According to our basic Cauchy Criterion for Integrability, a *bounded(!)* function $f$ on an interval $[a, b]$ is integrable if and only if, for all $\epsilon > 0$, there exists $\delta > 0$ such that for all partitions $\pi_1$, $\pi_2$,

$$\text{if } \operatorname{mesh}(\pi_1) < \delta \text{ and } \operatorname{mesh}(\pi_2) < \delta, \text{ then } |R(f, \pi_1) - R(f, \pi_2)| < \epsilon.$$

We will prove a *very* useful estimate for $|R(f, \pi_1) - R(f, \pi_2)|$. To do so, we'll work with the **oscillations** of $f$ on the intervals of the two partitions. To define the oscillations we will let $I_i := [x_{i-1}, x_i]$ and $J_j := [y_{j-1}, y_j]$ denote the *intervals* of the respective partitions, and let

$$\omega_{1i} := \sup_{x,\, x' \in I_i} f(x) - f(x'), \quad \omega_{2j} := \sup_{y,\, y' \in J_j} f(y) - f(y'),$$

denote the *oscillations* of $f$ on the intervals of the partitions $\pi_1$ and $\pi_2$ respectively. Please note that the subscripts here are *double* subscripts. Perhaps they should be written $\omega_{1,i}$ and $\omega_{2,j}$. I hope my choice will strain your eyes less. In the *statement* of the next Theorem, the dependence on the selection vectors will be omitted.

**Theorem:** *If $f$ is a real-valued function defined on the interval $[a, b]$, $\pi_1$ and $\pi_2$ are partitions of $[a, b]$, then*

$$\big|R(f, \pi_1) - R(f, \pi_2)\big| \le \sum_{i=1}^{n_{\pi_1}} \omega_{1i}\, \Delta x_i + \sum_{j=1}^{n_{\pi_2}} \omega_{2j}\, \Delta y_j, \tag{1}$$

*where $\omega_{1i}$ and $\omega_{2j}$ denote the oscillations of $f$ on the intervals $I_i$ and $J_j$ of the respective partitions.*

*Proof:* We will use the partition $\pi$ that consists of all the points of $\pi_1$ and of $\pi_2$, the *common refinement* of $\pi_1$ and $\pi_2$. We can write $\pi = \pi_1 \cup \pi_2$ to express this. We will denote the points of $\pi$ by $z_k$ and the intervals of $\pi$ by $K_k$. Then for each $i$ we can write

$$\Delta x_i = \sum_{K_k \subseteq I_i} \Delta z_k. \tag{2}$$

That is, the length of $I_i$, denoted $\Delta x_i$, is the sum of the lengths, $\Delta z_k$, of the intervals $K_k$ of $\pi$ that are contained in $I_i$. The formula (2) is true because each interval $K_k$ of $\pi$ is contained in exactly one interval $I_{i(k)}$ of $\pi_1$. In other words, $i(k)$ is the $i$ such that $K_k \subseteq I_{i(k)}$. There is a similar formula for $\Delta y_j$ and a similar function $j(k)$.

Now we need to bring in arbitrary "selection vectors" $\xi_1$ and $\xi_2$ to construct our Riemann sums. Thus for $R(f, \pi_1)$ we write

$$R(f, \pi_1, \xi_1) = \sum_{i=1}^{n_{\pi_1}} f(\xi_{1i}) \Delta x_i = \sum_{i=1}^{n_{\pi_1}} f(\xi_{1i}) \sum_{K_k \subseteq I_i} \Delta z_k = \sum_{k=1}^{n_\pi} f(\xi_{1i(k)}) \Delta z_k.$$

The last sum may need some explanation! Each interval $K_k$ is contained in one and only one of the intervals $I_i$. This $i$ we called $i(k)$. We can't reverse this though, because each $i$ can be associated with several $k$. Finally, each $k$ gets used, exactly once. This allows us to express $R(f, \pi_1)$ in a way that will become compatible with $R(f, \pi_2)$. Again, $i(k)$ is the $i$ such that $K_k \subseteq I_{i(k)}$. Notice that $\xi_{1i(k)}$ need not belong to $K_k$!

Now it is time to be clever. We choose a selection vector $\xi$ for $\pi$. Here, $\xi_k \in K_k$ for every $k$. Then

$$R(f, \pi_1, \xi_1) - R(f, \pi, \xi) = \sum_{k=1}^{n_\pi} \big(f(\xi_{1i(k)}) - f(\xi_k)\big) \Delta z_k.$$

We can regroup this sum once again, to return to the point of view of $\pi_1$:

$$R(f, \pi_1, \xi_1) - R(f, \pi, \xi) = \sum_{i=1}^{n_{\pi_1}} \sum_{K_k \subseteq I_i} \big(f(\xi_{1i}) - f(\xi_k)\big) \Delta z_k.$$

Each difference $f(\xi_{1i}) - f(\xi_k)$ uses two function values whose arguments are both in $I_i$, so $|f(\xi_{1i}) - f(\xi_k)| \le \omega_{1i}$. Our sum can now be replaced by an estimate:

$$|R(f, \pi_1, \xi_1) - R(f, \pi, \xi)| \le \sum_{i=1}^{n_{\pi_1}} \sum_{K_k \subseteq I_i} \omega_{1i} \Delta z_k = \sum_{i=1}^{n_{\pi_1}} \omega_{1i} \Delta x_i,$$

the equality here is true because of (2).

We complete the proof of the Theorem using a similar argument involving $\pi_2$ and $\pi$, together with the triangle inequality:

$$|R(f, \pi_1, \xi_1) - R(f, \pi_2, \xi_2)| \le |R(f, \pi_1, \xi_1) - R(f, \pi, \xi)| + |R(f, \pi, \xi) - R(f, \pi_2, \xi_2)| \le \sum_{i=1}^{n_{\pi_1}} \omega_{1i} \Delta x_i + \sum_{j=1}^{n_{\pi_2}} \omega_{2j} \Delta y_j.$$

**Applications of the estimate**

In one application, the smallness of the oscillations will be the key. In the other, the smallness of the mesh size will be important, because the sum of the oscillations can be estimated.

**Theorem:** *If $f$ is continuous on $[a, b]$ then $f$ is Riemann integrable on $[a, b]$.*

*Proof:* Because it is continuous $f$ is *uniformly* continuous on $[a, b]$ so if the meshes of partitions $\pi_1$ and $\pi_2$ are both sufficiently small, then the oscillations are small so by (1)

$$|R(f, \pi_1) - R(f, \pi_2)| \le \sum_{i=1}^{n_{\pi_1}} \omega_{1i} \Delta x_i + \sum_{j=1}^{n_{\pi_2}} \omega_{2j} \Delta y_j < 2\omega(b - a),$$

where $\omega$ denotes the largest of the oscillations appearing in the two sums. Since $\omega$ can be made as small as we wish, the Theorem follows.

**Theorem:** *If $f$ is monotone on $[a, b]$ then $f$ is Riemann integrable on $[a, b]$.*

*Proof:* Without loss of generality we assume that $f$ is increasing. Then for a partition $\pi$ of $[a, b]$,

$$\sum_{i=1}^{n_\pi} \omega_i = f(b) - f(a), \quad \text{so} \quad \sum_{i=1}^{n_\pi} \omega_i \Delta x_i \le (f(b) - f(a)) \operatorname{mesh}(\pi).$$

The Theorem follows now as an application of (1), by choosing a small enough mesh size.

**A criterion for Riemann integrability**

The next Theorem says, very roughly, that for Riemann integrability to hold, oscillations that are large can be gathered into a "small" closed set. It main use will be as a tool to prove a more understandable Theorem.

**Theorem (The Oscillation Criterion):** *A real-valued function $f$ on $[a, b]$ is Riemann integrable on $[a, b]$ if and only if $f$ is bounded and*

(3) *for all positive $\beta$ and $\eta$ there exists $\delta > 0$ such that $\operatorname{mesh}(\pi) < \delta$ implies that* $\displaystyle\sum_{\omega_i > \eta} \Delta x_i < \beta.$

*Proof:* If $f$ is Riemann integrable on $[a, b]$ then $f$ is bounded. We take, as our $\epsilon$ in the Cauchy Criterion for integrability, $\frac{1}{2}\eta\beta$. Then there exists $\delta > 0$ such that $\operatorname{mesh}(\pi_1) < \delta$ and $\operatorname{mesh}(\pi_2) < \delta$ implies $|R(f, \pi_1) - R(f, \pi_2)| < \frac{1}{2}\eta\beta$. We can choose $\pi_2 = \pi_1 = \pi$, and construct selection vectors $\xi_1$ and $\xi_2$ so that $f(\xi_{1i}) - f(\xi_{2i}) \ge \frac{1}{2}\omega_i$. Then

$$\frac{1}{2}\eta\beta > |R(f, \pi_1, \xi_1) - R(f, \pi_2, \xi_2)| = \sum_{i=1}^{n_{\pi_1}} \big(f(\xi_{1i}) - f(\xi_{2i})\big) \Delta x_i \ge \sum_{i=1}^{n_{\pi_1}} \frac{1}{2}\omega_i \Delta x_i \ge \sum_{\omega_i > \eta} \frac{1}{2}\eta \Delta x_i,$$

so $f$ is bounded and (3) holds if $f$ is Riemann integrable on $[a, b]$.

If $f$ is bounded on $[a, b]$ and (3) holds then when $\text{mesh}(\pi_1) < \delta$ and $\text{mesh}(\pi_2) < \delta$ we use (1) and find that

$$\begin{aligned} |R(f,\pi_1) - R(f,\pi_2)| &\le \sum_{i=1}^{n_{\pi_1}} \omega_{1i}\Delta x_i + \sum_{j=1}^{n_{\pi_2}} \omega_{2j}\Delta y_j = \sum_{\omega_{1i}\le\eta} \omega_{1i}\Delta x_i + \sum_{\omega_{1i}>\eta} \omega_{1i}\Delta x_i + \text{terms}' \\ &\le \eta \sum_{\omega_{1i}\le\eta} \Delta x_i + \overline{\omega} \sum_{\omega_{1i}>\eta} \Delta x_i + (\text{similar terms})' \\ &< 2\eta(b-a) + 2\overline{\omega}\beta, \end{aligned}$$

where $\overline{\omega}$ denotes the oscillation of $f$ over $[a, b]$. The "twos" arise to include the primed terms. We can now choose $\beta$ and $\eta$ to ensure that this last quantity is as small as desired. The Oscillation Criterion Theorem follows.

The following Corollary describes the fact that $\omega_i \Delta x_i$ is the area of a rectangle that contains the part of the graph of $f$ that lies above $I_i = [x_{i-1}, x_i]$.

**Corollary (Graph-Covering Criterion):** *A real-valued function $f$ on $[a, b]$ is Riemann integrable on $[a, b]$ if and only if $f$ is bounded and*

$(3')$ *for all* $\epsilon > 0$ *there exists* $\delta > 0$ *such that* $\text{mesh}(\pi) < \delta$ *implies that* $\displaystyle\sum_{i=1}^{n_\pi} \omega_i \Delta x_i < \epsilon.$

*Proof:* If $f$ is bounded and $(3')$ holds then (1) can be applied to show that the Cauchy criterion is satisfied.

The other part of the proof is contained in the first part of the proof of the Oscillation Criterion Theorem. If $f$ is Riemann integrable on $[a, b]$ then $f$ is bounded. We take, as our $\epsilon$ in the Cauchy Criterion for integrability, $\frac{1}{2}\epsilon$. Then there exists $\delta > 0$ such that $\text{mesh}(\pi_1) < \delta$ and $\text{mesh}(\pi_2) < \delta$ implies $|R(f,\pi_1) - R(f,\pi_2)| < \frac{1}{2}\epsilon$. We can choose $\pi_2 = \pi_1 = \pi$, and construct selection vectors $\xi_1$ and $\xi_2$ so that $f(\xi_{1i}) - f(\xi_{2i}) \ge \frac{1}{2}\omega_i$. Then

$$\frac{1}{2}\epsilon > |R(f,\pi_1,\xi_1) - R(f,\pi_2,\xi_2)| = \sum_{i=1}^{n_\pi} \big(f(\xi_{1i}) - f(\xi_{2i})\big)\Delta x_i \ge \sum_{i=1}^{n_{\pi_1}} \frac{1}{2}\omega_i \Delta x_i.$$

Since our partition $\pi$ is an arbitrary one whose mesh is less than $\delta$, $(3')$ holds.

**An application of the Corollary**

**Theorem:** *If $f$ is Riemann integrable on $[a, b]$ then so is $|f|$, and*

$$\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx.$$

*Proof:* To show that $|f|$ is Riemann integrable we note that it is bounded because $f$ is. If we let $|\omega|_i$ denote the oscillation of $|f|$ on the interval $I_i = [x_{i-1}, x_i]$ then (using the triangle inequality variant $||u| - |v|| \le |u - v|$)

$$|\omega|_i = \sup_{x,\ y\in I_i} \big||f(x)| - |f(y)|\big| \le \sup_{x,\ y\in I_i} \big|f(x) - f(y)\big| = \omega_i.$$

Then $(3')$ holds for $|f|$ if it holds for $f$.

To verify the inequality we note that $|R(f,\, \pi,\, \xi)| \le R(|f|,\, \pi,\, \xi)$. The terms on each side approach the desired integralss as the mesh size approaches zero.

**Three problems whose assertions are useful in the work that follows.**

**Problem 1:** *A collection $\mathcal{J}$ of intervals is called non-overlapping if, whenever $I \in \mathcal{J}$ and $J \in \mathcal{J}$, then $I \cap J$ is either empty or consists of a single point. Prove that if $\mathcal{J}$ is a finite, non-overlapping collection of compact intervals,* *all of which are contained in a compact interval* $[A,\, B]$, *then the sum of the lengths of the intervals in the collection is at most* $B - A$. Hint: Draw a picture! Recall telescoping sums! Think about the Alternating Series Test!

You might want to assign yourself two Exercises to help with the next problem: Find a nice "formula" for the intersection of two intervals (always an interval, since the empty set "is" an interval, e.g., $[1,\, -1] = \emptyset$). Find out what has to be true about two intervals in order to be sure that their union is an interval.

**Problem 2:** *Prove that if the union of a finite collection of bounded open intervals contains the compact interval* $[A,\, B]$, *then the sum of the lengths of the open intervals in that collection is strictly greater than* $B - A$. *Show also that if the intervals are not open, but are compact, then "strictly greater than" need not hold, and must be replaced by "at least."*

**Problem 3:** *Prove that if the union of a finite collection of compact intervals contains the union of a finite non-overlapping family of compact intervals, then the sum of the lengths of the intervals in the non-overlapping family is at most the sum of the lengths of the intervals in the "containing" collection.*

**Discontinuities, oscillations and: An Important Criterion for Riemann integrability**

Continuity plays an important rôle in Riemann integration. We will presently show that a Riemann integrable function has to be continuous except (possibly) on a "very small set" (see the Appendix now or later). To get started we will define a sort of "oscillation at a point" that we can use to identify points of discontinuity.

What does it mean to say that $f$ is *dis*continuous at $\widehat{x}_o$? Answer: *There exists* $\epsilon > 0$ *such that*

$$\text{(4)} \qquad \textit{for all } \delta > 0 \textit{ there exists } x \textit{ such that } |x - \widehat{x}_o| < \delta \textit{ and } |f(x) - f(\widehat{x}_o)| \geq \epsilon.$$

Let us set $\omega(\widehat{x}_o) := \sup\{\epsilon > 0 : (4) \text{ holds for } \epsilon \text{ and } f\}$. Then (using the convention that the supremum of the empty subset of the non-negative reals is 0) $\omega(\widehat{x}_o) > 0$ precisely when $f$ is discontinuous at $\widehat{x}_o$, or, $\omega(\widehat{x}_o) = 0$ precisely when $f$ is continuous at $\widehat{x}_o$. For the Dirichlet function, $\omega(\widehat{x}_o) = 1$ for all $\widehat{x}_o$. Now let $\pi$ be a partition of $[a,\, b]$. If $I_i := [x_{i-1},\, x_i]$ is one of the intervals of $\pi$ and $\omega_i \leq \epsilon$ then $|f(x) - f(y)| \leq \epsilon$ for all $x$ and $y$ in $I_i$, so no point $\hat{x}_o$ of $(x_{i-1},\, x_i)$ can have $\omega(\widehat{x}_o) > \epsilon$. It *is* possible that $\omega(x_{i-1}) > \epsilon$ or $\omega(x_i) > \epsilon$, but then it would have to be true that $\omega_{i-1} > \epsilon$ or that $\omega_{i+1} > \epsilon$. We can formalize what we have discovered:

*If* $\pi$ *is a partition of* $[a,\, b]$ *then every point* $\widehat{x}_o$ *at which* $\omega(\widehat{x}_o) > \epsilon$ *belongs to an interval* $I_i$ *of* $\pi$ *whose oscillation,* $\omega_i$, *exceeds* $\epsilon$. In symbols,

$$\text{(5)} \qquad \{\widehat{x}_o : \omega(\widehat{x}_o) > \epsilon\} \subseteq \bigcup_{\omega_i > \epsilon} I_i.$$

We will connect this "covering property," (5), and the "oscillation criterion" (3). What follows is a Theorem whose converse we will also prove. "Continuous almost everywhere" is explained in the Appendix.

**Theorem:** *If* $f$ *is Riemann integrable on* $[a,\, b]$ *then* $f$ *is continuous almost everywhere on* $[a,\, b]$.

*Proof:* Suppose $f$ is Riemann integrable on $[a,\, b]$. Let us denote by $D$ the set of discontinuities of $f$. We can then write (Divide and Conquer)

$$D = \bigcup_{n=1}^{\infty} D_n, \quad \text{where} \quad D_n = \left\{\widehat{x}_o \in [a,\, b] : \omega(\widehat{x}_o) > \frac{1}{n}\right\}.$$

Our objective is to show that each set $D_n$ is a null set. Then, as explained in the Appendix, $D$ is a null set because $D$ is the countable union of null sets. For a fixed $n$, say $n = N$, and any given $\beta > 0$, no matter how small, we can define $\eta = 1/N$ and apply (3) (Oscillation Criterion) to find the $\delta > 0$ guaranteed by (3). Then we choose *any* partition $\pi$ such that $\text{mesh}(\pi) < \delta$. We next use the covering property (5) and the conclusion of (3), which gives (since $\Delta x_i$ is the length of $I_i$)

$$D_N = \{\widehat{x}_o : \omega(\widehat{x}_o) > 1/N\} \subseteq \bigcup_{\omega_i > 1/N} I_i \quad \text{and} \quad \sum_{\omega_i > 1/N} \Delta x_i < \beta,$$

so that $D_N$ is a set of measure zero, or a null set (this is explained in the Appendix). Since the union of an at most countable collection of null sets is a null set, $D$ is a null set, which is what we had to prove.

The converse of this Theorem has a more involved proof that has some features in common with the proof that a finite open cover of a compact set has a Lebesgue number. In fact the proof will make direct use of the compactness of $[a, b]$. The proof also uses this: the length of an interval stretched about its center is equal to the stretching factor times the length of the interval stretched.

**Theorem:** *If $f$ is bounded and continuous almost everywhere on $[a, b]$ then $f$ is Riemann integrable on $[a, b]$.*

*Proof:* To prove this Theorem we will show that the "oscillation criterion" in (3) holds. Thus we let $\eta > 0$ and $\beta > 0$ be given.

Again let us denote by $D$ the set of discontinuities of $f$. We can (in principle) find an at most countable cover of $D$ by open intervals $(c_n, d_n)$ whose series of lengths sums to less than $\beta/3$:

$$D \subseteq \bigcup_{n=1}^{\infty} (c_n, d_n) \quad \text{and} \quad \sum_{n=1}^{\infty} (d_n - c_n) < \beta/3.$$

For each $x \in [a, b] \setminus D$, $f$ is continuous, so there exists $\lambda_x > 0$ such that for all $t \in (x - \lambda_x, x + \lambda_x)$ we have $|f(t) - f(x)| < \eta/4$.

The intervals $(c_n, d_n)$ and the intervals $(x - \lambda_x/2, x + \lambda_x/2)$ with $x \in [a, b] \setminus D$ comprise an open cover of $[a, b]$. We can (in principle) find a finite subcover. Some of the intervals in the finite subcover may be among the $(c_n, d_n)$; let us call their closures "$D$-intervals." The closures of the intervals in the finite subcover of the form $(x - \lambda_x/2, x + \lambda_x/2)$, with $x \in [a, b] \setminus D$, we will call "$C$-intervals." Note: we cut the intervals in half because we need for the closures of the "halves" to to be contained in the "wholes." If an interval is simultaneously of both kinds we will call it a $C$-interval and forget that it is also a $D$-interval.

Next we use $a$ and $b$ and *all* the endpoints *(that lie in* $[a, b]$) of the $C$ and $D$ intervals to form a partition $\pi_o$ of $[a, b]$. We will denote the points of $\pi_o$ by $x_k^o$. This is where we will find our $\delta$. We let $\delta$ be the length of the *shortest* interval in $\pi_o$.

$$\text{That is, } \delta := \min_{1 \le k \le n_{\pi_o}} x_k^o - x_{k-1}^o.$$

How does this particular choice help? It gives us two pieces of information:

*The length of each $C$-interval and that of each $D$-interval is at least $\delta$;*

*If $\pi$ is a partition of $[a, b]$ and $\text{mesh}(\pi) < \delta$ then no interval $I_i$ of $\pi$ can contain more than one point of $\pi_o$.*

**The hard work begins:** we suppose that $\text{mesh}(\pi) < \delta$. We examine an interval $I_i$ from $\pi$. There are two possibilities to consider: whether or not $I_i$ meets one of the $D$-intervals.

**Case 1:** *$I_i$ meets a $D$-interval.* Then there is some point $\overline{x}$ in $I_i \cap [c, d]$, where $[c, d]$ is a $D$-interval. For brevity, let $L := d - c$, the length of $[c, d]$. Then $[c - L, d + L]$ is a "stretched out" version of $[c, d]$, whose length is $3L$. The point of this stretching is that $I_i$ has to be contained in $[c - L, d + L]$. This is so because the length of $I_i$ is *less* than $\delta$, and $\delta \le L$. Then every point of $I_i$ is within $\delta$ of $\overline{x}$, and $\overline{x} \in [c, d]$, so every point of $I_i$ is either within $L$ of $c$ or is within $L$ of $d$, which means that $(c - L, d + L)$ has to swallow up the interval $I_i$. The sum of the lengths of the tripled $D$-intervals is less than $3 \cdot (\beta/3) = \beta$, and all the intervals of $\pi$ that meet some $D$-interval are non-overlapping, and are contained in the union of the tripled $D$-intervals. Thus (by Problem 3) the sum of the lengths of the intervals of $\pi$ that meet some $D$-interval is less than $\beta$:

$$\sum_{I_i \text{ meets a } D\text{-interval}} \Delta x_i < \beta. \tag{6}$$

**Case 2:** *$I_i$ meets no $D$-interval.* In this case, we want to show that $\omega_i \le \eta$, because then all the $I_i$ with $\omega_i > \eta$ will have to be among those that meet a $D$-interval, and the sum of the lengths of *those* $I_i$ is already known to be less than $\beta$. The key in this Case is that $I_i$ is contained in the union of the $C$-intervals that $I_i$ meets.

Now $I_i$ either contains one point of $\pi_o$ or no points of $\pi_o$. If $I_i$ contains no points of $\pi_o$ then $x_{i-1}$ belongs to some $C$-interval, say to $[\widehat{x}-\widehat{\lambda}/2,\, \widehat{x}+\widehat{\lambda}/2]$. The endpoints of this interval are in $\pi_o$ (or one or both lie outside $[a,\, b]$, in which case an endpoint of $[\widehat{x}-\widehat{\lambda}/2,\, \widehat{x}+\widehat{\lambda}/2]$ that lies outside $[a,\, b]$ is replaced by $a$ or by $b$). We can then say that $\widehat{x}-\widehat{\lambda}/2$ " = " $x^o_\ell$ and $\widehat{x}+\widehat{\lambda}/2$ " = " $x^o_{\ell'}$, where $\ell < \ell'$. Then $x^o_\ell < x_{i-1} < x^o_{\ell'}$ We cannot have $x_i \ge x^o_{\ell+1}$ because then $I_i$ would contain a point of $\pi_o$. Thus $x^o_\ell < x_{i-1} < x_i < x^o_{\ell+1} \le x^o_{\ell'}$, so $I_i$ is not only contained in the union of $C$-intervals, $I_i$ is contained in a *single* $C$-interval, $[\widehat{x}-\widehat{\lambda}/2,\, \widehat{x}+\widehat{\lambda}/2]$. Therefore,

(7) if $x \in I_i$ and $y \in I_i$ then $|f(x)-f(y)| \le |f(x)-f(\widehat{x})| + |f(\widehat{x})-f(y)| < \eta/4+\eta/4 = \eta/2$ so $\omega_i \le \eta/2$.

Even if $I_i$ contains a point $x^o_k$ of $\pi_o$ it is still possible that $I_i$ is contained in a single $C$-interval. If so, (7) shows that $\omega_i \le \eta/2$.

We can say for sure that if $I_i$ meets no $D$-interval, and a point $x^o_k$ of $\pi_o$ is in $I_i$, then $x^o_k$ has to be an endpoint of a $C$-interval!

Thus if $I_i$ contains a point $x^o_k$ of $\pi_o$ but $I_i$ is *not* contained in a single $C$-interval then $x^o_k$ must be the right end-point of one $C$-interval, say $[\widehat{x}-\widehat{\lambda}/2,\, \widehat{x}+\widehat{\lambda}/2]$, and $x^o_k$ must be the left end-point of another $C$-interval, say $[\widehat{y}-\widehat{\mu}/2,\, \widehat{y}+\widehat{\mu}/2]$. Hence $\widehat{x}+\widehat{\lambda}/2 = x^o_k = \widehat{y}-\widehat{\mu}/2$. To finish the proof, we consider any two points, $x$ and $y$, that are both in $I_i \subseteq [\widehat{x}-\widehat{\lambda}/2,\, x^o_k] \cup [x^o_k,\, \widehat{y}+\widehat{\mu}/2]$. If both $x$ and $y$ are in the same one of the two intervals, then we can use (7). Otherwise, suppose $x < x^o_k < y$. Then $x \in [\widehat{x}-\widehat{\lambda}/2,\, \widehat{x}+\widehat{\lambda}/2]$ and $y \in [\widehat{y}-\widehat{\mu}/2,\, \widehat{y}+\widehat{\mu}/2]$ so

$$\begin{aligned}|f(x)-f(y)| &\le |f(x)-f(x^o_k)| + |f(x^o_k)-f(y)| \\ &\le |f(x)-f(\widehat{x})| + |f(\widehat{x})-f(x^o_k)| \\ &\quad + |f(x^o_k)-f(\widehat{y})| + |f(\widehat{y})-f(y)| \\ &< \eta/4+\eta/4+\eta/4+\eta/4 = \eta, \quad \text{and} \quad \omega_i \le \eta.\end{aligned}$$

As mentioned at the start of Case 2, this was all that we had to show.

The two Theorems we have proved can be combined to give a succinct Riemann integrability criterion. The Theorem we state next can also be obtained as a consequence of the theory of the Lebesgue integral. Our proof, while somewhat difficult, is "elementary," which, as you can now see, does not necessarily mean "easy!"

**Theorem:** *If $f$ is bounded and real-valued on a closed and bounded interval $[a,\, b]$, then $f$ is Riemann integrable there if and only if $f$ is continuous almost everywhere on $[a,\, b]$.*

**Appendix: on sets of measure zero, and "almost everywhere"**

A set $S$ of real numbers is a *set of Lebesgue measure zero* if for *every* $\epsilon > 0$ there exists a sequence of intervals $I_n$ (that can be closed, open, half open-half closed), with endpoints $a_n < b_n$, such that $\sum_{n=1}^{\infty} b_n - a_n < \epsilon$, AND (important!) every point of $S$ is in at least one of the intervals $I_n$. Another way to say the last condition is "$S$ is contained in the (union of the) intervals $I_n$." In symbols, this can be compactly written $S \subseteq \bigcup_{n=1}^{\infty} I_n$.

Note: The sequence $\{I_n\}$ of intervals can be a *finite* sequence. By the way, $\{I_n\}$ usually depends on $\epsilon$!

**Examples**
**(A1)** *A singleton, namely a set that contains only one point, is a set of Lebesgue measure zero, or set of measure zero, for short, or (even shorter) a null set.* Thus if $S$ contains only the point $x_o$, the single interval $I_1 := (x_o-\epsilon/3,\, x_o+\epsilon/3)$ gives us "a sequence of intervals $I_n$, with endpoints $a_n \le b_n$, such that $\sum_{n=1}^{\infty} b_n - a_n < \epsilon$." This is so because here $\sum_{n=1}^{\infty} b_n - a_n = b_1 - a_1 = 2\epsilon/3$. Of course, the interval $[x_o,\, x_o]$ works for all $\epsilon$, while $[x_o,\, x_o)$ does *not* work.
Exercise: Verify that a finite set is a null set.
**(2)** *A countable set is a null set.* Thus, $S = \{x_1,\, x_2,\, \ldots\, x_n,\, \ldots\}$, and no two of the $x_n$'s are the same. The set $\mathbb{Z}$ of integers is countable. So is the set $\mathbb{Q}$ of rational numbers. Let us see why an at most countable set is a null set. Suppose $\epsilon > 0$ is given. We consider the intervals $I_n := (x_n - \epsilon/2^{n+1},\, x_n + \epsilon/2^{n+1})$, and by construction, we have $S \subseteq \bigcup_{n=1}^{\infty} I_n$. The sum of the lengths of all these intervals is given by the series $\sum_{n=1}^{\infty} \epsilon 2^{-n} = \epsilon$. This does not *quite* meet the conditions of the definition (we have sum $= \epsilon$, not sum $< \epsilon$. So, we might change the definition

to "$\sum_{n=1}^{\infty} b_n - a_n \leq \epsilon$," or we can just make $I_1$ have length $\epsilon/3$ instead of $\epsilon/2$. I prefer to change the definition, but the one given here is the one that is usually found in textbooks.
**(A3) Exercise:** *Prove that the Cantor set is a null set.*

**There are two basic theorems about sets of measure zero.**

**(A4) Theorem:** *The union of a finite or countable collection of null sets is a null set.*

**(A5) Theorem:** *A set $S$ of real numbers is a null set if and only if there exists a sequence $\{I_n\}$ of intervals, the sum of whose lengths is finite, such that every point of the set $S$ is contained in infinitely many of the intervals $I_n$.*

Theorem (A4) is proved with the method used to show that an at most countable set is a null set – but the $n$-th set of the collection is "enclosed" in a *sequence* of intervals with total-of-lengths less than $\epsilon/2^n$. Then we note that we can arrange a sequence of sequences into a one sequence by taking the first term of the first sequence as our first item, taking our second item from the second sequence, but we take as the *third* item the *second* term of the *first* sequence. Then we choose as our fourth item the second term of the second sequence, choose the first term of the third sequence as our fifth item, and so on. This is easy to see if you draw a picture with the sequences arranged as rows of dots. . .

Theorem (A5) is a bit harder to prove, and is used less frequently than Theorem (A4). We'll skip it.

**almost everywhere**

We say that a statement about the real numbers $x$ in an interval $I$ is *"true almost everywhere in $I$"* if the set $S$ of numbers $x$ for which the statement is NOT true is a null set. Thus, we can say that a function that is continuous except in a null set, i.e., set of measure zero, is "continuous almost everywhere."

**Acknowledgement:** The Spring 2004 Math 5616H class made important contributions to this note!