Line Integrals

11.1 Basic Properties

11.1.1 Length

I will give a discussion of what is meant by a line integral which is independent of the earlier material on Lebesgue integration. Line integrals are of fundamental importance in physics and in the theory of functions of a complex variable.

Definition 11.1.1 *Let $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ be a function. Then $\boldsymbol{\gamma}$ is of bounded variation if*

$$\sup \left\{ \sum_{i=1}^{n} |\boldsymbol{\gamma}(t_i) - \boldsymbol{\gamma}(t_{i-1})| : a = t_0 < \cdots < t_n = b \right\} \equiv V(\boldsymbol{\gamma}, [a,b]) < \infty$$

where the sums are taken over all possible lists, $\{a = t_0 < \cdots < t_n = b\}$. The set of points traced out will be denoted by $\boldsymbol{\gamma}^* \equiv \boldsymbol{\gamma}([a,b])$. The function $\boldsymbol{\gamma}$ is called a parameterization of $\boldsymbol{\gamma}^*$. The set of points $\boldsymbol{\gamma}^*$ is called a rectifiable curve. If a set of points $\boldsymbol{\gamma}^* = \boldsymbol{\gamma}([a,b])$ where $\boldsymbol{\gamma}$ is continuous and $\boldsymbol{\gamma}$ is one to one on (a,b) such that also $\boldsymbol{\gamma}(t) \neq \boldsymbol{\gamma}(a)$ if $t \in (a,b)$ and $\boldsymbol{\gamma}(t) \neq \boldsymbol{\gamma}(b)$ if $t \in (a,b)$, then $\boldsymbol{\gamma}^*$ is called a simple curve. A closed curve is one which has a parameterization $\boldsymbol{\gamma}$ defined on an interval $[a,b]$ such that $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$.

The case of most interest is for simple curves. It turns out that in this case, the above concept of length is a property which $\boldsymbol{\gamma}^*$ possesses independent of the parameterization $\boldsymbol{\gamma}$ used to describe the set of points $\boldsymbol{\gamma}^*$. To show this, it is helpful to use the following lemma.

Lemma 11.1.2 *Let $\phi : [a,b] \to \mathbb{R}$ be a continuous function and suppose ϕ is $1-1$ on (a,b). Then ϕ is either strictly increasing or strictly decreasing on $[a,b]$. Furthermore, ϕ^{-1} is continuous.*

Proof: First it is shown that ϕ is either strictly increasing or strictly decreasing on (a,b).

If ϕ is not strictly decreasing on (a,b), then there exists $x_1 < y_1$, $x_1, y_1 \in (a,b)$ such that

$$(\phi(y_1) - \phi(x_1))(y_1 - x_1) > 0.$$

If for some other pair of points, $x_2 < y_2$ with $x_2, y_2 \in (a,b)$, the above inequality does not hold, then since ϕ is $1-1$,

$$(\phi(y_2) - \phi(x_2))(y_2 - x_2) < 0.$$

Let $x_t \equiv tx_1 + (1-t)x_2$ and $y_t \equiv ty_1 + (1-t)y_2$. Then $x_t < y_t$ for all $t \in [0,1]$ because

$$tx_1 \leq ty_1 \text{ and } (1-t)x_2 \leq (1-t)y_2$$

with strict inequality holding for at least one of these inequalities since not both t and $(1-t)$ can equal zero. Now define

$$h(t) \equiv (\phi(y_t) - \phi(x_t))(y_t - x_t).$$

Since h is continuous and $h(0) < 0$, while $h(1) > 0$, there exists $t \in (0,1)$ such that $h(t) = 0$. Therefore, both x_t and y_t are points of (a,b) and $\phi(y_t) - \phi(x_t) = 0$ contradicting the assumption that ϕ is one to one. It follows ϕ is either strictly increasing or strictly decreasing on (a,b).

This property of being either strictly increasing or strictly decreasing on (a,b) carries over to $[a,b]$ by the continuity of ϕ. Suppose ϕ is strictly increasing on (a,b), a similar argument holding for ϕ strictly decreasing on (a,b). If $x > a$, then pick $y \in (a,x)$ and from the above, $\phi(y) < \phi(x)$. Now by continuity of ϕ at a,

$$\phi(a) = \lim_{x \to a+} \phi(z) \le \phi(y) < \phi(x).$$

Therefore, $\phi(a) < \phi(x)$ whenever $x \in (a,b)$. Similarly $\phi(b) > \phi(x)$ for all $x \in (a,b)$.

It only remains to verify ϕ^{-1} is continuous. Suppose then that $s_n \to s$ where s_n and s are points of $\phi([a,b])$. It is desired to verify that $\phi^{-1}(s_n) \to \phi^{-1}(s)$. If this does not happen, there exists $\varepsilon > 0$ and a subsequence, still denoted by s_n such that $\left|\phi^{-1}(s_n) - \phi^{-1}(s)\right| \ge \varepsilon$. Using the sequential compactness of $[a,b]$ there exists a further subsequence, still denoted by n, such that $\phi^{-1}(s_n) \to t_1 \in [a,b], t_1 \ne \phi^{-1}(s)$. Then by continuity of ϕ, it follows $s_n \to \phi(t_1)$ and so $s = \phi(t_1)$. Therefore, $t_1 = \phi^{-1}(s)$ after all. This proves the lemma.

Now suppose $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$ are two parameterizations of the simple curve $\boldsymbol{\gamma}^*$ as described above. Thus $\boldsymbol{\gamma}([a,b]) = \boldsymbol{\gamma}^* = \boldsymbol{\eta}([c,d])$ and the two continuous functions $\boldsymbol{\gamma}, \boldsymbol{\eta}$ are of bounded variation and one to one on their respective open intervals. I need to show the two definitions of length yield the same thing with either parameterization. Since $\boldsymbol{\gamma}^*$ is compact, it follows from Theorem 5.1.3 on Page 84, both $\boldsymbol{\gamma}^{-1}$ and $\boldsymbol{\eta}^{-1}$ are continuous. Thus $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta} : [c,d] \to [a,b]$ is continuous. It is also uniformly continuous because $[c,d]$ is compact. Let $\mathcal{P} \equiv \{t_0, \cdots, t_n\}$ be a partition of $[a,b], t_0 < t_1 < \cdots$, such that

$$0 \le V(\boldsymbol{\gamma}, [a,b]) - \sum_{k=1}^{n} |\boldsymbol{\gamma}(t_k) - \boldsymbol{\gamma}(t_{k-1})| < \varepsilon$$

Note the sums approximating the total variation are all no larger than the total variation because when another point is added in to the partition, it is an easy exercise in the triangle inequality to show the corresponding sum either becomes larger or stays the same.

Let $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}(s_k) = t_k$ so that $\{s_0, \cdots, s_n\}$ is a partition of $[c,d]$. By the lemma, the s_k are either strictly decreasing or strictly increasing as a function of k, depending on whether $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}$ is increasing or decreasing. Thus $\boldsymbol{\gamma}(t_k) = \boldsymbol{\eta}(s_k)$ and so

$$V(\boldsymbol{\gamma}, [a,b]) - V(\boldsymbol{\eta}, [c,d]) \le V(\boldsymbol{\gamma}, [a,b]) - \sum_{k=1}^{n} |\boldsymbol{\eta}(s_k) - \boldsymbol{\eta}(s_{k-1})| < \varepsilon$$

It follows

$$V(\boldsymbol{\gamma}, [a,b]) \le V(\boldsymbol{\eta}, [c,d]) + \varepsilon$$

and since ε is arbitrary, this shows $V(\boldsymbol{\gamma}, [a,b]) \le V(\boldsymbol{\eta}, [c,d])$. Turning the argument around reverses the inequality. This proves the following fundamental theorem.

Theorem 11.1.3 *Let Γ be a simple curve and let $\boldsymbol{\gamma}$ be a parameterization for Γ where $\boldsymbol{\gamma}$ is one to one on (a,b), continuous on $[a,b]$ and of bounded variation. Then the total variation*

$$V(\boldsymbol{\gamma}, [a,b])$$

can be used as a definition for the length of Γ *in the sense that if* $\Gamma = \boldsymbol{\eta}([c,d])$ *where* $\boldsymbol{\eta}$ *is a bounded variation continuous function which is one to one on* (c,d) *with* $\boldsymbol{\eta}([c,d]) = \Gamma$,

$$V(\boldsymbol{\gamma}, [a,b]) = V(\boldsymbol{\eta}, [c,d]).$$

This common value can be denoted by $V(\Gamma)$ *and is called the length of* Γ.

The length is not dependent on parameterization. Simple curves which have such parameterizations are called rectifiable.

11.1.2 Orientation

There is another notion called orientation. For simple rectifiable curves, you can think of it as a direction of motion over the curve but what does this really mean for a wriggly curve? A precise description is needed.

Definition 11.1.4 *Let* $\boldsymbol{\eta}, \boldsymbol{\gamma}$ *be continuous one to one parameterizations for a simple rectifiable curve. If* $\boldsymbol{\eta}^{-1} \circ \boldsymbol{\gamma}$ *is increasing, then* $\boldsymbol{\gamma}$ *and* $\boldsymbol{\eta}$ *are said to be equivalent parameterizations and this is written as* $\boldsymbol{\gamma} \sim \boldsymbol{\eta}$. *It is also said that the two parameterizations give the same orientation for the curve when* $\boldsymbol{\gamma} \sim \boldsymbol{\eta}$.

When the parameterizations are equivalent, they preserve the direction of motion along the curve and this also shows there are exactly two orientations of the curve since either $\boldsymbol{\eta}^{-1} \circ \boldsymbol{\gamma}$ is increasing or it is decreasing thanks to Lemma 11.1.2. In simple language, the message is that there are exactly two directions of motion along a simple curve.

Lemma 11.1.5 *The following hold for* $\sim$.

$$\boldsymbol{\gamma} \sim \boldsymbol{\gamma}, \tag{11.1}$$

$$\text{If } \boldsymbol{\gamma} \sim \boldsymbol{\eta} \text{ then } \boldsymbol{\eta} \sim \boldsymbol{\gamma}, \tag{11.2}$$

$$\text{If } \boldsymbol{\gamma} \sim \boldsymbol{\eta} \text{ and } \boldsymbol{\eta} \sim \boldsymbol{\theta}, \text{ then } \boldsymbol{\gamma} \sim \boldsymbol{\theta}. \tag{11.3}$$

Proof: Formula 11.1 is obvious because $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\gamma}(t) = t$ so it is clearly an increasing function. If $\boldsymbol{\gamma} \sim \boldsymbol{\eta}$ then $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}$ is increasing. Now $\boldsymbol{\eta}^{-1} \circ \boldsymbol{\gamma}$ must also be increasing because it is the inverse of $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}$. This verifies 11.2. To see 11.3, $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\theta} = (\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}) \circ (\boldsymbol{\eta}^{-1} \circ \boldsymbol{\theta})$ and so since both of these functions are increasing, it follows $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\theta}$ is also increasing. This proves the lemma.

Definition 11.1.6 *Let* Γ *be a simple rectifiable curve and let* $\boldsymbol{\gamma}$ *be a parameterization for* Γ. *Denoting by* $[\boldsymbol{\gamma}]$ *the equivalence class of parameterizations determined by the above equivalence relation, the following pair will be called an oriented curve.*

$$(\Gamma, [\boldsymbol{\gamma}])$$

In simple language, an oriented curve is one which has a direction of motion specified.

Actually, people usually just write Γ and there is understood a direction of motion or orientation on Γ. How can you identify which orientation is being considered?

Proposition 11.1.7 *Let* $(\Gamma, [\boldsymbol{\gamma}])$ *be an oriented simple curve and let* $\mathbf{p}, \mathbf{q}$ *be any two distinct points of* Γ. *Then* $[\boldsymbol{\gamma}]$ *is determined by the order of* $\boldsymbol{\gamma}^{-1}(\mathbf{p})$ *and* $\boldsymbol{\gamma}^{-1}(\mathbf{q})$. *This means that* $\boldsymbol{\eta} \in [\boldsymbol{\gamma}]$ *if and only if* $\boldsymbol{\eta}^{-1}(\mathbf{p})$ *and* $\boldsymbol{\eta}^{-1}(\mathbf{q})$ *occur in the same order as* $\boldsymbol{\gamma}^{-1}(\mathbf{p})$ *and* $\boldsymbol{\gamma}^{-1}(\mathbf{q})$.

Proof: Suppose $\boldsymbol{\gamma}^{-1}(\mathbf{p}) < \boldsymbol{\gamma}^{-1}(\mathbf{q})$ and let $\boldsymbol{\eta} \in [\boldsymbol{\gamma}]$. Is it true that $\boldsymbol{\eta}^{-1}(\mathbf{p}) < \boldsymbol{\eta}^{-1}(\mathbf{q})$? Of course it is because $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}$ is increasing. Therefore, if $\boldsymbol{\eta}^{-1}(\mathbf{p}) > \boldsymbol{\eta}^{-1}(\mathbf{q})$ it would follow

$$\boldsymbol{\gamma}^{-1}(\mathbf{p}) = \boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}\left(\boldsymbol{\eta}^{-1}(\mathbf{p})\right) > \boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}\left(\boldsymbol{\eta}^{-1}(\mathbf{q})\right) = \boldsymbol{\gamma}^{-1}(\mathbf{q})$$

which is a contradiction. Thus if $\boldsymbol{\gamma}^{-1}(\mathbf{p}) < \boldsymbol{\gamma}^{-1}(\mathbf{q})$ for one $\boldsymbol{\gamma} \in [\boldsymbol{\gamma}]$, then this is true for all $\boldsymbol{\eta} \in [\boldsymbol{\gamma}]$.

Now suppose $\boldsymbol{\eta}$ is a parameterization for Γ defined on $[c,d]$ which has the property that

$$\boldsymbol{\eta}^{-1}(\mathbf{p}) < \boldsymbol{\eta}^{-1}(\mathbf{q})$$

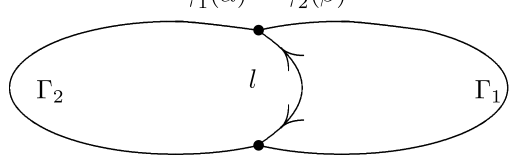
Does it follow $\boldsymbol{\eta} \in [\boldsymbol{\gamma}]$? Is $\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}$ increasing? By Lemma 11.1.2 it is either increasing or decreasing. Thus it suffices to test it on two points of $[c,d]$. Pick the two points $\boldsymbol{\eta}^{-1}(\mathbf{p}), \boldsymbol{\eta}^{-1}(\mathbf{q})$. Is

$$\boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}\left(\boldsymbol{\eta}^{-1}(\mathbf{p})\right) < \boldsymbol{\gamma}^{-1} \circ \boldsymbol{\eta}\left(\boldsymbol{\eta}^{-1}(\mathbf{q})\right)?$$

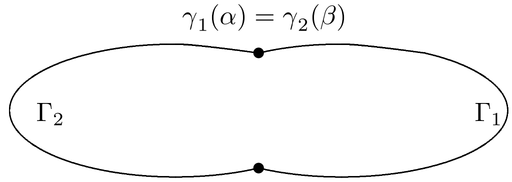
Yes because these reduce to $\boldsymbol{\gamma}^{-1}(\mathbf{p})$ on the left and $\boldsymbol{\gamma}^{-1}(\mathbf{q})$ on the right. It is given that $\boldsymbol{\gamma}^{-1}(\mathbf{p}) < \boldsymbol{\gamma}^{-1}(\mathbf{q})$. This proves the lemma.

This shows that the direction of motion on the curve is determined by any two points and the determination of which is encountered first by any parameterization in the equivalence class of parameterizations which determines the orientation. Sometimes people indicate this direction of motion by drawing an arrow.

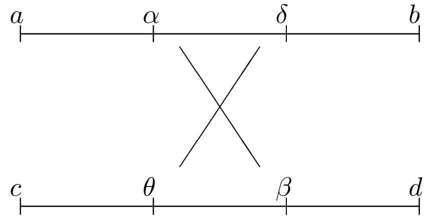
Now here is an interesting observation relative to two simple closed rectifiable curves. The situation is illustrated by the following picture.



Proposition 11.1.8 *Let Γ_1 and Γ_2 be two simple closed rectifiable oriented curves and let their intersection be l. Suppose also that l is itself a simple curve. Also suppose the orientation of l when considered a part of Γ_1 is opposite its orientation when considered a part of Γ_2. Then if the open segment (l except for its endpoints) of l is removed, the result is a simple closed rectifiable curve Γ. This curve has a parameterization $\boldsymbol{\gamma}$ with the property that on $\boldsymbol{\gamma}_j^{-1}(\Gamma \cap \Gamma_j)$, $\boldsymbol{\gamma}^{-1}\boldsymbol{\gamma}_j$ is increasing. In other words, Γ has an orientation consistent with that of Γ_1 and Γ_2. Furthermore, if Γ has such a consistent orientation, then the orientations of l as part of the two simple closed curves, Γ_1 and Γ_2 are opposite.*



Proof: Let $\Gamma_1 = \boldsymbol{\gamma}_1([a,b]), \boldsymbol{\gamma}_1(a) = \boldsymbol{\gamma}_1(b)$, and $\Gamma_2 = \boldsymbol{\gamma}_2([c,d]), \boldsymbol{\gamma}_2(c) = \boldsymbol{\gamma}_2(d)$, with $l = \boldsymbol{\gamma}_1([\alpha,\delta]) = \boldsymbol{\gamma}_2([\theta,\beta])$. (Recall continuous images of connected sets are connected and the connected sets on the real line are intervals.) By the assumption the two orientations are opposite, someting can be said about the relationship of $\alpha, \delta, \theta, \beta$. Suppose without loss of generality that $\alpha < \delta$. Then because of this assumption it follows $\theta < \beta$. The following diagram might be useful to summarize what was just said.



Note the first of the interval $[\beta, d]$ matches the last of the interval $[a, \alpha]$ and the first of $[\delta, \beta]$ matches the last of $[c, \theta]$, all this in terms of where these points are sent. If the orientations for l were not opposite, such a thing would not happen.

Now I need to describe the parameterization of $\Gamma \equiv \Gamma_1 \cup \Gamma_2$. To verify it is a simple closed curve, I must produce an interval and a mapping from this interval to Γ which satisfies the conditions needed for $\boldsymbol{\gamma}$ to be a simple closed rectifiable curve. The following is the definition as well as a description of which part of Γ_j is being obtained. Then $\boldsymbol{\gamma}(t)$ is given by

$$\boldsymbol{\gamma}(t) \equiv$$

$$\begin{cases} \boldsymbol{\gamma}_1(t), t \in [a, \alpha], \boldsymbol{\gamma}_1(a) \to \boldsymbol{\gamma}_1(\alpha) = \boldsymbol{\gamma}_2(\beta) \\ \boldsymbol{\gamma}_2(t + \beta - \alpha), t \in [\alpha, \alpha + d - \beta], \boldsymbol{\gamma}_2(\beta) \to \boldsymbol{\gamma}_2(d) = \boldsymbol{\gamma}_2(c) \\ \boldsymbol{\gamma}_2(t + c - \alpha - d + \beta), t \in [\alpha + d - \beta, \alpha + d - \beta + \theta - c], \\ \boldsymbol{\gamma}_2(c) = \boldsymbol{\gamma}_2(d) \to \boldsymbol{\gamma}_2(\theta) = \boldsymbol{\gamma}_1(\delta) \\ \boldsymbol{\gamma}_1(t - \alpha - d + \beta - \theta + c + \delta), t \in [\alpha + d - \beta + \theta - c, \alpha + d - \beta + \theta - c + b - \delta], \\ \boldsymbol{\gamma}_1(\delta) \to \boldsymbol{\gamma}_1(b) = \boldsymbol{\gamma}_1(a) \end{cases}$$

The construction shows $\boldsymbol{\gamma}$ is one to one on

$$(a, \alpha + d - \beta + \theta - c + b - \delta)$$

and if t is in this open interval, then

$$\boldsymbol{\gamma}(t) \neq \boldsymbol{\gamma}(a) = \boldsymbol{\gamma}_1(a)$$

and

$$\boldsymbol{\gamma}(t) \neq \boldsymbol{\gamma}(\alpha + d - \beta + \theta - c + b - \delta) = \boldsymbol{\gamma}_1(b).$$

Also

$$\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}_1(a) = \boldsymbol{\gamma}(\alpha + d - \beta + \theta - c + b - \delta) = \boldsymbol{\gamma}_1(b)$$

so it is a simple closed curve. The claim about preserving the orientation is also obvious from the formula. Note that t is never subtracted.

It only remains to prove the last claim. Suppose then that it is not so and l has the same orientation as part of each Γ_j. Then from a repeat of the above argument, you could change the orientation of l relative to Γ_2 and obtain an orientation of Γ which is consistent with that of Γ_1 and Γ_2. Call a parameterization which has this new orientation $\boldsymbol{\gamma}_n$ while $\boldsymbol{\gamma}$ is the one which is assumed to exist. This new orientation of l changes the orientation of Γ_2 because there are two points in l. Therefore on $\boldsymbol{\gamma}_2^{-1}(\Gamma \cap \Gamma_2)$, $\boldsymbol{\gamma}_n^{-1}\boldsymbol{\gamma}_2$ is decreasing while $\boldsymbol{\gamma}^{-1}\boldsymbol{\gamma}_2$ is assumed to be increasing. Hence $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}_n$ are not equivalent. However, the above construction would leave the orientation of both $\boldsymbol{\gamma}_1([a, \alpha])$ and $\boldsymbol{\gamma}_1([\delta, b])$ unchanged and at least one of these must have at least two points. Thus the orientation of Γ must be the same for $\boldsymbol{\gamma}_n$ as for $\boldsymbol{\gamma}$. That is, $\boldsymbol{\gamma} \sim \boldsymbol{\gamma}_n$. This is a contradiction. This proves the proposition.

There is a slightly different aspect of the above proposition which is interesting. It involves using the shared segment to orient the simple closed curve Γ.

Corollary 11.1.9 *Let the intersection of simple closed rectifiable curves, Γ_1 and Γ_2 consist of the simple curve l. Then place opposite orientations on l, and use these two different orientations to specify orientations of Γ_1 and Γ_2. Then letting Γ denote the simple closed curve which is obtained from deleting the open segment of l, there exists an orientation for Γ which is consistent with the orientations of Γ_1 and Γ_2 obtained from the given specification of opposite orientations on l.*

11.2 The Line Integral

Now I will return to considering the more general notion of bounded variation parameterizations without worrying about whether $\boldsymbol{\gamma}$ is one to one on the open interval. The line integral and its properties are presented next.

Definition 11.2.1 *Let $\boldsymbol{\gamma}:[a,b]\rightarrow\mathbb{R}^n$ be of bounded variation and let $\mathbf{f}:\boldsymbol{\gamma}^*\rightarrow\mathbb{R}^n$. Letting $\mathcal{P}\equiv\{t_0,\cdots,t_n\}$ where $a=t_0<t_1<\cdots<t_n=b$, define*

$$||\mathcal{P}||\equiv\max\{|t_j-t_{j-1}|:j=1,\cdots,n\}$$

and the Riemann Stieltjes sum by

$$S(\mathcal{P})\equiv\sum_{j=1}^{n}\mathbf{f}(\boldsymbol{\gamma}(\tau_j))\cdot(\boldsymbol{\gamma}(t_j)-\boldsymbol{\gamma}(t_{j-1}))$$

where $\tau_j\in[t_{j-1},t_j]$. (Note this notation is a little sloppy because it does not identify the specific point, τ_j used. It is understood that this point is arbitrary.) Define $\int_{\boldsymbol{\gamma}}\mathbf{f}\cdot d\boldsymbol{\gamma}$ as the unique number which satisfies the following condition. For all $\varepsilon>0$ there exists a $\delta>0$ such that if $||\mathcal{P}||\leq\delta$, then

$$\left|\int_{\boldsymbol{\gamma}}\mathbf{f}\cdot d\boldsymbol{\gamma}-S(\mathcal{P})\right|<\varepsilon.$$

Sometimes this is written as

$$\int_{\boldsymbol{\gamma}}\mathbf{f}\cdot d\boldsymbol{\gamma}\equiv\lim_{||\mathcal{P}||\rightarrow 0}S(\mathcal{P}).$$

Then $\boldsymbol{\gamma}^*$ is a set of points in $\mathbb{R}^n$ and as t moves from a to b, $\boldsymbol{\gamma}(t)$ moves from $\boldsymbol{\gamma}(a)$ to $\boldsymbol{\gamma}(b)$. Thus $\boldsymbol{\gamma}^*$ has a first point and a last point. (In the case of a closed curve these are the same point.) If $\phi:[c,d]\rightarrow[a,b]$ is a continuous nondecreasing function, then $\boldsymbol{\gamma}\circ\phi:[c,d]\rightarrow\mathbb{R}^n$ is also of bounded variation and yields the same set of points in $\mathbb{R}^n$ with the same first and last points.

Theorem 11.2.2 *Let ϕ and $\boldsymbol{\gamma}$ be as just described. Then assuming that*

$$\int_{\boldsymbol{\gamma}}\mathbf{f}\cdot d\boldsymbol{\gamma}$$

exists, so does

$$\int_{\boldsymbol{\gamma}\circ\phi}\mathbf{f}\cdot d(\boldsymbol{\gamma}\circ\phi)$$

and

$$\int_{\boldsymbol{\gamma}}\mathbf{f}\cdot d\boldsymbol{\gamma}=\int_{\boldsymbol{\gamma}\circ\phi}\mathbf{f}\cdot d(\boldsymbol{\gamma}\circ\phi). \tag{11.4}$$

Proof: There exists $\delta > 0$ such that if $\mathcal{P}$ is a partition of $[a,b]$ such that $||\mathcal{P}|| < \delta$, then

$$\left|\int_{\boldsymbol{\gamma}} \mathbf{f}\cdot d\boldsymbol{\gamma} - S(\mathcal{P})\right| < \varepsilon.$$

By continuity of ϕ, there exists $\sigma > 0$ such that if $\mathcal{Q}$ is a partition of $[c,d]$ with $||\mathcal{Q}|| < \sigma, \mathcal{Q} = \{s_0, \cdots, s_n\}$, then $|\phi(s_j) - \phi(s_{j-1})| < \delta$. Thus letting $\mathcal{P}$ denote the points in $[a,b]$ given by $\phi(s_j)$ for $s_j \in \mathcal{Q}$, it follows that $||\mathcal{P}|| < \delta$ and so

$$\left|\int_{\boldsymbol{\gamma}} \mathbf{f}\cdot d\boldsymbol{\gamma} - \sum_{j=1}^{n} \mathbf{f}(\boldsymbol{\gamma}(\phi(\tau_j))) \cdot (\boldsymbol{\gamma}(\phi(s_j)) - \boldsymbol{\gamma}(\phi(s_{j-1})))\right| < \varepsilon$$

where $\tau_j \in [s_{j-1}, s_j]$. Therefore, from the definition 11.4 holds and

$$\int_{\boldsymbol{\gamma}\circ\phi} \mathbf{f}\cdot d(\boldsymbol{\gamma}\circ\phi)$$

exists. This proves the theorem.

This theorem shows that $\int_{\boldsymbol{\gamma}} \mathbf{f}\cdot d\boldsymbol{\gamma}$ is independent of the particular parameterization $\boldsymbol{\gamma}$ used in its computation to the extent that if ϕ is any nondecreasing continuous function from another interval, $[c,d]$, mapping to $[a,b]$, then the same value is obtained by replacing $\boldsymbol{\gamma}$ with $\boldsymbol{\gamma}\circ\phi$. In other words, this line integral depends only on $\boldsymbol{\gamma}^*$ and the order in which $\boldsymbol{\gamma}(t)$ encounters the points of $\boldsymbol{\gamma}^*$ as t moves from one end to the other of the interval. For the case of an oriented rectifiable curve Γ this shows the line integral is dependent only on the set of points and the orientation of Γ.

The fundamental result in this subject is the following theorem.

Theorem 11.2.3 *Let $\mathbf{f} : \boldsymbol{\gamma}^* \to \mathbb{R}^n$ be continuous and let $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ be continuous and of bounded variation. Then $\int_{\boldsymbol{\gamma}} \mathbf{f}\cdot d\boldsymbol{\gamma}$ exists. Also letting $\delta_m > 0$ be such that $|t-s| < \delta_m$ implies $|\mathbf{f}(\boldsymbol{\gamma}(t)) - \mathbf{f}(\boldsymbol{\gamma}(s))| < \frac{1}{m}$,*

$$\left|\int_{\boldsymbol{\gamma}} \mathbf{f} d\boldsymbol{\gamma} - S(\mathcal{P})\right| \leq \frac{2V(\boldsymbol{\gamma}, [a,b])}{m}$$

whenever $||\mathcal{P}|| < \delta_m$.

Proof: The function, $\mathbf{f}\circ\boldsymbol{\gamma}$, is uniformly continuous because it is defined on a compact set. Therefore, there exists a decreasing sequence of positive numbers, $\{\delta_m\}$ such that if $|s-t| < \delta_m$, then

$$|\mathbf{f}(\boldsymbol{\gamma}(t)) - \mathbf{f}(\boldsymbol{\gamma}(s))| < \frac{1}{m}.$$

Let

$$F_m \equiv \overline{\{S(\mathcal{P}) : ||\mathcal{P}|| < \delta_m\}}.$$

Thus F_m is a closed set. (The symbol, $S(\mathcal{P})$ in the above definition, means to include all sums corresponding to $\mathcal{P}$ for any choice of τ_j.) It is shown that

$$\operatorname{diam}(F_m) \leq \frac{2V(\boldsymbol{\gamma}, [a,b])}{m} \tag{11.5}$$

and then it will follow there exists a unique point, $I \in \cap_{m=1}^{\infty} F_m$. This is because $\mathbb{R}$ is complete. It will then follow $I = \int_{\boldsymbol{\gamma}} \mathbf{f}(t)\, d\boldsymbol{\gamma}(t)$. To verify 11.5, it suffices to verify that whenever $\mathcal{P}$ and $\mathcal{Q}$ are partitions satisfying $||\mathcal{P}|| < \delta_m$ and $||\mathcal{Q}|| < \delta_m$,

$$|S(\mathcal{P}) - S(\mathcal{Q})| \leq \frac{2}{m} V(\boldsymbol{\gamma}, [a,b]). \tag{11.6}$$

Suppose $||\mathcal{P}|| < \delta_m$ and $\mathcal{Q} \supseteq \mathcal{P}$. Then also $||\mathcal{Q}|| < \delta_m$. To begin with, suppose that $\mathcal{P} \equiv \{t_0, \cdots, t_p, \cdots, t_n\}$ and $\mathcal{Q} \equiv \{t_0, \cdots, t_{p-1}, t^*, t_p, \cdots, t_n\}$. Thus $\mathcal{Q}$ contains only one more point than $\mathcal{P}$. Letting $S(\mathcal{Q})$ and $S(\mathcal{P})$ be Riemann Stieltjes sums,

$$S(\mathcal{Q}) \equiv \sum_{j=1}^{p-1} \mathbf{f}(\boldsymbol{\gamma}(\sigma_j)) \cdot (\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})) + \mathbf{f}(\boldsymbol{\gamma}(\sigma_*))(\boldsymbol{\gamma}(t^*) - \boldsymbol{\gamma}(t_{p-1}))$$

$$+\mathbf{f}(\boldsymbol{\gamma}(\sigma^*)) \cdot (\boldsymbol{\gamma}(t_p) - \boldsymbol{\gamma}(t^*)) + \sum_{j=p+1}^{n} \mathbf{f}(\boldsymbol{\gamma}(\sigma_j)) \cdot (\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})),$$

$$S(\mathcal{P}) \equiv \sum_{j=1}^{p-1} \mathbf{f}(\boldsymbol{\gamma}(\tau_j)) \cdot (\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})) +$$

$$\overbrace{\mathbf{f}(\boldsymbol{\gamma}(\tau_p)) \cdot (\boldsymbol{\gamma}(t^*) - \boldsymbol{\gamma}(t_{p-1})) + \mathbf{f}(\boldsymbol{\gamma}(\tau_p)) \cdot (\boldsymbol{\gamma}(t_p) - \boldsymbol{\gamma}(t^*))}^{=\mathbf{f}(\boldsymbol{\gamma}(\tau_p)) \cdot (\boldsymbol{\gamma}(t_p) - \boldsymbol{\gamma}(t_{p-1}))}$$

$$+ \sum_{j=p+1}^{n} \mathbf{f}(\boldsymbol{\gamma}(\tau_j)) \cdot (\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})).$$

Therefore,

$$|S(\mathcal{P}) - S(\mathcal{Q})| \leq \sum_{j=1}^{p-1} \frac{1}{m} |\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})| + \frac{1}{m} |\boldsymbol{\gamma}(t^*) - \boldsymbol{\gamma}(t_{p-1})| +$$

$$\frac{1}{m} |\boldsymbol{\gamma}(t_p) - \boldsymbol{\gamma}(t^*)| + \sum_{j=p+1}^{n} \frac{1}{m} |\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})| \leq \frac{1}{m} V(\boldsymbol{\gamma}, [a, b]). \tag{11.7}$$

Clearly the extreme inequalities would be valid in 11.7 if $\mathcal{Q}$ had more than one extra point. You simply do the above trick more than one time. Let $S(\mathcal{P})$ and $S(\mathcal{Q})$ be Riemann Stieltjes sums for which $||\mathcal{P}||$ and $||\mathcal{Q}||$ are less than δ_m and let $\mathcal{R} \equiv \mathcal{P} \cup \mathcal{Q}$. Then from what was just observed,

$$|S(\mathcal{P}) - S(\mathcal{Q})| \leq |S(\mathcal{P}) - S(\mathcal{R})| + |S(\mathcal{R}) - S(\mathcal{Q})| \leq \frac{2}{m} V(\boldsymbol{\gamma}, [a, b]).$$

and this shows 11.6 which proves 11.5. Therefore, there exists a unique number, $I \in \cap_{m=1}^{\infty} F_m$ which satisfies the definition of $\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}$. This proves the theorem.

Note this is a general sort of result. It is not assumed that $\boldsymbol{\gamma}$ is one to one anywhere in the proof. The following theorem follows easily from the above definitions and theorem. This theorem is used to establish estimates.

Theorem 11.2.4 *Let $\mathbf{f}$ be a continuous function defined on $\boldsymbol{\gamma}^*$, denoted as $C(\boldsymbol{\gamma}^*)$ where $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ is of bounded variation and continuous. Let*

$$M \geq \max \{|\mathbf{f} \circ \boldsymbol{\gamma}(t)| : t \in [a, b]\}. \tag{11.8}$$

Then

$$\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} \right| \leq MV(\boldsymbol{\gamma}, [a, b]). \tag{11.9}$$

Also if $\{\mathbf{f}_n\}$ is a sequence of functions of $C(\boldsymbol{\gamma}^*)$ which is converging uniformly to the function, $\mathbf{f}$ on $\boldsymbol{\gamma}^*$, then

$$\lim_{n \to \infty} \int_{\boldsymbol{\gamma}} \mathbf{f}_n \cdot d\boldsymbol{\gamma} = \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}. \tag{11.10}$$

In case $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$ *so the curve is a closed curve and for* f_k *the* k^{th} *component of* $\mathbf{f}$,

$$m_k \leq f_k(\mathbf{x}) \leq M_k$$

for all $\mathbf{x} \in \boldsymbol{\gamma}^*$, *it also follows*

$$\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} \right| \leq \frac{1}{2} \left(\sum_{k=1}^{n} (M_k - m_k)^2 \right)^{1/2} V(\boldsymbol{\gamma}, [a, b]) \tag{11.11}$$

Proof: Let 11.8 hold. From the proof of Theorem 11.2.3, when $||\mathcal{P}|| < \delta_m$,

$$\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} - S(\mathcal{P}) \right| \leq \frac{2}{m} V(\boldsymbol{\gamma}, [a, b])$$

and so

$$\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} \right| \leq |S(\mathcal{P})| + \frac{2}{m} V(\boldsymbol{\gamma}, [a, b])$$

Using the Cauchy Schwarz inequality and the above estimate in $S(\mathcal{P})$,

$$\begin{aligned} &\leq \sum_{j=1}^{n} M |\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})| + \frac{2}{m} V(\boldsymbol{\gamma}, [a, b]) \\ &\leq MV(\boldsymbol{\gamma}, [a, b]) + \frac{2}{m} V(\boldsymbol{\gamma}, [a, b]). \end{aligned}$$

This proves 11.9 since m is arbitrary. To verify 11.10 use the above inequality to write

$$\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} - \int_{\boldsymbol{\gamma}} \mathbf{f}_n \cdot d\boldsymbol{\gamma} \right| = \left| \int_{\boldsymbol{\gamma}} (\mathbf{f} - \mathbf{f}_n) \cdot d\boldsymbol{\gamma}(t) \right|$$

$$\leq \max \{ |\mathbf{f} \circ \boldsymbol{\gamma}(t) - \mathbf{f}_n \circ \boldsymbol{\gamma}(t)| : t \in [a, b] \} V(\boldsymbol{\gamma}, [a, b]).$$

Since the convergence is assumed to be uniform, this proves 11.10.

It only remains to verify 11.11. In this case $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$ and so for each vector $\mathbf{c}$

$$\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} = \int_{\boldsymbol{\gamma}} (\mathbf{f} - \mathbf{c}) \cdot d\boldsymbol{\gamma}$$

for any constant vector $\mathbf{c}$. Let

$$c_k = \frac{1}{2} (M_k + m_k)$$

Then for $t \in [a, b]$

$$\begin{aligned} |\mathbf{f}(\boldsymbol{\gamma}(t)) - \mathbf{c}|^2 &= \sum_{k=1}^{n} \left| f_k(\boldsymbol{\gamma}(t)) - \frac{1}{2}(M_k + m_k) \right|^2 \\ &\leq \sum_{k=1}^{n} \left(\frac{1}{2}(M_k - m_k) \right)^2 = \frac{1}{4} \sum_{k=1}^{n} (M_k - m_k)^2 \end{aligned}$$

Then with this choice of $\mathbf{c}$, it follows from 11.9 that

$$\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} \right| = \left| \int_{\boldsymbol{\gamma}} (\mathbf{f} - \mathbf{c}) \cdot d\boldsymbol{\gamma} \right|$$

$$\leq \frac{1}{2}\left(\sum_{k=1}^{n}(M_k - m_k)^2\right)^{1/2} V(\boldsymbol{\gamma},[a,b])$$

This proves the lemma.

It turns out to be much easier to evaluate such integrals in the case where $\boldsymbol{\gamma}$ is also $C^1([a,b])$. The following theorem about approximation will be very useful but first here is an easy lemma.

Lemma 11.2.5 *Let $\boldsymbol{\gamma}:[a,b]\to\mathbb{R}^n$ be in $C^1([a,b])$. Then $V(\boldsymbol{\gamma},[a,b])<\infty$ so $\boldsymbol{\gamma}$ is of bounded variation.*

Proof: This follows from the following

$$\begin{aligned}\sum_{j=1}^{n}|\boldsymbol{\gamma}(t_j)-\boldsymbol{\gamma}(t_{j-1})| &= \sum_{j=1}^{n}\left|\int_{t_{j-1}}^{t_j}\boldsymbol{\gamma}'(s)\,ds\right| \\ &\leq \sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}|\boldsymbol{\gamma}'(s)|\,ds \\ &\leq \sum_{j=1}^{n}\int_{t_{j-1}}^{t_j}||\boldsymbol{\gamma}'||_\infty\,ds \\ &= ||\boldsymbol{\gamma}'||_\infty (b-a).\end{aligned}$$

where

$$||\boldsymbol{\gamma}'||_\infty \equiv \max\{|\boldsymbol{\gamma}'(t)| : t\in[a,b]\}.$$

Therefore it follows $V(\boldsymbol{\gamma},[a,b])\leq ||\boldsymbol{\gamma}'||_\infty (b-a)$.

The following is a useful theorem for reducing bounded variation curves to ones which have a C^1 parameterization.

Theorem 11.2.6 *Let $\boldsymbol{\gamma}:[a,b]\to\mathbb{R}^n$ be continuous and of bounded variation. Let Ω be an open set containing $\boldsymbol{\gamma}^*$ and let $\mathbf{f}:\Omega\to\mathbb{R}^n$ be continuous, and let $\varepsilon>0$ be given. Then there exists $\boldsymbol{\eta}:[a,b]\to\mathbb{R}^n$ such that $\boldsymbol{\eta}(a)=\boldsymbol{\gamma}(a)$, $\boldsymbol{\gamma}(b)=\boldsymbol{\eta}(b)$, $\boldsymbol{\eta}\in C^1([a,b])$, and*

$$||\boldsymbol{\gamma}-\boldsymbol{\eta}||<\varepsilon, \tag{11.12}$$

where $||\boldsymbol{\gamma}-\boldsymbol{\eta}||\equiv\max\{|\boldsymbol{\gamma}(t)-\boldsymbol{\eta}(t)| : t\in[a,b]\}$. Also

$$\left|\int_{\boldsymbol{\gamma}}\mathbf{f}\cdot d\boldsymbol{\gamma}-\int_{\boldsymbol{\eta}}\mathbf{f}\cdot d\boldsymbol{\eta}\right|<\varepsilon, \tag{11.13}$$

$$V(\boldsymbol{\eta},[a,b])\leq V(\boldsymbol{\gamma},[a,b]), \tag{11.14}$$

Proof: Extend $\boldsymbol{\gamma}$ to be defined on all $\mathbb{R}$ according to the rule $\boldsymbol{\gamma}(t)=\boldsymbol{\gamma}(a)$ if $t<a$ and $\boldsymbol{\gamma}(t)=\boldsymbol{\gamma}(b)$ if $t>b$. Now define

$$\boldsymbol{\gamma}_h(t)\equiv\frac{1}{2h}\int_{-2h+t+\frac{2h}{(b-a)}(t-a)}^{t+\frac{2h}{(b-a)}(t-a)}\boldsymbol{\gamma}(s)\,ds.$$

where the integral is defined in the obvious way, that is componentwise. Since $\boldsymbol{\gamma}$ is continuous, this is certainly possible. Then

$$\boldsymbol{\gamma}_h(b)\equiv\frac{1}{2h}\int_b^{b+2h}\boldsymbol{\gamma}(s)\,ds=\frac{1}{2h}\int_b^{b+2h}\boldsymbol{\gamma}(b)\,ds=\boldsymbol{\gamma}(b),$$

$$\boldsymbol{\gamma}_h(a) \equiv \frac{1}{2h}\int_{a-2h}^{a} \boldsymbol{\gamma}(s)\,ds = \frac{1}{2h}\int_{a-2h}^{a} \boldsymbol{\gamma}(a)\,ds = \boldsymbol{\gamma}(a).$$

Also, because of continuity of $\boldsymbol{\gamma}$ and the fundamental theorem of calculus,

$$\boldsymbol{\gamma}_h'(t) = \frac{1}{2h}\left\{\boldsymbol{\gamma}\left(t + \frac{2h}{b-a}(t-a)\right)\left(1 + \frac{2h}{b-a}\right) - \boldsymbol{\gamma}\left(-2h + t + \frac{2h}{b-a}(t-a)\right)\left(1 + \frac{2h}{b-a}\right)\right\}$$

and so $\boldsymbol{\gamma}_h \in C^1([a,b])$. The following lemma is significant.

Lemma 11.2.7 $V(\boldsymbol{\gamma}_h, [a,b]) \leq V(\boldsymbol{\gamma}, [a,b])$.

Proof: Let $a = t_0 < t_1 < \cdots < t_n = b$. Then using the definition of $\boldsymbol{\gamma}_h$ and changing the variables to make all integrals over $[0, 2h]$,

$$\sum_{j=1}^{n} |\boldsymbol{\gamma}_h(t_j) - \boldsymbol{\gamma}_h(t_{j-1})| =$$

$$\sum_{j=1}^{n}\left|\frac{1}{2h}\int_0^{2h}\left[\boldsymbol{\gamma}\left(s - 2h + t_j + \frac{2h}{b-a}(t_j - a)\right) - \boldsymbol{\gamma}\left(s - 2h + t_{j-1} + \frac{2h}{b-a}(t_{j-1} - a)\right)\right]\right|$$

$$\leq \frac{1}{2h}\int_0^{2h}\sum_{j=1}^{n}\left|\boldsymbol{\gamma}\left(s - 2h + t_j + \frac{2h}{b-a}(t_j - a)\right) - \boldsymbol{\gamma}\left(s - 2h + t_{j-1} + \frac{2h}{b-a}(t_{j-1} - a)\right)\right| ds.$$

For a given $s \in [0, 2h]$, the points, $s - 2h + t_j + \frac{2h}{b-a}(t_j - a)$ for $j = 1, \cdots, n$ form an increasing list of points in the interval $[a - 2h, b + 2h]$ and so the integrand is bounded above by $V(\boldsymbol{\gamma}, [a - 2h, b + 2h]) = V(\boldsymbol{\gamma}, [a,b])$. It follows

$$\sum_{j=1}^{n} |\boldsymbol{\gamma}_h(t_j) - \boldsymbol{\gamma}_h(t_{j-1})| \leq V(\boldsymbol{\gamma}, [a,b])$$

which proves the lemma.

With this lemma the proof of the theorem can be completed without too much trouble. Let H be an open set containing $\boldsymbol{\gamma}^*$ such that $\overline{H}$ is a compact subset of Ω. Let $0 < \varepsilon < \operatorname{dist}(\boldsymbol{\gamma}^*, H^C)$. Then there exists δ_1 such that if $h < \delta_1$, then for all t,

$$\begin{aligned} |\boldsymbol{\gamma}(t) - \boldsymbol{\gamma}_h(t)| &\leq \frac{1}{2h}\int_{-2h+t+\frac{2h}{(b-a)}(t-a)}^{t+\frac{2h}{(b-a)}(t-a)} |\boldsymbol{\gamma}(s) - \boldsymbol{\gamma}(t)|\,ds \\ &< \frac{1}{2h}\int_{-2h+t+\frac{2h}{(b-a)}(t-a)}^{t+\frac{2h}{(b-a)}(t-a)} \varepsilon\,ds = \varepsilon \end{aligned} \tag{11.15}$$

due to the uniform continuity of $\boldsymbol{\gamma}$. This proves 11.12.

Using the estimate from Theorem 11.2.3, 11.5, the uniform continuity of $\mathbf{f}$ on H, and the above lemma, there exists δ such that if $||\mathcal{P}|| < \delta$, then

$$\left|\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}(t) - S(\mathcal{P})\right| < \frac{\varepsilon}{3}, \left|\int_{\boldsymbol{\gamma}_h} \mathbf{f} \cdot d\boldsymbol{\gamma}_h(t) - S_h(\mathcal{P})\right| < \frac{\varepsilon}{3}$$

for all $h < 1$. Here $S(\mathcal{P})$ is a Riemann Stieltjes sum of the form

$$\sum_{i=1}^{n} \mathbf{f}(\boldsymbol{\gamma}(\tau_i)) \cdot (\boldsymbol{\gamma}(t_i) - \boldsymbol{\gamma}(t_{i-1}))$$

and $S_h(\mathcal{P})$ is a similar Riemann Stieltjes sum taken with respect to $\boldsymbol{\gamma}_h$ instead of $\boldsymbol{\gamma}$. Because of 11.15 $\boldsymbol{\gamma}_h(t)$ has values in $H \subseteq \Omega$. Therefore, fix the partition, $\mathcal{P}$, and choose h small enough that in addition to this, the following inequality is valid.

$$|S(\mathcal{P}) - S_h(\mathcal{P})| < \frac{\varepsilon}{3}$$

This is possible because of 11.15 and the uniform continuity of $\mathbf{f}$ on $\overline{H}$. It follows

$$\left|\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}(t) - \int_{\boldsymbol{\gamma}_h} \mathbf{f} \cdot d\boldsymbol{\gamma}_h(t)\right| \leq$$

$$\left|\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}(t) - S(\mathcal{P})\right| + |S(\mathcal{P}) - S_h(\mathcal{P})|$$

$$+ \left|S_h(\mathcal{P}) - \int_{\boldsymbol{\gamma}_h} \mathbf{f} \cdot d\boldsymbol{\gamma}_h(t)\right| < \varepsilon.$$

Let $\boldsymbol{\eta} \equiv \boldsymbol{\gamma}_h$. Formula 11.14 follows from the lemma. This proves the theorem.

This is a very useful theorem because if $\boldsymbol{\gamma}$ is $C^1([a,b])$, it is easy to calculate $\int_{\boldsymbol{\gamma}} \mathbf{f} d\boldsymbol{\gamma}$ and the above theorem allows a reduction to the case where $\boldsymbol{\gamma}$ is C^1. The next theorem shows how easy it is to compute these integrals in the case where $\boldsymbol{\gamma}$ is C^1. First note that if $\mathbf{f}$ is continuous and $\boldsymbol{\gamma} \in C^1([a,b])$, then by Lemma 11.2.5 and the fundamental existence theorem, Theorem 11.2.3, $\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}$ exists.

Theorem 11.2.8 *If $\mathbf{f} : \boldsymbol{\gamma}^* \to X$ is continuous and $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ is in $C^1([a,b])$, then*

$$\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} = \int_a^b \mathbf{f}(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{\gamma}'(t)\, dt. \tag{11.16}$$

Proof: Let $\mathcal{P}$ be a partition of $[a,b]$, $\mathcal{P} = \{t_0, \cdots, t_n\}$ and $||\mathcal{P}||$ is small enough that whenever $|t - s| < ||\mathcal{P}||$,

$$|\mathbf{f}(\boldsymbol{\gamma}(t)) - \mathbf{f}(\boldsymbol{\gamma}(s))| < \varepsilon \tag{11.17}$$

and

$$\left|\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} - \sum_{j=1}^{n} \mathbf{f}(\boldsymbol{\gamma}(\tau_j)) \cdot (\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1}))\right| < \varepsilon.$$

Now

$$\sum_{j=1}^{n} \mathbf{f}(\boldsymbol{\gamma}(\tau_j)) \cdot (\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1}))$$

$$= \int_a^b \sum_{j=1}^{n} \mathbf{f}(\boldsymbol{\gamma}(\tau_j)) \cdot \mathcal{X}_{[t_{j-1}, t_j]}(s)\, \boldsymbol{\gamma}'(s)\, ds$$

where here

$$\mathcal{X}_{[p,q]}(s) \equiv \begin{cases} 1 \text{ if } s \in [p,q] \\ 0 \text{ if } s \notin [p,q] \end{cases}.$$

Also,

$$\int_a^b \mathbf{f}(\boldsymbol{\gamma}(s)) \cdot \boldsymbol{\gamma}'(s)\, ds = \int_a^b \sum_{j=1}^n \mathbf{f}(\boldsymbol{\gamma}(s)) \cdot \mathcal{X}_{[t_{j-1},t_j]}(s)\, \boldsymbol{\gamma}'(s)\, ds$$

and thanks to 11.17,

$$\left| \overbrace{\int_a^b \sum_{j=1}^n \mathbf{f}(\boldsymbol{\gamma}(\tau_j)) \cdot \mathcal{X}_{[t_{j-1},t_j]}(s)\, \boldsymbol{\gamma}'(s)\, ds}^{=\sum_{j=1}^n \mathbf{f}(\boldsymbol{\gamma}(\tau_j))\cdot(\boldsymbol{\gamma}(t_j)-\boldsymbol{\gamma}(t_{j-1}))} - \overbrace{\int_a^b \sum_{j=1}^n \mathbf{f}(\boldsymbol{\gamma}(s)) \cdot \mathcal{X}_{[t_{j-1},t_j]}(s)\, \boldsymbol{\gamma}'(s)\, ds}^{=\int_a^b \mathbf{f}(\boldsymbol{\gamma}(s))\cdot\boldsymbol{\gamma}'(s)ds} \right|$$

$$\begin{aligned} &\leq \sum_{j=1}^n \int_{t_{j-1}}^{t_j} |\mathbf{f}(\boldsymbol{\gamma}(\tau_j)) - \mathbf{f}(\boldsymbol{\gamma}(s))|\, |\boldsymbol{\gamma}'(s)|\, ds \\ &\leq ||\boldsymbol{\gamma}'||_\infty \sum_j \varepsilon (t_j - t_{j-1}) \\ &= \varepsilon\, ||\boldsymbol{\gamma}'||_\infty (b-a). \end{aligned}$$

It follows that

$$\begin{aligned} &\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} - \int_a^b \mathbf{f}(\boldsymbol{\gamma}(s)) \cdot \boldsymbol{\gamma}'(s)\, ds \right| \\ \leq &\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} - \sum_{j=1}^n \mathbf{f}(\boldsymbol{\gamma}(\tau_j)) \cdot (\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})) \right| \\ &+ \left| \sum_{j=1}^n \mathbf{f}(\boldsymbol{\gamma}(\tau_j)) \cdot (\boldsymbol{\gamma}(t_j) - \boldsymbol{\gamma}(t_{j-1})) - \int_a^b \mathbf{f}(\boldsymbol{\gamma}(s)) \cdot \boldsymbol{\gamma}'(s)\, ds \right| \\ \leq &\ \varepsilon\, ||\boldsymbol{\gamma}'||_\infty (b-a) + \varepsilon. \end{aligned}$$

Since ε is arbitrary, this verifies 11.16.

You can piece bounded variation curves together to get another bounded variation curve. You can also take the integral in the opposite direction along a given curve. There is also something called a potential.

Definition 11.2.9 *A function $\mathbf{f} : \Omega \to \mathbb{R}^n$ for Ω an open set in $\mathbb{R}^n$ has a potential if there exists a function, F, the potential, such that $\nabla F = \mathbf{f}$. Also if $\boldsymbol{\gamma}_k : [a_k, b_k] \to \mathbb{R}^n$ is continuous and of bounded variation, for $k = 1, \cdots, m$ and $\boldsymbol{\gamma}_k(b_k) = \boldsymbol{\gamma}_{k+1}(a_k)$, define*

$$\int_{\sum_{k=1}^m \boldsymbol{\gamma}_k} \mathbf{f} \cdot d\boldsymbol{\gamma}_k \equiv \sum_{k=1}^m \int_{\boldsymbol{\gamma}_k} \mathbf{f} \cdot d\boldsymbol{\gamma}_k. \tag{11.18}$$

In addition to this, for $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$*, define* $-\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ *by* $-\boldsymbol{\gamma}(t) \equiv \boldsymbol{\gamma}(b+a-t)$*. Thus* $\boldsymbol{\gamma}$ *simply traces out the points of* $\boldsymbol{\gamma}^*$ *in the opposite order.*

The following lemma is useful and follows quickly from Theorem 11.2.2.

Lemma 11.2.10 *In the above definition, there exists a continuous bounded variation function,* $\boldsymbol{\gamma}$ *defined on some closed interval,* $[c,d]$*, such that* $\boldsymbol{\gamma}([c,d]) = \cup_{k=1}^m \boldsymbol{\gamma}_k([a_k,b_k])$ *and* $\boldsymbol{\gamma}(c) = \boldsymbol{\gamma}_1(a_1)$ *while* $\boldsymbol{\gamma}(d) = \boldsymbol{\gamma}_m(b_m)$*. Furthermore,*

$$\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} = \sum_{k=1}^m \int_{\boldsymbol{\gamma}_k} \mathbf{f} \cdot d\boldsymbol{\gamma}_k.$$

If $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ *is of bounded variation and continuous, then*

$$\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} = -\int_{-\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}.$$

The following theorem shows that it is very easy to compute a line integral when the function has a potential.

Theorem 11.2.11 *Let* $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ *be continuous and of bounded variation. Also suppose* $\nabla F = \mathbf{f}$ *on* Ω*, an open set containing* $\boldsymbol{\gamma}^*$ *and* $\mathbf{f}$ *is continuous on* Ω*. Then*

$$\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} = F(\boldsymbol{\gamma}(b)) - F(\boldsymbol{\gamma}(a)).$$

Proof: By Theorem 11.2.6 there exists $\boldsymbol{\eta} \in C^1([a,b])$ such that $\boldsymbol{\gamma}(a) = \boldsymbol{\eta}(a)$, and $\boldsymbol{\gamma}(b) = \boldsymbol{\eta}(b)$ such that

$$\left| \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} - \int_{\boldsymbol{\eta}} \mathbf{f} \cdot d\boldsymbol{\eta} \right| < \varepsilon.$$

Then from Theorem 11.2.8, since $\boldsymbol{\eta}$ is in $C^1([a,b])$, it follows from the chain rule and the fundamental theorem of calculus that

$$\begin{aligned} \int_{\boldsymbol{\eta}} \mathbf{f} \cdot d\boldsymbol{\eta} &= \int_a^b \mathbf{f}(\boldsymbol{\eta}(t)) \boldsymbol{\eta}'(t)\, dt = \int_a^b \frac{d}{dt} F(\boldsymbol{\eta}(t))\, dt \\ &= F(\boldsymbol{\eta}(b)) - F(\boldsymbol{\eta}(a)) = F(\boldsymbol{\gamma}(b)) - F(\boldsymbol{\gamma}(a)). \end{aligned}$$

Therefore,

$$\left\| (F(\boldsymbol{\gamma}(b)) - F(\boldsymbol{\gamma}(a))) - \int_{\boldsymbol{\gamma}} \mathbf{f}(z)\, dz \right\| < \varepsilon$$

and since $\varepsilon > 0$ is arbitrary, this proves the theorem.

Corollary 11.2.12 *If* $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ *is continuous, has bounded variation, is a closed curve,* $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$*, and* $\boldsymbol{\gamma}^* \subseteq \Omega$ *where* Ω *is an open set on which* $\nabla F = \mathbf{f}$*, then*

$$\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} = 0.$$

Theorem 11.2.13 *Let* Ω *be a connected open set and let* $\mathbf{f} : \Omega \to \mathbb{R}^n$ *be continuous. Then* $\mathbf{f}$ *has a potential* F *if and only if*

$$\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}$$

is path independent for all $\boldsymbol{\gamma}$ *a bounded variation curve such that* $\boldsymbol{\gamma}^*$ *is contained in* Ω*. This means the above line integral depends only on* $\boldsymbol{\gamma}(a)$ *and* $\boldsymbol{\gamma}(b)$*.*

Proof: The first part was proved in Theorem 11.2.11. It remains to verify the existence of a potential in the situation of path independence.

Let $x_0 \in \Omega$ be fixed. Let S be the points $\mathbf{x}$ of Ω which have the property there is a bounded variation curve joining $\mathbf{x}_0$ to $\mathbf{x}$. Let $\boldsymbol{\gamma}_{\mathbf{x}_0\mathbf{x}}$ denote such a curve. Note first that S is nonempty. To see this, $B(\mathbf{x}_0, r) \subseteq \Omega$ for r small enough. Every $\mathbf{x} \in B(\mathbf{x}_0, r)$ is in S. Then S is open because if $\mathbf{x} \in S$, then $B(\mathbf{x}, r) \subseteq \Omega$ for small enough r and if $\mathbf{y} \in B(\mathbf{x}, r)$, you could go take $\boldsymbol{\gamma}_{\mathbf{x}_0\mathbf{x}}$ and from $\mathbf{x}$ follow the straight line segment joining $\mathbf{x}$ to $\mathbf{y}$. In addition to this, $\Omega \setminus S$ must also be open because if $\mathbf{x} \in \Omega \setminus S$, then choosing $B(\mathbf{x}, r) \subseteq \Omega$, no point of $B(\mathbf{x}, r)$ can be in S because then you could take the straight line segment from that point to $\mathbf{x}$ and conclude that $\mathbf{x} \in S$ after all. Therefore, since Ω is connected, it follows $\Omega \setminus S = \emptyset$. Thus for every $\mathbf{x} \in S$, there exists $\boldsymbol{\gamma}_{\mathbf{x}_0\mathbf{x}}$, a bounded variation curve from $\mathbf{x}_0$ to $\mathbf{x}$.

Define

$$F(\mathbf{x}) \equiv \int_{\boldsymbol{\gamma}_{\mathbf{x}_0\mathbf{x}}} \mathbf{f} \cdot d\boldsymbol{\gamma}_{\mathbf{x}_0\mathbf{x}}$$

F is well defined by assumption. Now let $l_{\mathbf{x}(\mathbf{x}+t\mathbf{e}_k)}$ denote the linear segment from $\mathbf{x}$ to $\mathbf{x} + t\mathbf{e}_k$. Thus to get to $\mathbf{x} + t\mathbf{e}_k$ you could first follow $\boldsymbol{\gamma}_{\mathbf{x}_0\mathbf{x}}$ to $\mathbf{x}$ and from there follow $l_{\mathbf{x}(\mathbf{x}+t\mathbf{e}_k)}$ to $\mathbf{x} + t\mathbf{e}_k$. Hence

$$\frac{F(\mathbf{x}+t\mathbf{e}_k) - F(\mathbf{x})}{t} = \frac{1}{t}\int_{l_{\mathbf{x}(\mathbf{x}+t\mathbf{e}_k)}} \mathbf{f} \cdot dl_{\mathbf{x}(\mathbf{x}+t\mathbf{e}_k)}$$

$$= \frac{1}{t}\int_0^t \mathbf{f}(\mathbf{x} + s\mathbf{e}_k) \cdot \mathbf{e}_k ds \to f_k(\mathbf{x})$$

by continuity of $\mathbf{f}$. Thus $\nabla F = \mathbf{f}$ and this proves the theorem.

Corollary 11.2.14 *Let Ω be a connected open set and $\mathbf{f} : \Omega \to \mathbb{R}^n$. Then $\mathbf{f}$ has a potential if and only if every closed, $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$, bounded variation curve has the property that*

$$\int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma} = 0$$

Proof: Using Lemma 11.2.10, this condition about closed curves is equivalent to the condition that the line integrals of the above theorem are path independent. This proves the corollary.

Such a vector valued function is called conservative.

11.3 Simple Closed Rectifiable Curves

There are examples of space filling continuous curves. However, bounded variation curves are not like this. In fact, one can even say the two dimensional Lebesgue measure of a bounded variation curve is 0.

Theorem 11.3.1 *Let $\boldsymbol{\gamma} : [a, b] \to \boldsymbol{\gamma}^* \subseteq \mathbb{R}^n$ where $n \geq 2$ is a continuous bounded variation curve. Then*

$$m_n(\boldsymbol{\gamma}^*) = 0$$

where m_n denotes n dimensional Lebesgue measure.

Proof: Let $\varepsilon > 0$ be given. Let $t_0 \equiv a$ and if $t_0, \cdots, t_k$ have been chosen, let t_{k+1} be the first number larger than t_k such that

$$|\boldsymbol{\gamma}(t_{k+1}) - \boldsymbol{\gamma}(t_k)| = \varepsilon.$$

If the set of t such that $|\boldsymbol{\gamma}(t)-\boldsymbol{\gamma}(t_k)|=\varepsilon$ is nonempty, then this set is clearly closed and so such a t_{k+1} exists until k is such that

$$\boldsymbol{\gamma}^* \subseteq \cup_{j=0}^{k} B(\boldsymbol{\gamma}(t_j),\varepsilon)$$

Let m be the last index of this process where t_{m+1} does not exist. How large is m? This can be estimated because

$$V(\boldsymbol{\gamma},[a,b]) \geq \sum_{k=0}^{m} |\boldsymbol{\gamma}(t_{k+1})-\boldsymbol{\gamma}(t_k)| = m\varepsilon$$

and so $m \leq V(\boldsymbol{\gamma},[a,b])/\varepsilon$. Since $\boldsymbol{\gamma}^* \subseteq \cup_{j=0}^{m} B(\boldsymbol{\gamma}(t_j),\varepsilon)$,

$$\begin{aligned} m_n(\boldsymbol{\gamma}^*) &\leq \sum_{j=0}^{m} m_n(B(\boldsymbol{\gamma}(t_j),\varepsilon)) \\ &\leq \frac{V(\boldsymbol{\gamma},[a,b])}{\varepsilon} c_n \varepsilon^n = c_n V(\boldsymbol{\gamma},[a,b])\varepsilon^{n-1} \end{aligned}$$

Since ε was arbitrary, this proves the theorem.

Since a ball has positive measure, this proves the following corollary.

Corollary 11.3.2 *Let $\boldsymbol{\gamma} : [a,b] \to \boldsymbol{\gamma}^* \subseteq \mathbb{R}^n$ where $n \geq 2$ is a continuous bounded variation curve. Then $\boldsymbol{\gamma}^*$ has empty interior.*

Lemma 11.3.3 *Let Γ be a simple closed curve. Then there exists a mapping $\boldsymbol{\theta} : C \to \Gamma$ where C is the unit circle*

$$\left\{(x,y) : x^2+y^2=1\right\},$$

such that $\boldsymbol{\theta}$ is one to one and continuous.

Proof: Since Γ is a simple closed curve, there is a parameterization $\boldsymbol{\gamma}$ and an interval $[a,b]$ such that $\boldsymbol{\gamma}$ is continuous and one to one on (a,b) and $\boldsymbol{\gamma}(a)=\boldsymbol{\gamma}(b)$. Also $\boldsymbol{\gamma}(t) \neq \boldsymbol{\gamma}(a)=\boldsymbol{\gamma}(b)$ if $t \neq a$ and if $t \neq b$. Define $\boldsymbol{\theta}^{-1} : \Gamma \to C$ by

$$\boldsymbol{\theta}^{-1}(\mathbf{x}) \equiv \left(\cos\left(\frac{2\pi}{b-a}\left(\boldsymbol{\gamma}^{-1}(\mathbf{x})-a\right)\right), \sin\left(\frac{2\pi}{b-a}\left(\boldsymbol{\gamma}^{-1}(\mathbf{x})-a\right)\right)\right)$$

Note that $\boldsymbol{\theta}^{-1}$ is onto C. The function is well defined because it sends the point $\boldsymbol{\gamma}(a)=\boldsymbol{\gamma}(b)$ to the same point, $(1,0)$. It is also one to one. To see this note $\boldsymbol{\gamma}^{-1}$ is one to one on $\Gamma \setminus \{\boldsymbol{\gamma}(a),\boldsymbol{\gamma}(b)\}$. What about the case where $\mathbf{x} \neq \boldsymbol{\gamma}(a)=\boldsymbol{\gamma}(b)$? Could $\boldsymbol{\theta}^{-1}(\mathbf{x}) = \boldsymbol{\theta}^{-1}(\boldsymbol{\gamma}(a))$? In this case, $\boldsymbol{\gamma}^{-1}(\mathbf{x})$ is in (a,b) while $\boldsymbol{\gamma}^{-1}(\boldsymbol{\gamma}(a))=a$ so

$$\boldsymbol{\theta}^{-1}(\mathbf{x}) \neq \boldsymbol{\theta}^{-1}(\boldsymbol{\gamma}(a)) = (1,0).$$

Thus $\boldsymbol{\theta}^{-1}$ is one to one on Γ.

Why is $\boldsymbol{\theta}^{-1}$ continuous? Suppose $\mathbf{x}_n \to \boldsymbol{\gamma}(a)=\boldsymbol{\gamma}(b)$ first. Why does $\boldsymbol{\theta}^{-1}(\mathbf{x}_n) \to (1,0) = \boldsymbol{\theta}^{-1}(\boldsymbol{\gamma}(a))$? Let $\{\mathbf{x}_n\}$ denote any subsequence of the given sequence. Then by compactness of $[a,b]$ there exists a further subsequence, still denoted by $\mathbf{x}_n$ such that

$$\boldsymbol{\gamma}^{-1}(\mathbf{x}_n) \to t \in [a,b]$$

Hence by continuity of $\boldsymbol{\gamma}$, $\mathbf{x}_n \to \boldsymbol{\gamma}(t)$ and so $\boldsymbol{\gamma}(t)$ must equal $\boldsymbol{\gamma}(a)=\boldsymbol{\gamma}(b)$. It follows from the assumption of what a simple curve is that $t \in \{a,b\}$. Hence $\boldsymbol{\theta}^{-1}(\mathbf{x}_n)$ converges to either

$$\left(\cos\left(\frac{2\pi}{b-a}(a-a)\right), \sin\left(\frac{2\pi}{b-a}(a-a)\right)\right)$$

or

$$\left(\cos\left(\frac{2\pi}{b-a}(b-a)\right), \sin\left(\frac{2\pi}{b-a}(b-a)\right)\right)$$

but these are the same point. This has shown that if $\mathbf{x}_n \to \boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$, there is a subsequence such that $\boldsymbol{\theta}^{-1}(\mathbf{x}_n) \to \boldsymbol{\theta}^{-1}(\boldsymbol{\gamma}(a))$. Thus $\boldsymbol{\theta}^{-1}$ is continuous at $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$. Next suppose $\mathbf{x}_n \to \mathbf{x} \neq \boldsymbol{\gamma}(a) \equiv \mathbf{p}$. Then there exists $B(\mathbf{p}, r)$ such that for all n large enough, $\mathbf{x}_n$ and $\mathbf{x}$ are contained in the compact set $\Gamma \setminus B(\mathbf{p}, r) \equiv K$. Then $\boldsymbol{\gamma}$ is continuous and one to one on the compact set $\boldsymbol{\gamma}^{-1}(K) \subseteq (a, b)$ and so by Theorem 5.1.3 $\boldsymbol{\gamma}^{-1}$ is continuous on K. In particular it is continuous at $\mathbf{x}$ so $\boldsymbol{\theta}^{-1}(\mathbf{x}_n) \to \boldsymbol{\theta}^{-1}(\mathbf{x})$. This proves the lemma.

11.3.1 The Jordan Curve Theorem

The following theorem includes the Jordan Curve theorem, a major result for simple closed curves in the plane. In this theorem and in what follows U_i will denote the inside of a simple closed curve and U_o will denote the outside.

Theorem 11.3.4 *Let C denote the unit circle, $\left\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\right\}$. Suppose $\boldsymbol{\gamma} : C \to \Gamma \subseteq \mathbb{R}^2$ is one to one onto and continuous. Then $\mathbb{R}^n \setminus \Gamma$ consists of two components, a bounded component (called the inside) U_i and an unbounded component (called the outside), U_o. Also the boundary of each of these two components of $\mathbb{R}^n \setminus \Gamma$ is Γ and Γ has empty interior.*

Proof: That $\mathbb{R}^n \setminus \Gamma$ consists of two components, U_o and U_i follows from the Jordan separation theorem. There is exactly one unbounded component because Γ is bounded and so U_i is defined as the bounded component. It remains to verify the assertion about Γ being the boundary. Let $\mathbf{x}$ be a limit point of U_i. Then it can't be in U_o because these are both open sets. Therefore, all the limit points of U_i are in $U_i \cup \Gamma$. Similarly all the limit points of U_o are in $U_o \cup \Gamma$. Thus $\partial U_i \subseteq \Gamma$ and $\partial U_o \subseteq \Gamma$.

I claim Γ has empty interior. This follows because by Theorem 5.1.3 on Page 84, $\boldsymbol{\gamma}$ and $\boldsymbol{\gamma}^{-1}$ must both be continuous since C is compact. Thus if B is an open ball contained in Γ, it follows from invariance of domain that $\boldsymbol{\gamma}^{-1}(B)$ is an open set in $\mathbb{R}^2$. But this needs to be contained in C which is a contradiction because C has empty interior obviously.

Now let $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{x} \notin U_o \cup U_i$. Then $\mathbf{x} \in \Gamma$ and must be a limit point of either U_o or U_i since if this were not so, Γ would be forced to have nonempty interior. Hence $\overline{U_1} \cup \overline{U_2} = \mathbb{R}^2$. Next I will show $\partial U_i = \partial U_o$. Suppose then that

$$\mathbf{p} \in \partial U_i \setminus \partial U_o$$

Then $\mathbf{p} \notin U_o$ because $\mathbf{p} \in \partial U_i \subseteq \Gamma$ which is disjoint from U_o. Thus $\mathbf{p}$ is not in $\overline{U_o}$ because it is given to not be in ∂U_o. Hence there is a ball centered at $\mathbf{p}$, $B(\mathbf{p}, r)$ which contains no points of $\overline{U_o}$. Thus $B(\mathbf{p}, r) \subseteq \overline{U_i}$ and so $\mathbf{p}$ is an interior point of $\overline{U_i}$ which implies $\mathbf{p}$ is actually in U_i, a contradiction to $\mathbf{p} \in \partial U_i$. Thus $\partial U_i \setminus \partial U_o = \emptyset$ and a similar argument shows $\partial U_o \setminus \partial U_i = \emptyset$. Thus $\partial U_i = \partial U_o$ and so if $\mathbf{x} \in \Gamma$, then it is in either ∂U_i or ∂U_o and these are equal. Thus

$$\partial U_i = \partial U_o \subseteq \Gamma \subseteq \partial U_i \cup \partial U_o = \partial U_i = \partial U_o.$$

This proves the theorem.

The following lemma will be of importance in what follows. There are of course more general versions of this lemma. I am only presenting what will be needed. To say two sets are homeomorphic is to say there is a one to one continuous onto mapping from one to the other which also has continuous inverse. Clearly the statement that two sets are homeomorphic defines an equivalence relation.

Lemma 11.3.5 *In the situation of Theorem 11.3.4, let* Γ *be a simple closed curve and let* $\boldsymbol{\gamma}^*$ *be a straight line segment such that the open segment,* $\boldsymbol{\gamma}^*$ *without its endpoints, is contained in* U_i *such that the intersection of* $\boldsymbol{\gamma}^*$ *with* Γ *equals* $\{\mathbf{p},\mathbf{q}\}$ *. Then this line segment divides* U_i *into two connected open sets which are the interiors of two simple closed curves.*

Proof: Suppose $\boldsymbol{\theta}$ is a one to one onto continuous mapping of C, the unit circle, to Γ. Say $\boldsymbol{\theta}(\mathbf{a}) = \mathbf{p}$ and $\boldsymbol{\theta}(\mathbf{b}) = \mathbf{q}$ where $\mathbf{a},\mathbf{b}$ are points of C. Then $\Gamma\cup\boldsymbol{\gamma}^*$ is homeomorphic to the $C\cup l^*$ where l^* is the straight line joining $\mathbf{a}$ and $\mathbf{b}$. This is easy to see because l^* is clearly homeomorphic to $\boldsymbol{\gamma}^*$ so all that is required is to extend $\boldsymbol{\theta}$ to all of $C\cup l^*$. By the Jordan separation theorem, $\mathbb{R}^2\setminus(\Gamma\cup\boldsymbol{\gamma}^*)$ is the union of three disjoint open connected sets, exactly one of which is unbounded. This is because this is obviously true for the circle and secant line l^*. Also, denoting one of the arcs of C joining $\mathbf{a}$ to $\mathbf{b}$ by C_1 and the other arc by C_2, it follows $\boldsymbol{\theta}(C_j)\cup\boldsymbol{\gamma}^*$ is a simple closed curve because it is homeomorphic to the half of the circle $C_j\cup l^*$ which is clearly homeomorphic to the unit circle. Let $\Gamma_j\equiv\boldsymbol{\theta}(C_j)$. The two connected open sets are the insides of $\Gamma_1\cup\boldsymbol{\gamma}^*$ and $\Gamma_2\cup\boldsymbol{\gamma}^*$ respectively and $\Gamma\cup\boldsymbol{\gamma}^* = (\Gamma_1\cup\boldsymbol{\gamma}^*)\cup(\Gamma_2\cup\boldsymbol{\gamma}^*)$.

Here is why. Denote them by U_{1i} and U_{2i} respectively. First note they can have no point in common for if $\mathbf{x}$ were such a point, then since both of these sets are open connected components, this would require them to coincide. Hence $\mathbb{R}^2\setminus(\Gamma\cup\boldsymbol{\gamma}^*)$ would only have two components, a bounded and an unbounded component contrary to what was shown above.

If $\mathbf{x}$ is a point of U_i which is not on $\boldsymbol{\gamma}^*$, why must it be in one of U_{1i} or U_{2i}? If it is in neither, then it is in $U_{1o}\cap U_{2o}$ the intersection of the two unbounded components. I claim this intersection equals U_o. To see this, note the unbounded component of $\mathbb{R}^2\setminus(\Gamma\cup\boldsymbol{\gamma}^*)$ equals the unbounded component of $\mathbb{R}^2\setminus\Gamma$ which equals U_o because $\boldsymbol{\gamma}^*$ is contained entirely in $U_i\cup\Gamma$. But the unbounded component of $\mathbb{R}^2\setminus(\Gamma\cup\boldsymbol{\gamma}^*)$ equals everything which is not in the union of the interior components and the boundary. Thus this unbounded component equals

$$
\begin{aligned}
&((\Gamma_1\cup\boldsymbol{\gamma}^*)\cup(\Gamma_2\cup\boldsymbol{\gamma}^*)\cup U_{1i}\cup U_{2i})^C \\
=\;& ((\Gamma_1\cup\boldsymbol{\gamma}^*)\cup U_{1i})^C\cap((\Gamma_2\cup\boldsymbol{\gamma}^*)\cup U_{2i})^C \\
=\;& U_{1o}\cap U_{2o}
\end{aligned}
$$

Now this is a contradiction because $\mathbf{x}\in U_i$ and so is not in U_o. This proves the lemma.

The following lemma has to do with decomposing the inside and boundary of a simple closed rectifiable curve into small pieces. The argument is like one given in Apostol [3]. In doing this I will refer to a region as the union of a connected open set with its boundary. Also, two regions will be said to be non overlapping if they either have empty intersection or the intersection is contained in the intersection of their boundaries.The height of a set A equals $\sup\{|y_1-y_2| : (x_1,y_1),(x_2,y_2)\in A\}$. The width of A will be defined similarly.

Lemma 11.3.6 *Let* Γ *be a simple closed rectifiable curve. Also let* $\delta>0$ *be given such that* 2δ *is smaller than both the height and width of* Γ*. Then there exist finitely many non overlapping regions* $\{R_k\}_{k=1}^n$ *consisting of simple closed curves along with their interiors whose union equals* $U_i\cup\Gamma$*. These regions consist of two kinds, those contained in* U_i *and those with nonempty intersection with* Γ*. These latter regions are called "border" regions. The boundary of a border region consists of straight line segments parallel to the coordinate axes of the form* $x=m\delta$ *or* $y=k\delta$ *for* m,k *integers along with arcs from* Γ*. The regions contained in* U_i *consist of rectangles. Thus all of these regions have boundaries which are rectifiable simple closed curves. Also all regions are contained in a square having sides of length no more than* 2δ*. There are at most*

$$4\left(\frac{V(\Gamma)}{\delta}+1\right)$$

border regions. The construction also yields an orientation for Γ and for all these regions and the orientations for any segment shared by two regions are opposite.

Proof: Let $\Gamma = \boldsymbol{\gamma}([a,b])$ where $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2)$. Let

$$y_1 \equiv \max\{\boldsymbol{\gamma}_2(t) : t \in [a,b]\}$$

and let

$$y_2 \equiv \min\{\boldsymbol{\gamma}_2(t) : t \in [a,b]\}.$$

Thus (x_1, y_1) is the "top" point of Γ while (x_2, y_2) is the "bottom" point of Γ. Consider the lines $y = y_1$ and $y = y_2$. By assumption $|y_1 - y_2| > 2\delta$. Consider the line l given by $y = m\delta$ where m is chosen to make $m\delta$ as close as possible to $(y_1 + y_2)/2$. Thus $y_1 > m\delta > y_2$. By Theorem 11.3.4 (x_j, y_j) $j = 1, 2$ are neither of them interior points of Γ and so by Theorem 11.3.4 again, there exist points $\mathbf{p}_j \in U_i$ such that $\mathbf{p}_1$ is above l and $\mathbf{p}_2$ is below l. (Simply pick $\mathbf{p}_j$ very close to (x_j, y_j) and yet in U_i and this will take place.) Therefore, the horizontal line l must have nonempty intersection with U_i because U_i is connected. If it had empty intersection it would be possible to separate U_i into two nonempty open sets, one containing $\mathbf{p}_1$ and the other containing $\mathbf{p}_2$.

Let $\mathbf{q}$ be a point of U_i which is also in l. Then there exists a maximal segment of the line l containing $\mathbf{q}$ which is contained in $U_i \cup \Gamma$. This segment, $\boldsymbol{\gamma}^*$ satisfies the conditions of Lemma 11.3.5 and so it divides U_i into disjoint open connected sets whose boundaries are simple rectifiable closed curves. Note the line segment has finite length. Letting Γ_j be the simple closed curve which contains $\mathbf{p}_j$, orient $\boldsymbol{\gamma}^*$ as part of Γ_2 such that motion is from right to left. As part of Γ_1 the motion along the curve is from left to right. By Proposition 11.1.7 this provides an orientation to each Γ_j. By Proposition 11.1.8 there exists an orientation for Γ which is consistent with these two orientations on the Γ_j.

Now do the same process to the two simple closed curves just obtained and continue till all regions have height less than 2δ. Each application of the process yields two new non overlapping regions of the desired sort in place of an earlier region of the desired sort except possibly the regions might have excessive height. The orientation of a new line segment in the construction is determined from the orientations of the simple closed curves obtained earlier. By Proposition 11.1.7 the orientations of the segments shared by two regions are opposite so the line integrals over these segments cancel. Eventually this process ends because all regions have "height" less than 2δ. The reason for this is that if it did not end, the curve Γ could not have finite total variation because there would exist an arbitrarily large number of non overlapping regions each of which have a pair of points which are farther apart than 2δ. This takes care of finding the subregions so far as height is concerned.

Now follow the same process just described on each of the non overlapping "short" regions just obtained using vertical rather than horizontal lines, letting the orientation of the vertical edges be determined from the orientation already obtained, but this time feature width instead of height and let the lines be vertical of the form $x = k\delta$ where k is an integer.

How many border regions are there? Denote by $V(\Gamma)$ the length of Γ. Now decompose Γ into N arcs of length δ with maybe one having length less than δ. Thus $N - 1 \leq \frac{V(\Gamma)}{\delta}$ and so

$$N \leq \frac{V(\Gamma)}{\delta} + 1$$

Each of these N arcs can't intersect any more than four of the boxes in the construction. Therefore, at most $4N$ boxes of the construction can intersect Γ. Thus there are no more than

$$4\left(\frac{V(\Gamma)}{\delta} + 1\right)$$

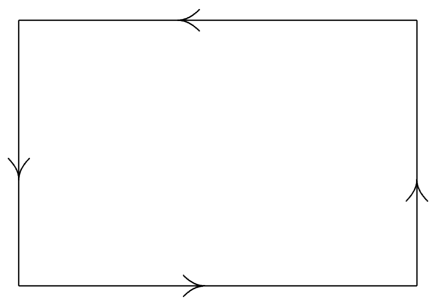
border regions. This proves the lemma.

11.3.2 Orientation And Green's Formula

How do you describe the orientation of a simple closed rectifiable curve analytically? The above process did it but I want another way to identify this which is more geometrically appealing. For simple examples, this is not too hard but it becomes less obvious when you consider the general case. The problem is the simple closed curve could be very wriggly.

The orientation of a rectifiable simple closed curve will be defined in terms of a very important formula known as Green's formula. First I will present Green's formula for a rectangle. In this lemma, it is very easy to understand the orientation of the bounding curve. The direction of motion is counter clockwise. As described in Proposition 11.1.7 it suffices to describe a direction of motion along the curve using any two points.

Lemma 11.3.7 *Let $R = [a, b] \times [c, d]$ be a rectangle and let P, Q be functions which are C^1 in some open set containing R. Orient the boundary of R as shown in the following picture. This is called the counter clockwise direction or the positive orientation*



Then letting $\boldsymbol{\gamma}$ denote the oriented boundary of R as shown,

$$\int_R (Q_x(x, y) - P_y(x, y))\, dm_2 = \int_{\boldsymbol{\gamma}} \mathbf{f} \cdot d\boldsymbol{\gamma}$$

where

$$\mathbf{f}(x, y) \equiv (P(x, y), Q(x, y)).$$

In this context the line integral is usually written using the notation

$$\int_{\partial R} P dx + Q dy.$$

Proof: This follows from direct computation. A parameterization for the bottom line of R is

$$\boldsymbol{\gamma}_B(t) = (a + t(b - a), c),\ t \in [0, 1]$$

A parameterization for the top line of R with the given orientation is

$$\boldsymbol{\gamma}_T(t) = (b + t(a - b), d),\ t \in [0, 1]$$

A parameterization for the line on the right side is

$$\boldsymbol{\gamma}_R(t) = (b, c + t(d - c)),\ t \in [0, 1]$$

and a parameterization for the line on the left side is

$$\boldsymbol{\gamma}_L(t) = (a, d + t(c - d)),\ t \in [0, 1]$$

Now it is time to do the computations using Theorem 11.2.8.

$$\begin{aligned}\int_{\gamma} \mathbf{f} \cdot d\boldsymbol{\gamma} &= \int_0^1 P(a+t(b-a), c)(b-a)\,dt \\ &\quad + \int_0^1 P(b+t(a-b), d)(a-b)\,dt\end{aligned}$$

$$+ \int_0^1 Q(b, c+t(d-c))(d-c)\,dt + \int_0^1 Q(a, d+t(c-d))(c-d)\,dt$$

Changing the variables and combining the integrals, this equals

$$= \int_a^b P(x,c)\,dx - \int_a^b P(x,d)\,dx + \int_c^d Q(b,y)\,dy - \int_c^d Q(a,y)\,dy$$

$$\begin{aligned} &= -\int_a^b \int_c^d P_y(x,y)\,dydx + \int_c^d \int_a^b Q_x(x,y)\,dxdy \\ &= \int_R (Q_x - P_y)\,dm_2 \end{aligned}$$

By Fubini's theorem, Theorem 9.2.3 on Page 204. (To use this theorem you can extend the functions to equal 0 off R.) This proves the lemma.

Note that if the rectangle were oriented in the opposite way, you would get

$$\int_{\gamma} \mathbf{f} \cdot d\boldsymbol{\gamma} = \int_R (P_y - Q_x)\,dm_2$$

With this lemma, it is possible to prove Green's theorem and also give an analytic criterion which will distinguish between different orientations of a simple closed rectifiable curve. First here is a discussion which amounts to a computation.

Let Γ be a rectifiable simple closed curve with inside U_i and outside U_o. Let $\{R_k\}_{k=1}^{n_\delta}$ denote the non overlapping regions of Lemma 11.3.6 all oriented as explained there and let Γ also be oriented as explained there. It could be shown that all the regions contained in U_i have positive orientation but this will not be fussed over here. What can be said with no fussing is that since the shared edges have opposite orientations, all these interior regions are either oriented positively or they are all oriented negatively.

Let $\mathcal{B}_\delta$ be the set of border regions and let $\mathcal{I}_\delta$ be the rectangles contained in U_i. Thus in taking the sum of the line integrals over the boundaries of the interior rectangles, the integrals over the "interior edges" cancel out and you are left with a line integral over the exterior edges of a polygon which is composed of the union of the squares in $\mathcal{I}_\delta$.

Now let $\mathbf{f}(x,y) = (P(x,y), Q(x,y))$ be a vector field which is C^1 on U_i, and suppose also that both P_y and Q_x are in $L^1(U_i)$ and that P, Q are continuous on $U_i \cup \Gamma$. (An easy way to get all this to happen is to let P, Q be restrictions to $U_i \cup \Gamma$ of functions which are C^1 on some open set containing $U_i \cup \Gamma$.) Note that

$$\cup_{\delta>0} \{R : R \in \mathcal{I}_\delta\} = U_i$$

and that for

$$I_\delta \equiv \cup \{R : R \in \mathcal{I}_\delta\},$$

the following pointwise convergence holds.

$$\lim_{\delta \to 0} \mathcal{X}_{I_\delta}(\mathbf{x}) = \mathcal{X}_{U_i}(\mathbf{x}).$$

By the dominated convergence theorem,

$$\begin{aligned}\lim_{\delta\to 0}\int_{I_\delta}(Q_x-P_y)\,dm_2 &= \int_{U_i}(Q_x-P_y)\,dm_2\\ \lim_{\delta\to 0}\int_{I_\delta}(P_y-Q_x)\,dm_2 &= \int_{U_i}(P_y-Q_x)\,dm_2\end{aligned}$$

Let ∂R denote the boundary of R for R one of these regions of Lemma 11.3.6 oriented as described. Let $w_\delta(R)^2$ denote

$$\begin{aligned}&(\max\{Q(\mathbf{x}):\mathbf{x}\in\partial R\}-\min\{Q(\mathbf{x}):\mathbf{x}\in\partial R\})^2\\ &+(\max\{P(\mathbf{x}):\mathbf{x}\in\partial R\}-\min\{P(\mathbf{x}):\mathbf{x}\in\partial R\})^2\end{aligned}$$

By uniform continuity of P,Q on the compact set $U_i\cup\Gamma$, if δ is small enough, $w_\delta(R)<\varepsilon$ for all $R\in\mathcal{B}_\delta$. Then for $R\in\mathcal{B}_\delta$, it follows from Theorem 11.2.4

$$\left|\int_{\partial R}\mathbf{f}\cdot d\boldsymbol{\gamma}\right|\leq\frac{1}{2}w_\delta(R)(V(\partial R))<\varepsilon(V(\partial R)) \tag{11.19}$$

whenever δ is small enough. Always let δ be this small.

Also since the line integrals cancel on shared edges

$$\sum_{R\in\mathcal{I}_\delta}\int_{\partial R}\mathbf{f}\cdot d\boldsymbol{\gamma}+\sum_{R\in\mathcal{B}_\delta}\int_{\partial R}\mathbf{f}\cdot d\boldsymbol{\gamma}=\int_\Gamma\mathbf{f}\cdot d\boldsymbol{\gamma} \tag{11.20}$$

Consider the second sum on the left. From 11.19

$$\left|\sum_{R\in\mathcal{B}_\delta}\int_{\partial R}\mathbf{f}\cdot d\boldsymbol{\gamma}\right|\leq\sum_{R\in\mathcal{B}_\delta}\left|\int_{\partial R}\mathbf{f}\cdot d\boldsymbol{\gamma}\right|\leq\varepsilon\sum_{R\in\mathcal{B}_\delta}(V(\partial R))$$

Denote by Γ_R the part of Γ which is contained in $R\in\mathcal{B}_\delta$ and $V(\Gamma_R)$ is its length. Then the above sum equals

$$\varepsilon\left(\sum_{R\in\mathcal{B}_\delta}V(\Gamma_R)+B_\delta\right)=\varepsilon(V(\Gamma)+B_\delta)$$

where B_δ is the sum of the lengths of the straight edges. This is easy to estimate. Recall from 11.3.6 there are no more than

$$4\left(\frac{V(\Gamma)}{\delta}+1\right)$$

of these border regions. Furthermore, the sum of the lengths of all four edges of one of these is no more than 8δ and so

$$B_\delta\leq 4\left(\frac{V(\Gamma)}{\delta}+1\right)8\delta=32V(\Gamma)+32\delta.$$

Thus the absolute value of the second sum on the right in 11.20 is dominated by

$$\varepsilon(33V(\Gamma)+32\delta)$$

Since ε was arbitrary, this formula implies with Green's theorem proved above for squares

$$\int_\Gamma\mathbf{f}\cdot d\boldsymbol{\gamma}=\lim_{\delta\to 0}\sum_{R\in\mathcal{I}_\delta}\int_{\partial R}\mathbf{f}\cdot d\boldsymbol{\gamma}+\lim_{\delta\to 0}\sum_{R\in\mathcal{B}_\delta}\int_{\partial R}\mathbf{f}\cdot d\boldsymbol{\gamma}$$

$$= \lim_{\delta \to 0} \sum_{R \in \mathcal{I}_\delta} \int_{\partial R} \mathbf{f} \cdot d\boldsymbol{\gamma} = \lim_{\delta \to 0} \int_{I_\delta} \pm (Q_x - P_y) \, dm_2 = \int_{U_i} \pm (Q_x - P_y) \, dm_2$$

where the $\pm$ adusts for whether the interior rectangles are all oriented positively or all oriented negatively. This has proved the general form of Green's theorem which is stated in the following theorem.

Theorem 11.3.8 *Let Γ be a rectifiable simple closed curve in $\mathbb{R}^2$ having inside U_i and outside U_o. Let P, Q be functions with the property that*

$$Q_x, P_y \in L^1(U_i)$$

and P, Q are C^1 on U_i. Assume also P, Q are continuous on $\Gamma \cup U_i$. Then there exists an orientation for Γ (Remember there are only two.) such that for

$$\mathbf{f}(x, y) = (P(x, y), Q(x, y)),$$

$$\int_\Gamma \mathbf{f} \cdot d\boldsymbol{\gamma} = \int_{U_i} (Q_x - P_y) \, dm_2.$$

Proof: In the construction of the regions, an orientation was imparted to Γ. The above computation shows

$$\int_\Gamma \mathbf{f} \cdot d\boldsymbol{\gamma} = \int_{U_i} \pm (Q_x - P_y) \, dm_2$$

If the area integral equals

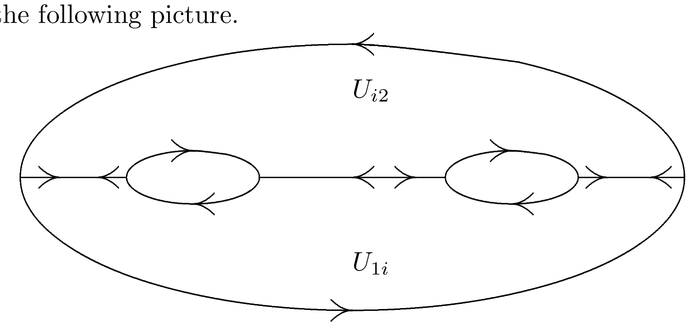
$$\int_{U_i} -(Q_x - P_y) \, dm_2,$$

just take the other orientation for Γ. This proves the theorem.

With this wonderful theorem, it is possible to give an analytic description of the two different orientations of a rectifiable simple closed curve. The positive orientation is the one for which Greens theorem holds and the other one, called the negative orientation is the one for which

$$\int_\Gamma \mathbf{f} \cdot d\boldsymbol{\gamma} = \int_{U_i} (P_y - Q_x) \, dm_2.$$

There are other regions for which Green's theorem holds besides just the inside and boundary of a simple closed curve. For Γ a simple closed curve and U_i its inside, lets refer to $U_i \cup \Gamma$ as a Jordan region. When you have two non overlapping Jordan regions which intersect in a finite number of simple curves, you can delete the interiors of these simple curves and what results will also be a region for which Green's theorem holds. This is illustrated in the following picture.

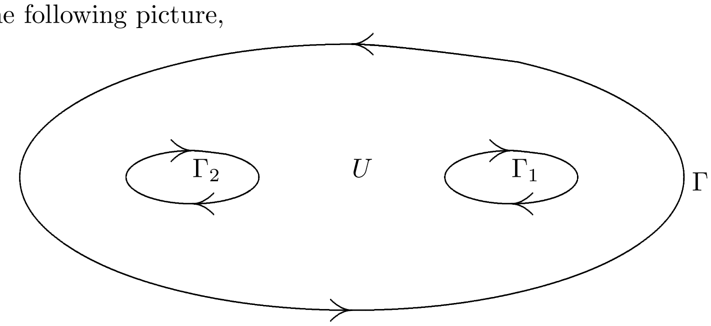


There are two Jordan regions here with insides U_{1i} and U_{2i} and these regions intersect in three simple curves. As indicated in the picture, opposite orientations are given to each of these three simple curves. Then the line integrals over these cancel. The area integrals add. Recall the two dimensional area of a bounded variation curve equals 0.

Denote by Γ the curve on the outside of the whole thing and Γ_1 and Γ_2 the oriented boundaries of the two holes which result when the curves of intersection are removed, the orientations as shown. Then letting $\mathbf{f}(x,y) = (P(x,y), Q(x,y))$, and

$$U = U_{1i} \cup U_{2i} \cup \{\text{Open segments of intersection}\}$$

as shown in the following picture,



it follows from applying Green's theorem to both of the Jordan regions,

$$\begin{aligned}\int_\Gamma \mathbf{f}\cdot d\boldsymbol{\gamma} + \int_{\Gamma_1} \mathbf{f}\cdot d\boldsymbol{\gamma}_1 + \int_{\Gamma_2} \mathbf{f}\cdot d\gamma_2 &= \int_{U_{1i}\cup U_{2i}} (Q_x - P_y)\, dm_2 \\ &= \int_U (Q_x - P_y)\, dm_2\end{aligned}$$

To make this simpler, just write it in the form

$$\int_{\partial U} \mathbf{f}\cdot d\boldsymbol{\gamma} = \int_U (Q_x - P_y)\, dm_2$$

where ∂U is oriented as indicated in the picture and involves the three oriented curves $\Gamma, \Gamma_1, \Gamma_2$.

11.4 Stoke's Theorem

Stokes theorem is usually presented in calculus courses under far more restrictive assumptions than will be used here. It turns out that all the hard questions are related to Green's theorem and that when you have the general version of Green's theorem this can be used to obtain a general version of Stoke's theorem using a simple identity. This is because Stoke's theorem is really just a three dimensional version of the two dimensional Green's theorem. This will be made more precise below.

To begin with suppose Γ is a rectifiable curve in $\mathbb{R}^2$ having parameterization $\boldsymbol{\alpha} : [a,b] \to \Gamma$ for $\boldsymbol{\alpha}$ a continuous function. Let $\mathbf{R} : U \to \mathbb{R}^n$ be a C^1 function where U contains $\boldsymbol{\alpha}^*$. Then one could define a curve

$$\boldsymbol{\gamma}(t) \equiv \mathbf{R}(\boldsymbol{\alpha}(t)),\ t \in [a,b].$$

Lemma 11.4.1 *The curve $\boldsymbol{\gamma}^*$ where $\boldsymbol{\gamma}$ is as just described is a rectifiable curve. If $\mathbf{F}$ is defined and continuous on $\boldsymbol{\gamma}^*$ then*

$$\int_{\boldsymbol{\gamma}} \mathbf{F}\cdot d\boldsymbol{\gamma} = \int_{\boldsymbol{\alpha}} ((\mathbf{F}\circ\mathbf{R})\cdot\mathbf{R}_u, (\mathbf{F}\circ\mathbf{R})\cdot\mathbf{R}_v)\cdot d\boldsymbol{\alpha}$$

where $\mathbf{R}_u$ signifies the partial derivative of $\mathbf{R}$ with respect to the variable u.

Proof: Let

$$K \equiv \left\{ \mathbf{y} \in \mathbb{R}^2 : \text{dist}\left(\mathbf{y}, \boldsymbol{\alpha}^*\right) \leq r \right\}$$

where r is small enough that $K \subseteq U$. This is easily done because $\boldsymbol{\alpha}^*$ is compact. Let

$$C_K \equiv \max \left\{ \left|\left| D\mathbf{R}\left(\mathbf{x}\right)\right|\right| : \mathbf{x} \in K \right\}$$

Consider

$$\sum_{j=0}^{n-1} \left|\mathbf{R}\left(\boldsymbol{\alpha}\left(t_{j+1}\right)\right) - \mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right| \tag{11.21}$$

where $\{t_0, \cdots, t_n\}$ is a partition of $[a, b]$. Since $\boldsymbol{\alpha}$ is continuous, there exists a δ such that if $||\mathcal{P}|| < \delta$, then the segment

$$\left\{\boldsymbol{\alpha}\left(t_j\right) + s\left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right) : s \in [0, 1]\right\}$$

is contained in K. Therefore, by the mean value inequality, Theorem 6.4.2,

$$\sum_{j=0}^{n-1} \left|\mathbf{R}\left(\boldsymbol{\alpha}\left(t_{j+1}\right)\right) - \mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right| \leq \sum_{j=0}^{n-1} C_K \left|\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right|$$

Now if $\mathcal{P}$ is any partition, 11.21 can always be made larger by adding in points to $\mathcal{P}$ till $||\mathcal{P}|| < \delta$ and so this shows

$$V\left(\boldsymbol{\gamma}, [a, b]\right) \leq C_K V\left(\boldsymbol{\alpha}, [a, b]\right).$$

This proves the first part.

Next consider the claim about the integral. Let

$$G\left(\mathbf{v}, \mathbf{x}\right) \equiv \mathbf{R}\left(\mathbf{x} + \mathbf{v}\right) - \mathbf{R}\left(\mathbf{x}\right) - D\mathbf{R}\left(\mathbf{x}\right)\left(\mathbf{v}\right).$$

Then

$$D_1 G\left(\mathbf{v}, \mathbf{x}\right) = D\mathbf{R}\left(\mathbf{x} + \mathbf{v}\right) - D\mathbf{R}\left(\mathbf{x}\right)$$

and so by uniform continuity of $D\mathbf{R}$ on the compact set K, it follows there exists $\delta > 0$ such that if $|\mathbf{v}| < \delta$, then for all $\mathbf{x} \in \boldsymbol{\alpha}^*$,

$$\left|\left|D\mathbf{R}\left(\mathbf{x} + \mathbf{v}\right) - D\mathbf{R}\left(\mathbf{x}\right)\right|\right| = \left|\left|D_1 G\left(\mathbf{v}, \mathbf{x}\right)\right|\right| < \varepsilon.$$

By Theorem 6.4.2 again it follows that for all $\mathbf{x} \in \boldsymbol{\alpha}^*$ and $|\mathbf{v}| < \delta$,

$$\left|G\left(\mathbf{v}, \mathbf{x}\right)\right| = \left|\mathbf{R}\left(\mathbf{x} + \mathbf{v}\right) - \mathbf{R}\left(\mathbf{x}\right) - D\mathbf{R}\left(\mathbf{x}\right)\left(\mathbf{v}\right)\right| \leq \varepsilon \left|\mathbf{v}\right| \tag{11.22}$$

Letting $||\mathcal{P}||$ be small enough, it follows from the continuity of $\boldsymbol{\alpha}$ that

$$\left|\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right| < \delta$$

Therefore for such $\mathcal{P}$,

$$\sum_{j=0}^{n-1} \mathbf{F}\left(\boldsymbol{\gamma}\left(t_j\right)\right) \cdot \left(\boldsymbol{\gamma}\left(t_{j+1}\right) - \boldsymbol{\gamma}\left(t_j\right)\right)$$

$$= \sum_{j=0}^{n-1} \mathbf{F}\left(\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right) \cdot \left(\mathbf{R}\left(\boldsymbol{\alpha}\left(t_{j+1}\right)\right) - \mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right)$$

$$= \sum_{j=0}^{n-1} \mathbf{F}\left(\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right) \cdot \left[D\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right) + \mathbf{o}\left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right)\right]$$

where

$$\mathbf{o}\left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right) = \mathbf{R}\left(\boldsymbol{\alpha}\left(t_{j+1}\right)\right) - \mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right) - D\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right)$$

and by 11.22,

$$\left|\mathbf{o}\left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right)\right| < \varepsilon \left|\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right|$$

It follows

$$\left|\sum_{j=0}^{n-1} \mathbf{F}\left(\boldsymbol{\gamma}\left(t_j\right)\right) \cdot \left(\boldsymbol{\gamma}\left(t_{j+1}\right) - \boldsymbol{\gamma}\left(t_j\right)\right) - \sum_{j=0}^{n-1} \mathbf{F}\left(\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right) \cdot D\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right)\right| \tag{11.23}$$

$$\leq \sum_{j=0}^{n-1} \left|\mathbf{o}\left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right)\right| \leq \sum_{j=0}^{n-1} \varepsilon \left|\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right| \leq \varepsilon V\left(\boldsymbol{\alpha}, [a, b]\right)$$

Consider the second sum in 11.23. A term in the sum equals

$$\mathbf{F}\left(\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right) \cdot \left(\mathbf{R}_u\left(\boldsymbol{\alpha}\left(t_j\right)\right)\left(\alpha_1\left(t_{j+1}\right) - \alpha_1\left(t_j\right)\right) + \mathbf{R}_v\left(\boldsymbol{\alpha}\left(t_j\right)\right)\left(\alpha_2\left(t_{j+1}\right) - \alpha_2\left(t_j\right)\right)\right)$$

$$= \left(\mathbf{F}\left(\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right) \cdot \mathbf{R}_u\left(\boldsymbol{\alpha}\left(t_j\right)\right), \mathbf{F}\left(\mathbf{R}\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right) \cdot \mathbf{R}_v\left(\boldsymbol{\alpha}\left(t_j\right)\right)\right) \cdot \left(\boldsymbol{\alpha}\left(t_{j+1}\right) - \boldsymbol{\alpha}\left(t_j\right)\right)$$

By continuity of $\mathbf{F}$, $\mathbf{R}_u$ and $\mathbf{R}_v$, it follows that sum converges as $||\mathcal{P}|| \to 0$ to

$$\int_{\boldsymbol{\alpha}} \left(\left(\mathbf{F} \circ \mathbf{R}\right) \cdot \mathbf{R}_u, \left(\mathbf{F} \circ \mathbf{R}\right) \cdot \mathbf{R}_v\right) \cdot d\boldsymbol{\alpha}$$

Therefore, taking the limit as $||\mathcal{P}|| \to 0$ in 11.23

$$\left|\int_{\boldsymbol{\gamma}} \mathbf{F} \cdot d\boldsymbol{\gamma} - \int_{\boldsymbol{\alpha}} \left(\left(\mathbf{F} \circ \mathbf{R}\right) \cdot \mathbf{R}_u, \left(\mathbf{F} \circ \mathbf{R}\right) \cdot \mathbf{R}_v\right) \cdot d\boldsymbol{\alpha}\right| < \varepsilon V\left(\boldsymbol{\alpha}, [a, b]\right).$$

Since $\varepsilon > 0$ is arbitrary, this proves the lemma.

The following is a little identity which will allow a proof of Stoke's theorem to follow from Green's theorem. First recall the following definition from calculus of the curl of a vector field and the cross product of two vectors from calculus.

Definition 11.4.2 *Let $\mathbf{u} \equiv (a, b, c)$ and $\mathbf{v} \equiv (d, e, f)$ be two vectors in $\mathbb{R}^3$. Then*

$$\mathbf{u} \times \mathbf{v} \equiv \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a & b & c \\ d & e & f \end{vmatrix}$$

*where the determinant is expanded formally along the top row. Let $\mathbf{f} : U \to \mathbb{R}^3$ for $U \subseteq \mathbb{R}^3$ denote a vector field. The **curl** of the vector field yields another vector field and it is defined as follows.*

$$\left(\operatorname{curl}\left(\mathbf{f}\right)\left(\mathbf{x}\right)\right)_i \equiv \left(\nabla \times \mathbf{f}\left(\mathbf{x}\right)\right)_i$$

where here ∂_j means the partial derivative with respect to x_j and the subscript of i in $\left(\operatorname{curl}\left(\mathbf{f}\right)\left(\mathbf{x}\right)\right)_i$ means the i^{th} Cartesian component of the vector, $\operatorname{curl}\left(\mathbf{f}\right)\left(\mathbf{x}\right)$. Thus the curl is evaluated by expanding the following determinant along the top row.

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1\left(x, y, z\right) & f_2\left(x, y, z\right) & f_3\left(x, y, z\right) \end{vmatrix}.$$

Note the similarity with the cross product. More precisely and less evocatively,

$$\nabla \times \mathbf{f}(x,y,z) = \equiv \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k}.$$

In the above, $\mathbf{i} = \mathbf{e}_1, \mathbf{j} = \mathbf{e}_2$, *and* $\mathbf{k} = \mathbf{e}_3$ *the standard unit basis vectors for* $\mathbb{R}^3$.

With this definition, here is the identity.

Lemma 11.4.3 *Let* $\mathbf{R} : U \to V \subseteq \mathbb{R}^3$ *where* U *is an open subset of* $\mathbb{R}^2$ *and* V *is an open subset of* $\mathbb{R}^3$. *Suppose* $\mathbf{R}$ *is* C^2 *and let* $\mathbf{F}$ *be a* C^1 *vector field defined in* V.

$$(\mathbf{R}_u \times \mathbf{R}_v) \cdot (\nabla \times \mathbf{F})(\mathbf{R}(u,v)) = \left((\mathbf{F} \circ \mathbf{R})_u \cdot \mathbf{R}_v - (\mathbf{F} \circ \mathbf{R})_v \cdot \mathbf{R}_u\right)(u,v). \tag{11.24}$$

Proof: Letting x, y, z denote the components of $\mathbf{R}(\mathbf{u})$ and f_1, f_2, f_3 denote the components of $\mathbf{F}$, and letting a subscripted variable denote the partial derivative with respect to that variable, the left side of 11.24 equals

$$\begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_u & y_u & z_u \\ x_v & y_v & z_v \end{vmatrix} \cdot \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ f_1 & f_2 & f_3 \end{vmatrix}$$

$$= (f_{3y} - f_{2z})(y_u z_v - z_u y_v) + (f_{1z} - f_{3x})(z_u x_v - x_u z_v) + (f_{2x} - f_{1y})(x_u y_v - y_u x_v)$$

$$\begin{aligned} &= f_{3y} y_u z_v + f_{2z} z_u y_v + f_{1z} z_u x_v + f_{3x} x_u z_v + f_{2x} x_u y_v + f_{1y} y_u x_v \\ &\quad - (f_{2z} y_u z_v + f_{3y} z_u y_v + f_{1z} x_u z_v + f_{3x} z_u x_v + f_{2x} y_u x_v + f_{1y} x_u y_v) \end{aligned}$$

$$\begin{aligned} &= f_{1y} y_u x_v + f_{1z} z_u x_v + f_{2x} x_u y_v + f_{2z} z_u y_v + f_{3x} x_u z_v + f_{3y} y_u z_v \\ &\quad - (f_{1y} y_v x_u + f_{1z} z_v x_u + f_{2x} x_v y_u + f_{2z} z_v y_u + f_{3x} x_v z_u + f_{3y} y_v z_u) \end{aligned}$$

At this point add in and subtract off certain terms. Then the above equals

$$\begin{aligned} &= f_{1x} x_u x_v + f_{1y} y_u x_v + f_{1z} z_u x_v + f_{2x} x_u y_v + f_{2x} y_u y_v \\ &\quad + f_{2z} z_u y_v + f_{3x} x_u z_v + f_{3y} y_u z_v + f_{3z} z_u z_v \\ &\quad - \begin{pmatrix} f_{1x} x_v x_u + f_{1y} y_v x_u + f_{1z} z_v x_u + f_{2x} x_v y_u + f_{2x} y_v y_u \\ + f_{2z} z_v y_u + f_{3x} x_v z_u + f_{3y} y_v z_u + f_{3z} z_v z_u \end{pmatrix} \end{aligned}$$

$$\begin{aligned} &= \frac{\partial f_1 \circ \mathbf{R}(u,v)}{\partial u} x_v + \frac{\partial f_2 \circ \mathbf{R}(u,v)}{\partial u} y_v + \frac{\partial f_3 \circ \mathbf{R}(u,v)}{\partial u} z_v \\ &\quad - \left(\frac{\partial f_1 \circ \mathbf{R}(u,v)}{\partial v} x_u + \frac{\partial f_2 \circ \mathbf{R}(u,v)}{\partial v} y_u + \frac{\partial f_3 \circ \mathbf{R}(u,v)}{\partial v} z_u\right) \end{aligned}$$

$$= \left((\mathbf{F} \circ \mathbf{R})_u \cdot \mathbf{R}_v - (\mathbf{F} \circ \mathbf{R})_v \cdot \mathbf{R}_u\right)(u,v).$$

This proves the lemma.

Let U be a region in $\mathbb{R}^2$ for which Green's theorem holds. Thus Green's theorem says that for P, Q continuous on $U_i \cup \Gamma, P_v, Q_u \in L^1(U_i \cup \Gamma), P, Q$ being C^1 on U_i,

$$\int_U (Q_u - P_v)\, dm_2 = \int_{\partial U} \mathbf{f} \cdot d\boldsymbol{\alpha}$$

where ∂U consists of some simple closed rectifiable oriented curves as explained above. Here the u and v axes are in the same relation as the x and y axes.

Theorem 11.4.4 *(Stoke's Theorem) Let U be any region in $\mathbb{R}^2$ for which the conclusion of Green's theorem holds. Let $\mathbf{R} \in C^2\left(\overline{U}, \mathbb{R}^3\right)$ be a one to one function. Let*

$$\boldsymbol{\gamma}_j = \mathbf{R} \circ \boldsymbol{\alpha}_j,$$

where the $\boldsymbol{\alpha}_j$ are parameterizations for the oriented curves making up the boundary of U such that the conclusion of Green's theorem holds. Let S denote the surface,

$$S \equiv \{\mathbf{R}(u,v) : (u,v) \in U\},$$

Then for $\mathbf{F}$ a C^1 vector field defined near S,

$$\sum_{i=1}^{n} \int_{\boldsymbol{\gamma}_i} \mathbf{F} \cdot d\boldsymbol{\gamma}_i = \int_U (\mathbf{R}_u(u,v) \times \mathbf{R}_v(u,v)) \cdot (\nabla \times \mathbf{F}(\mathbf{R}(u,v)))\, dm_2$$

Proof: By Lemma 11.4.1,

$$\sum_{j=1}^{n} \int_{\boldsymbol{\gamma}_j} \mathbf{F} \cdot d\boldsymbol{\gamma}_j =$$

$$\sum_{j=1}^{n} \int_{\boldsymbol{\alpha}_j} ((\mathbf{F} \circ \mathbf{R}) \cdot \mathbf{R}_u, (\mathbf{F} \circ \mathbf{R}) \cdot \mathbf{R}_v) \cdot d\boldsymbol{\alpha}_j$$

By the assumption that the conclusion of Green's theorem holds for U, this equals

$$\begin{aligned} & \int_U \left[((\mathbf{F} \circ \mathbf{R}) \cdot \mathbf{R}_v)_u - ((\mathbf{F} \circ \mathbf{R}) \cdot \mathbf{R}_u)_v\right] dm_2 \\ = & \int_U \left[(\mathbf{F} \circ \mathbf{R})_u \cdot \mathbf{R}_v + (\mathbf{F} \circ \mathbf{R}) \cdot \mathbf{R}_{vu} - (\mathbf{F} \circ \mathbf{R}) \cdot \mathbf{R}_{uv} - (\mathbf{F} \circ \mathbf{R})_v \cdot \mathbf{R}_u\right] dm_2 \\ = & \int_U \left[(\mathbf{F} \circ \mathbf{R})_u \cdot \mathbf{R}_v - (\mathbf{F} \circ \mathbf{R})_v \cdot \mathbf{R}_u\right] dm_2 \end{aligned}$$

the last step holding by equality of mixed partial derivatives, a result of the assumption that $\mathbf{R}$ is C^2. Now by Lemma 11.4.3, this equals

$$\int_U (\mathbf{R}_u(u,v) \times \mathbf{R}_v(u,v)) \cdot (\nabla \times \mathbf{F}(\mathbf{R}(u,v)))\, dm_2$$

This proves Stoke's theorem.

With approximation arguments one can remove the assumption that $\mathbf{R}$ is C^2 and replace this condition with weaker conditions. This is not surprising because in the final result, only first derivatives of $\mathbf{R}$ occur.

11.5 Interpretation And Review

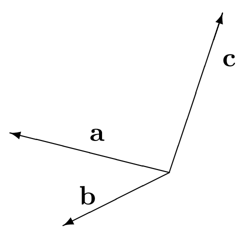
To understand the interpretation of Stoke's theorem in terms of an integral over the surface S, it is necessary to either do more theoretical development or to review some beginning calculus. I will do the latter here. First of all, it is important to understand the geometrical properties of the cross product. Those who have had a typical calculus course will probably not have seen this so I will present it here. It is elementary material which is a little out of place in an advanced calculus book but it is nevertheless useful and important and if you have not seen it, you should.

11.5.1 The Geometric Description Of The Cross Product

The cross product is a way of multiplying two vectors in $\mathbb{R}^3$. It is very different from the dot product in many ways. First the geometric meaning is discussed and then a description in terms of coordinates is given. Both descriptions of the cross product are important. The geometric description is essential in order to understand the applications to physics and geometry while the coordinate description is the only way to practically compute the cross product. In this presentation a vector is something which is characterized by direction and magnitude.

Definition 11.5.1 *Three vectors,* $\mathbf{a}, \mathbf{b}, \mathbf{c}$ *form a right handed system if when you extend the fingers of your right hand along the vector,* $\mathbf{a}$ *and close them in the direction of* $\mathbf{b}$, *the thumb points roughly in the direction of* $\mathbf{c}$.

For an example of a right handed system of vectors, see the following picture.



In this picture the vector $\mathbf{c}$ points upwards from the plane determined by the other two vectors. You should consider how a right hand system would differ from a left hand system. Try using your left hand and you will see that the vector, $\mathbf{c}$ would need to point in the opposite direction as it would for a right hand system.

From now on, the vectors, $\mathbf{i}, \mathbf{j}, \mathbf{k}$ will **always** form a right handed system. To repeat, if you extend the fingers of your right hand along $\mathbf{i}$ and close them in the direction $\mathbf{j}$, the thumb points in the direction of $\mathbf{k}$. Recall these are the basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$.

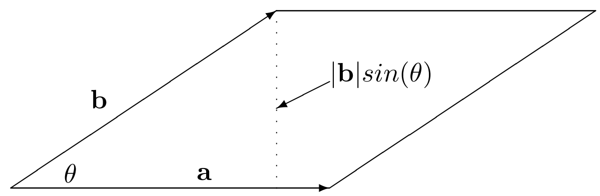
The following is the geometric description of the cross product. It gives both the direction and the magnitude and therefore specifies the vector.

Definition 11.5.2 *Let* $\mathbf{a}$ *and* $\mathbf{b}$ *be two vectors in* $\mathbb{R}^3$. *Then* $\mathbf{a} \times \mathbf{b}$ *is defined by the following two rules.*

1. $|\mathbf{a} \times \mathbf{b}| = |\mathbf{a}| |\mathbf{b}| \sin \theta$ *where* θ *is the included angle.*

2. $\mathbf{a} \times \mathbf{b} \cdot \mathbf{a} = 0$, $\mathbf{a} \times \mathbf{b} \cdot \mathbf{b} = 0$, *and* $\mathbf{a}, \mathbf{b}, \mathbf{a} \times \mathbf{b}$ *forms a right hand system.*

Note that $|\mathbf{a} \times \mathbf{b}|$ is the **area of the parallelogram** spanned by $\mathbf{a}$ and $\mathbf{b}$.



The cross product satisfies the following properties.

$$\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a}) \ , \ \mathbf{a} \times \mathbf{a} = \mathbf{0}, \tag{11.25}$$

For α a scalar,

$$(\alpha \mathbf{a}) \times \mathbf{b} = \alpha (\mathbf{a} \times \mathbf{b}) = \mathbf{a} \times (\alpha \mathbf{b}), \tag{11.26}$$

For $\mathbf{a}, \mathbf{b}$, and $\mathbf{c}$ vectors, one obtains the distributive laws,

$$\mathbf{a} \times (\mathbf{b} + \mathbf{c}) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}, \tag{11.27}$$

$$(\mathbf{b} + \mathbf{c}) \times \mathbf{a} = \mathbf{b} \times \mathbf{a} + \mathbf{c} \times \mathbf{a}. \tag{11.28}$$

Formula 11.25 follows immediately from the definition. The vectors $\mathbf{a} \times \mathbf{b}$ and $\mathbf{b} \times \mathbf{a}$ have the same magnitude, $|\mathbf{a}| |\mathbf{b}| \sin \theta$, and an application of the right hand rule shows they have opposite direction. Formula 11.26 is also fairly clear. If α is a nonnegative scalar, the direction of $(\alpha \mathbf{a}) \times \mathbf{b}$ is the same as the direction of $\mathbf{a} \times \mathbf{b}$,$\alpha (\mathbf{a} \times \mathbf{b})$ and $\mathbf{a} \times (\alpha \mathbf{b})$ while the magnitude is just α times the magnitude of $\mathbf{a} \times \mathbf{b}$ which is the same as the magnitude of $\alpha (\mathbf{a} \times \mathbf{b})$ and $\mathbf{a} \times (\alpha \mathbf{b})$. Using this yields equality in 11.26. In the case where $\alpha < 0$, everything works the same way except the vectors are all pointing in the opposite direction and you must multiply by $|\alpha|$ when comparing their magnitudes. The distributive laws are much harder to establish but the second follows from the first quite easily. Thus, assuming the first, and using 11.25,

$$\begin{aligned} (\mathbf{b} + \mathbf{c}) \times \mathbf{a} &= -\mathbf{a} \times (\mathbf{b} + \mathbf{c}) \\ &= -(\mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c}) \\ &= \mathbf{b} \times \mathbf{a} + \mathbf{c} \times \mathbf{a}. \end{aligned}$$

To verify the distributive law one can consider something called the box product.

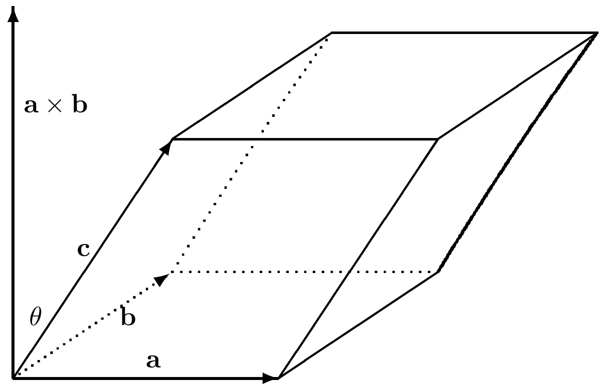
11.5.2 The Box Product, Triple Product

Definition 11.5.3 *A parallelepiped determined by the three vectors,* $\mathbf{a}, \mathbf{b}$, *and* $\mathbf{c}$ *consists of*

$$\{r\mathbf{a} + s\mathbf{b} + t\mathbf{c} : r, s, t \in [0, 1]\}.$$

That is, if you pick three numbers, $r, s,$ *and* t *each in* $[0, 1]$ *and form* $r\mathbf{a} + s\mathbf{b} + t\mathbf{c}$, *then the collection of all such points is what is meant by the parallelepiped determined by these three vectors.*

The following is a picture of such a thing.



You notice the area of the base of the parallelepiped, the parallelogram determined by the vectors, $\mathbf{a}$ and $\mathbf{b}$ has area equal to $|\mathbf{a}\times\mathbf{b}|$ while the altitude of the parallelepiped is $|\mathbf{c}|\cos\theta$ where θ is the angle shown in the picture between $\mathbf{c}$ and $\mathbf{a}\times\mathbf{b}$. Therefore, the volume of this parallelepiped is the area of the base times the altitude which is just

$$|\mathbf{a}\times\mathbf{b}|\,|\mathbf{c}|\cos\theta=\mathbf{a}\times\mathbf{b}\cdot\mathbf{c}.$$

This expression is known as the box product and is sometimes written as $[\mathbf{a},\mathbf{b},\mathbf{c}]$. You should consider what happens if you interchange the $\mathbf{b}$ with the $\mathbf{c}$ or the $\mathbf{a}$ with the $\mathbf{c}$. You can see geometrically from drawing pictures that this merely introduces a minus sign. In any case the box product of three vectors always equals either the volume of the parallelepiped determined by the three vectors or else minus this volume. From geometric reasoning like this you see that

$$\mathbf{a}\cdot\mathbf{b}\times\mathbf{c}=\mathbf{a}\times\mathbf{b}\cdot\mathbf{c}.$$

In other words, you can switch the $\times$ and the $\cdot$.

11.5.3 A Proof Of The Distributive Law For The Cross Product

Here is a proof of the distributive law for the cross product. Let $\mathbf{x}$ be a vector. From the above observation,

$$\begin{aligned}\mathbf{x}\cdot\mathbf{a}\times(\mathbf{b}+\mathbf{c}) &= (\mathbf{x}\times\mathbf{a})\cdot(\mathbf{b}+\mathbf{c})\\ &= (\mathbf{x}\times\mathbf{a})\cdot\mathbf{b}+(\mathbf{x}\times\mathbf{a})\cdot\mathbf{c}\\ &= \mathbf{x}\cdot\mathbf{a}\times\mathbf{b}+\mathbf{x}\cdot\mathbf{a}\times\mathbf{c}\\ &= \mathbf{x}\cdot(\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}).\end{aligned}$$

Therefore,

$$\mathbf{x}\cdot[\mathbf{a}\times(\mathbf{b}+\mathbf{c})-(\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c})]=0$$

for all $\mathbf{x}$. In particular, this holds for $\mathbf{x}=\mathbf{a}\times(\mathbf{b}+\mathbf{c})-(\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c})$ showing that $\mathbf{a}\times(\mathbf{b}+\mathbf{c})=\mathbf{a}\times\mathbf{b}+\mathbf{a}\times\mathbf{c}$ and this proves the distributive law for the cross product.

11.5.4 The Coordinate Description Of The Cross Product

Now from the properties of the cross product and its definition,

$$\begin{array}{ll}\mathbf{i}\times\mathbf{j}=\mathbf{k} & \mathbf{j}\times\mathbf{i}=-\mathbf{k}\\ \mathbf{k}\times\mathbf{i}=\mathbf{j} & \mathbf{i}\times\mathbf{k}=-\mathbf{j}\\ \mathbf{j}\times\mathbf{k}=\mathbf{i} & \mathbf{k}\times\mathbf{j}=-\mathbf{i}\end{array}$$

With this information, the following gives the coordinate description of the cross product.

Proposition 11.5.4 *Let* $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ *and* $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$ *be two vectors. Then*

$$\begin{aligned}\mathbf{a}\times\mathbf{b} = &\left(a_2b_3 - a_3b_2\right)\mathbf{i} + \left(a_3b_1 - a_1b_3\right)\mathbf{j} + \\ &+\left(a_1b_2 - a_2b_1\right)\mathbf{k}. \end{aligned} \tag{11.29}$$

Proof: From the above table and the properties of the cross product listed,

$$\begin{aligned}&\left(a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}\right)\times\left(b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}\right) = \\ &a_1b_2\mathbf{i}\times\mathbf{j} + a_1b_3\mathbf{i}\times\mathbf{k} + a_2b_1\mathbf{j}\times\mathbf{i} + a_2b_3\mathbf{j}\times\mathbf{k} + \\ &+a_3b_1\mathbf{k}\times\mathbf{i} + a_3b_2\mathbf{k}\times\mathbf{j} \\ &= a_1b_2\mathbf{k} - a_1b_3\mathbf{j} - a_2b_1\mathbf{k} + a_2b_3\mathbf{i} + a_3b_1\mathbf{j} - a_3b_2\mathbf{i} \\ &= \left(a_2b_3 - a_3b_2\right)\mathbf{i} + \left(a_3b_1 - a_1b_3\right)\mathbf{j} + \left(a_1b_2 - a_2b_1\right)\mathbf{k}\end{aligned} \tag{11.30}$$

This proves the proposition.

The easy way to remember the above formula is to write it as follows.

$$\mathbf{a}\times\mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \tag{11.31}$$

where you expand the determinant along the top row. This yields

$$\left(a_2b_3 - a_3b_2\right)\mathbf{i} - \left(a_1b_3 - a_3b_1\right)\mathbf{j} + \left(a_1b_2 - a_2b_1\right)\mathbf{k} \tag{11.32}$$

which is the same as 11.30.

11.5.5 The Integral Over A Two Dimensional Surface

First it is good to define what is meant by a smooth surface.

Definition 11.5.5 *Let S be a subset of $\mathbb{R}^3$. Then S is a **smooth surface** if there exists an open set, $U \subseteq \mathbb{R}^2$ and a C^1 function, $\mathbf{R}$ defined on U such that $\mathbf{R}(U) = S, \mathbf{R}$ is one to one, and for all $(u, v) \in U$,*

$$\mathbf{R}_u \times \mathbf{R}_v \neq \mathbf{0}. \tag{11.33}$$

This last condition ensures that there is always a well defined normal on S. This function, $\mathbf{R}$ is called a parameterization of the surface. It is just like a parameterization of a curve but here there are two parameters, u, v.

One way to think of this is that there is a piece of rubber occupying U in the plane and then it is taken and stretched in three dimensions. This gives S.

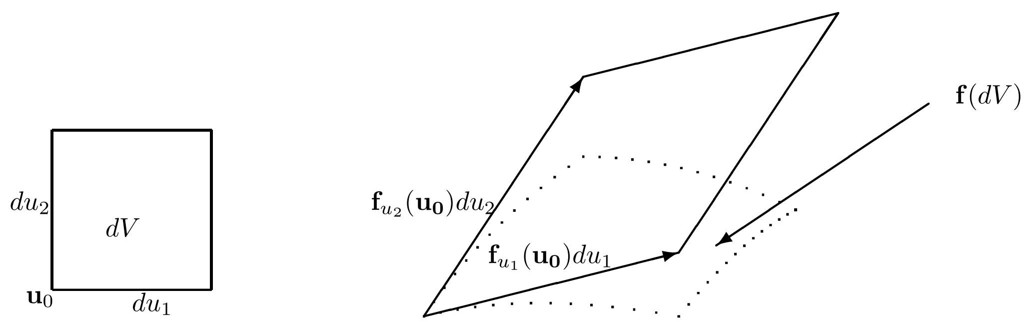
Definition 11.5.6 *Let $\mathbf{u}_1, \mathbf{u}_2$ be vectors in $\mathbb{R}^3$. The 2 dimensional parallelogram determined by these vectors will be denoted by $P(\mathbf{u}_1, \mathbf{u}_2)$ and it is defined as*

$$P(\mathbf{u}_1, \mathbf{u}_2) \equiv \left\{\sum_{j=1}^{2} s_j\mathbf{u}_j : s_j \in [0, 1]\right\}.$$

Then the area of this parallelogram is

$$\text{area } P(\mathbf{u}_1, \mathbf{u}_2) \equiv \left|\mathbf{u}_1 \times \mathbf{u}_2\right|.$$

Suppose then that $\mathbf{x} = \mathbf{R}(\mathbf{u})$ where $\mathbf{u} \in U$, a subset of $\mathbb{R}^2$ and $\mathbf{x}$ is a point in V, a subset of 3 dimensional space. Thus, letting the Cartesian coordinates of $\mathbf{x}$ be given by $\mathbf{x} = (x_1, x_2, x_3)^T$, each x_i being a function of $\mathbf{u}$, an infinitesimal rectangle located at $\mathbf{u}_0$ corresponds to an infinitesimal parallelogram located at $\mathbf{R}(\mathbf{u}_0)$ which is determined by the 2 vectors $\left\{\frac{\partial \mathbf{R}(\mathbf{u}_0)}{\partial u} du, \frac{\partial \mathbf{R}(\mathbf{u}_0)}{\partial v} dv\right\}$, each of which is tangent to the surface defined by $\mathbf{x} = \mathbf{R}(\mathbf{u})$. This is a very vague and unacceptable description. What exactly is an infinitesimal rectangle? However, it can all be made precise later and this is good motivation for the real thing.



From Definition 11.5.6, the volume of this infinitesimal parallelepiped located at $\mathbf{R}(\mathbf{u}_0)$ is given by

$$\left|\frac{\partial \mathbf{R}(\mathbf{u}_0)}{\partial u} du \times \frac{\partial \mathbf{R}(\mathbf{u}_0)}{\partial v} dv\right| = \left|\frac{\partial \mathbf{R}(\mathbf{u}_0)}{\partial u} \times \frac{\partial \mathbf{R}(\mathbf{u}_0)}{\partial v}\right| dudv \tag{11.34}$$

$$= |\mathbf{R}_u \times \mathbf{R}_v| \, dudv \tag{11.35}$$

This motivates the following definition of what is meant by the integral over a parametrically defined surface in $\mathbb{R}^3$.

Definition 11.5.7 *Suppose U is a subset of $\mathbb{R}^2$ and suppose $\mathbf{R} : U \to \mathbf{R}(U) = S \subseteq \mathbb{R}^3$ is a one to one and C^1 function. Then if $h : \mathbf{R}(U) \to \mathbb{R}$, define the 2 dimensional surface integral, $\int_{\mathbf{R}(U)} h(\mathbf{x}) \, dS$ according to the following formula.*

$$\int_{\mathbf{S}} h(\mathbf{x}) \, dS \equiv \int_U h(\mathbf{R}(\mathbf{u})) \left|\mathbf{R}_u(\mathbf{u}) \times \mathbf{R}_v(\mathbf{u})\right| dudv.$$

With this understanding, it becomes possible to interpret the meaning of Stoke's theorem. This is stated in the following theorem. Note that slightly more is assumed here than earlier. In particular, it is assumed that $\mathbf{R}_u \times \mathbf{R}_v \neq \mathbf{0}$. This allows the definition of a well defined normal vector which varies continuously over the surface, S.

Theorem 11.5.8 *(Stoke's Theorem) Let U be any region in $\mathbb{R}^2$ for which the conclusion of Green's theorem holds. Let $\mathbf{R} \in C^2\left(\overline{U_i}, \mathbb{R}^3\right)$ be a one to one function such that $\mathbf{R}_u \times \mathbf{R}_v \neq \mathbf{0}$ on U. Let*

$$\boldsymbol{\gamma}_j = \mathbf{R} \circ \boldsymbol{\alpha}_j,$$

where the $\boldsymbol{\alpha}_j$ are parameterizations for the oriented bounded variation curves bounding the region U oriented such that the conclusion of Green's theorem holds. Let S denote the surface,

$$S \equiv \{\mathbf{R}(u, v) : (u, v) \in U\},$$

Then for $\mathbf{F}$ *a* C^1 *vector field defined near* S,

$$\sum_{j=1}^{n}\int_{\gamma_j}\mathbf{F}\cdot d\boldsymbol{\gamma}_i = \int_U (\mathbf{R}_u\times\mathbf{R}_v)\cdot(\nabla\times\mathbf{F})(\mathbf{R}(u,v))\,dm_2 \tag{11.36}$$

$$= \int_S (\nabla\times\mathbf{F})\cdot\mathbf{n}dS \tag{11.37}$$

Proof: Formula 11.36 was established in Theorem 11.4.4. The unit normal of the point $\mathbf{R}(u,v)$ of S is $(\mathbf{R}_u\times\mathbf{R}_v)/|\mathbf{R}_u\times\mathbf{R}_v|$ and from the definition of the integral over the surface, Definition 11.5.7, Formula 11.37 follows.

11.6 Introduction To Complex Analysis

11.6.1 Basic Theorems, The Cauchy Riemann Equations

With Green's theorem and the technique of proof used in proving it, it is possible to present the most important parts of complex analysis almost effortlessly. I will do this here and leave some of the other parts for the exercises. Recall the complex numbers should be considered as points in the plane. Thus a complex number is of the form $x+iy$ where $i^2=-1$. The complex conjugate is defined by

$$\overline{x+iy}\equiv x-iy$$

and for z a complex number,

$$|z|\equiv(z\overline{z})^{1/2}=\sqrt{x^2+y^2}.$$

Thus when $x+iy$ is considered an ordered pair $(x,y)\in\mathbb{R}^2$ the magnitude of a complex number is nothing more than the usual norm of the ordered pair. Also for $z=x+iy, w=u+iv$,

$$|z-w|=\sqrt{(x-u)^2+(y-v)^2}$$

so in terms of all topological considerations, $\mathbb{R}^2$ is the same as $\mathbb{C}$. Thus to say $z\to f(z)$ is continuous, is the same as saying

$$(x,y)\to u(x,y),\ (x,y)\to v(x,y)$$

are continuous where $f(z)\equiv u(x,y)+iv(x,y)$ with u and v being called the real and imaginary parts of f. The only new thing is that writing an ordered pair (x,y) as $x+iy$ with the convention $i^2=-1$ makes $\mathbb{C}$ into a field. Now here is the definition of what it means for a function to be analytic.

Definition 11.6.1 *Let* U *be an open subset of* $\mathbb{C}$ $(\mathbb{R}^2)$ *and let* $f:U\to\mathbb{C}$ *be a function. Then* f *is said to be analytic on* U *if for every* $z\in U$,

$$\lim_{\Delta z\to 0}\frac{f(z+\Delta z)-f(z)}{\Delta z}\equiv f'(z)$$

exists and is a continuous function of $z\in U$. *For a function having values in* $\mathbb{C}$ *denote by* $u(x,y)$ *the real part of* f *and* $v(x,y)$ *the imaginary part. Both* u *and* v *have real values and*

$$f(x+iy)\equiv f(z)\equiv u(x,y)+iv(x,y)$$