \to \mathbf{R}$ on a rectangle $R \subset \mathbf{R}^n$ is integrable.*

Proof. Since R is compact and f is continuous, we know that f is actually *uniformly* continuous. So if $\epsilon > 0$ is given, we can choose $\delta > 0$ so that

$$x, y \in R \text{ and } \|x-y\| < \delta \text{ imply } |f(x) - f(y)| < \epsilon/\mathrm{Vol}\,R.$$

Now I let $\mathcal{P}$ be a partition of R whose rectangles Q all have 'diameter' smaller than δ. That is, if $Q \in \mathcal{P}$ and $x, y \in Q$, then $\|x-y\| < \delta$.

By the extreme value theorem f achieves a maximum and a minimum value on each $Q \in \mathcal{P}$. I denote these by M_Q and m_Q respectivly. My choice of δ and $\mathcal{P}$ guarantee that $M_Q - m_Q < \frac{\epsilon}{\mathrm{Vol}\,R}$. Now I define step functions

$$g := \sum_{Q\in\mathcal{P}} m_Q \mathbf{1}_Q \quad \text{and} \quad h := \sum_{Q\in\mathcal{P}} M_Q \mathbf{1}_Q.$$

It follows for each $Q \in \mathcal{P}$ and each $x \in Q$ that $g(x) = m_Q \le f(x) \le M_Q = h(x)$. Thus $g \le f \le h$ at all points in R. Moroever,

$$\int_R (h-g)\,dV = \sum_{Q\in\mathcal{P}} (M_Q - m_Q)\mathrm{Vol}\,Q < \frac{\epsilon}{\mathrm{Vol}\,R}\sum_{Q\in\mathcal{P}} \mathrm{Vol}\,Q = \frac{\epsilon}{\mathrm{Vol}\,R}\cdot \mathrm{Vol}\,R = \epsilon.$$

So by Proposition 0.8, f is integrable. □

The definition of Riemann integral, while an effective theoretical tool, is not a very practical means for actually computing integrals. In order to compute, we have a result that allows one to integrate functions 'one variable at a time.'

Theorem 0.11 (Fubini's Theorem)**.** *Let $R \subset \mathbf{R}^n$ be a rectangle of the form $R' \times [a_n, b_n]$ where $R' \subset \mathbf{R}^{n-1}$ and $[a_n, b_n] \subset \mathbf{R}$. Suppose that $f : R \to \mathbf{R}$ is an integrable function and that moreover*

- *For each fixed $x_n \in [a_n, b_n]$, the function $f(\cdot, x_n) : R' \to \mathbf{R}$ is integrable, and*
- *The integral $F(x_n) := \int_{R'} f(x', x_n)\,dV(x')$ is itself an integrable function on $[a_n, b_n]$.*

Then

$$\int_R f(x)\,dV(x) = \int_{a_n}^{b_n} F(x_n)\,dx_n = \int_{a_n}^{b_n}\left(\int_{R'} f(x', x_n)\,dV(x')\right)dx_n.$$

Proof. First I will show the theorem is true when f is a step function. Then I will show it holds more generally.

Step 1. Let $f = \sum_{Q\in\mathcal{P}} f_Q \mathbf{1}_Q$ be a step function subordinate to a partition $\mathcal{P}$ of R. Then there are partitions $\mathcal{P}'$ of R' and $\mathcal{I}$ of $[a_n, b_n]$ such that $Q \in \mathcal{P}$ if and only if $Q = Q' \times I$ for some $Q' \in \mathcal{P}'$ and $I \in \mathcal{I}$. In particular $\operatorname{Vol} Q = (\operatorname{Vol} Q')(\operatorname{Vol} I)$. Hence

$$\int_R f\, dV = \sum_{Q\in\mathcal{P}} f_Q \operatorname{Vol} Q = \sum_{I\in\mathcal{I}} \sum_{Q'\in\mathcal{P}'} f_{Q'\times I} (\operatorname{Vol} Q')(\operatorname{Vol} I).$$

Now for any $x_n \in [a_n, b_n]$, the 'slice' function $f(\cdot, x_n) : R' \to \mathbf{R}$ is a step function given by

$$f(\cdot, x_n) = \sum_{Q'\in\mathcal{P}'} f_{Q'\times I} \mathbf{1}_{Q'},$$

where $I \in \mathcal{I}$ is the interval containing x_n. Integrating over R' gives a function of x_n only:

$$F(x_n) := \int_{R'} f(x', x_n)\, dV(x') = \sum_{Q'\in\mathcal{P}'} f_{Q'\times I} \operatorname{Vol} Q'.$$

The value $F(x_n)$ depends only on the interval $I \in \mathcal{I}$ that contains x_n. That is, F is a step function on $[a_n, b_n]$ which can be written

$$F = \sum_{I\in\mathcal{I}} \left(\sum_{Q'\in\mathcal{P}'} f_{Q'\times I} \operatorname{Vol} Q' \right) \mathbf{1}_I.$$

Integrating over $[a_n, b_n]$ then gives

$$\int_{a_n}^{b_n} \int_{R'} f(x', x_n)\, dV(x')\, dx_n = \int_{a_n}^{b_n} F(x_n)\, dx_n = \sum_{I\in\mathcal{I}} \sum_{Q'\in\mathcal{P}'} f_{Q'\times I} (\operatorname{Vol} Q')(\operatorname{Vol} I) = \int_R f\, dV.$$

So Fubini's Theorem holds for step functions.

Step 2. Now suppose that $f : R \to \mathbf{R}$ is any integrable function satisfying the additional hypotheses of Fubini's Theorem. Then for any $\epsilon > 0$ there are step functions $g \le f \le h$ such that

$$\int_R (h - g)\, dV < \epsilon.$$

Observe that $g \le f \le h$ means in particular that for each fixed $x_n \in [a_n, b_n]$ that we have the inequality of 'slices' $g(\cdot, x_n) \le f(\cdot, x_n) \le h(\cdot, x_n)$. Thus we have that

$$\int_R g\, dV \le \int_R f\, dV \le \int_R h\, dV \tag{1}$$

and also for each fixed $x_n \in [a_n, b_n]$ that

$$\int_{R'} g(x', x_n)\, dV(x') \le \int_{R'} f(x', x_n)\, dV(x') \le \int_{R'} h(x', x_n)\, dV(x').$$

All three parts of this last inequality are integrable functions of x_n, the middle one by hypothesis and the left and right sides because they are step functions of x_n. Hence we further have

$$\int_{a_n}^{b_n} \int_{R'} g(x', x_n)\, dV(x')\, dx_n \le \int_{a_n}^{b_n} \int_{R'} f(x', x_n)\, dV(x')\, dx_n \le \int_{a_n}^{b_n} \int_{R'} h(x', x_n)\, dV(x')\, dx_n. \tag{2}$$

Since Fubini's Theorem holds for step functions, it follows that the left/right sides of (1) and (2) are the same. Hence

$$\left|\int_R f\,dV - \int_{a_n}^{b_n}\int_{R'} f(x',x_n)\,dV(x')\,dx_n\right| < \int_R (h-g)\,dV < \epsilon.$$

But $\epsilon > 0$ was arbitrary, so I conclude that $\int_R f\,dV = \int_{a_n}^{b_n}\int_{R'} f(x',x_n)\,dV(x')\,dx_n$. $\square$

0.3. **Integrating over regions.** Often we will want to integrate functions over bounded sets $\Omega \subset \mathbf{R}^n$ that are not rectangles. This can usually be done by choosing a rectangle R that contains Ω and then 'extending' the given function by declaring it's value to be 0 at all points outside Ω. Deciding when this actually works requires some more discussion.

Definition 0.12. *Let $R \subset \mathbf{R}^n$ be a rectangle and $S \subset R$ be any subset. We say that S has* zero volume *if for any $\epsilon > 0$ there exists a partition $\mathcal{P}$ of R such that*

$$\sum_{\substack{Q\in\mathcal{P}\\ Q\cap S\neq\emptyset}} \operatorname{Vol} Q < \epsilon.$$

With this definition in hand, we can generalize the fact that continuous functions are integrable.

Theorem 0.13. *If $f : R \to \mathbf{R}$ is a bounded function on a rectangle $R \subset \mathbf{R}^n$ and the subset of R where f fails to be continuous has zero volume, then f is integrable.*

The proof is an elaboration of the proof of Theorem 0.10.

Proof. Let $S \subset R$ be the set of points where f is discontinuous, and let $M > 0$ be an upper bound for $|f|$ on R. Then given $\epsilon > 0$, there is a partition $\mathcal{P}$ of R such that

$$\sum_{\substack{Q\in\mathcal{P}\\ Q\cap S\neq\emptyset}} \operatorname{Vol} Q < \epsilon/2M.$$

Combining all the other rectangles in $\mathcal{P}$, we get a compact set $K := \bigcup\{Q \in \mathcal{P} : Q \cap S = \emptyset\}$ on which f is continuous and therefore uniformly continuous. So as in the proof of Theorem 0.10, we can replace $\mathcal{P}$ by a refinement in order to arrange that for any $Q \in \mathcal{P}$ such that $Q \subset K$, the maximum and minimum values of f on Q' differ by less than $\frac{\epsilon}{2\operatorname{Vol} R}$.

Now we define step functions $g = \sum_{Q\in\mathcal{P}} g_Q$, $h = \sum_{Q\in\mathcal{P}} h_Q$, where

$$g_Q = \begin{cases} \min_{x\in Q} f(x) & \text{if } Q \subset K \\ -M & \text{otherwise} \end{cases} \quad \text{and} \quad h_Q = \begin{cases} \max_{x\in Q} f(x) & \text{if } Q \subset K \\ M & \text{otherwise.} \end{cases}$$

Then $g \le f \le h$ and

$$\begin{aligned} \int_R (h-g)\,dV &= \sum_{Q\subset K}(h_Q - g_Q)\operatorname{Vol} Q + \sum_{Q\not\subset K}(h_Q-g_Q)\operatorname{Vol} Q \\ &< \frac{\epsilon}{2\operatorname{Vol} R}\sum_{Q\subset K}\operatorname{Vol} Q + 2M\sum_{Q\not\subset K}\operatorname{Vol} Q \\ &< \frac{\epsilon}{2\operatorname{Vol} R}\cdot\operatorname{Vol} K + 2M\cdot\frac{\epsilon}{2M} \le \epsilon. \end{aligned}$$

So by Proposition 0.8, f is integrable. $\square$

In order to see the utility of this fact, we need one more definition.

Definition 0.14. *Let $R \subset \mathbf{R}^n$ be a rectangle and $U \subset R$ be an open set whose boundary ∂U has volume zero. Then we call the compact set $\Omega := \overline{U}$ a* region *in $\mathbf{R}^n$. The* volume *of Ω is the quantity*

$$\operatorname{Vol}\Omega := \int_R \mathbf{1}_\Omega \, dV.$$

More generally, if $f : \Omega \to \mathbf{R}$ is a continuous function, then we set

$$\int_\Omega f \, dV = \int_R \tilde{f} \, dV,$$

where $\tilde{f} : R \to \mathbf{R}$ is the function given by $\tilde{f}(x) = f(x)$ if $x \in \Omega$, and $\tilde{f}(x) = 0$ if $x \notin \Omega$.

Note that this definition is made possible by the previous theorem. That is, the functions $\mathbf{1}_\Omega$ and $\tilde{f}$ fail to be continuous only at points in $\partial\Omega$ which, by assumption, has volume zero. Note also that this definition is consistent with our previous definition of 'zero volume.'

Proposition 0.15. *If $R \subset \mathbf{R}^n$ is a rectangle and $S \subset R$ has zero volume (in the sense of Definition 0.12) then*

$$\int \mathbf{1}_S \, dV = 0.$$

That is, $\operatorname{Vol} S = 0$ *in the sense of Definition 0.14.*

Proof. Let $\epsilon > 0$ be given. By hypothesis, we can choose a partition $\mathcal{P}$ of R such that $\sum_{\substack{Q \in \mathcal{P} \\ Q \cap S \neq \emptyset}} \operatorname{Vol} Q < \epsilon$. So if we define the step function $h = \sum_{\substack{Q \in \mathcal{P} \\ Q \cap S \neq \emptyset}} \mathbf{1}_Q$, we find that

$$\int_R \mathbf{1}_S \, dV \leq \int_R h \, dV = \sum_{\substack{Q \in \mathcal{P} \\ Q \cap S \neq \emptyset}} \operatorname{Vol} Q < \epsilon.$$

As $\epsilon > 0$ was arbitrary, we must have $\int_R \mathbf{1}_S = 0$. □

There is one remaining loose end: how do we know when the boundary ∂U of an open set U has zero volume? This is hard to determine in general. However, the following will suffice for our purposes.

Theorem 0.16. *Let $R = R' \times [a_n, b_n]$ be a rectangle in $\mathbf{R}^n$ and $K' \subset R'$ be a compact set. If $f : K' \to [a_n, b_n]$, is continuous then the graph of f*

$$S := \{(x', f(x')) \in R : x' \in K'\}$$

has zero volume.

Proof. Given $\epsilon > 0$ we must find a partition of R such that the collective volume of those rectangles that meet S is smaller than ϵ.

First we choose a partition I of the target $[a_n, b_n]$ of f into intervals I of length no greater than $\epsilon' := \frac{\epsilon}{2 \operatorname{Vol} R'}$. Since K' is compact, the continuity of f on K' is *uniform*. That is, we may obtain $\delta > 0$ such that for any $x, y \in K'$, $\|x - y\| < \delta$ implies $\|f(x) - f(y)\| < \epsilon'$. Let $\mathcal{P}'$ be a partition of the 'base' rectangle R' such that all rectangles $Q' \in \mathcal{P}'$ have diameter smaller than δ. Then for any such Q', the maximum and minimum values of f on Q' differ by less than ϵ'.

Together, the partitions $\mathcal{I}$ of $[a_n, b_n]$ and $\mathcal{P}'$ of R' gives us a partition of R:

$$\mathcal{P} = \{Q = Q' \times I : Q' \in \mathcal{P}' \text{ and } I \in \mathcal{I}\}.$$

Then (and here is the most important point) for each $Q' \in \mathcal{P}'$, there are at most two intervals $I \in \mathcal{I}$ such that $Q = I \times Q' \in \mathcal{P}$ intersect S. Therefore,

$$\begin{aligned}
\sum_{\substack{Q \in \mathcal{P} \\ Q \cap S \neq \emptyset}} \operatorname{Vol} Q &= \sum_{Q' \in \mathcal{P}'} \sum_{\substack{I \in \mathcal{I} \\ (Q' \times I) \cap S \neq \emptyset}} (\operatorname{Vol} Q')(\operatorname{Vol} I) \\
&< \epsilon' \sum_{Q' \in \mathcal{P}'} \sum_{\substack{I \in \mathcal{I} \\ (Q' \times I) \cap S \neq \emptyset}} \operatorname{Vol} Q' \\
&\leq \epsilon' \sum_{Q' \in \mathcal{P}'} 2 \operatorname{Vol} Q' = \epsilon' \cdot 2 \operatorname{Vol} R' = \epsilon.
\end{aligned}$$

So $\mathcal{P}$ is the partition of R we seek. $\square$

The moral here is that we are now justified in trying to integrate continuous functions over any compact set $\Omega = \overline{U}$ where U is open and ∂U can be decomposed into finitely many pieces, each of which is the graph of a continuous function. Let us end by illustrating with an example: the unit ball $\Omega = \overline{B_3(0,1)} \subset \mathbf{R}^3$, which has boundary

$$\partial\Omega := \{(x,y,z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 = 1\} = \{(x,y,z) \in \mathbf{R}^3 : z = \pm\sqrt{1 - x^2 - y^2}\}.$$

Since both functions $-\sqrt{1-x^2-y^2}$ and $\sqrt{1-x^2-y^2}$ are continuous it follows that $\partial\Omega$ has volume zero. That is, $\overline{B_3(0,1)}$ is a region, and we are therefore entitled to try to integrate any continuous function $f : B_3(0,1) \to \mathbf{R}$.