13.1) Curves

1. **Parametric and Vector Equations for Curves**:

As discussed in Section 12.1, in two-dimensional space, an equation in x and y gives us a curve (possibly a line, which is a *straight* curve), whereas in three-dimensional space, an equation in x, y, and z gives us a surface (possibly a plane, which is a *flat* surface). What kind of equation will give us a curve in three-dimensional space?

In Sections 12.2 and 12.5, we saw that a line in either two or three dimensions can be represented by *parametric equations*, which express x, y, and (in the three-dimensional case) z in terms of an *independent parameter*, usually t. We also saw that the parametric equations can be combined into a single *vector equation* for a line.

This same approach will now be applied to curves in general, whether two-dimensional or three-dimensional. (Three-dimensional curves are known as **space curves**, while two-dimensional curves are known as **plane curves**.)

The process of writing parametric equations to represent a curve is known as *parameterizing* the curve. There are many different ways to parameterize a given curve.

As with a line, any parametrization implies a particular **orientation** or **direction** for the curve. The *positive* or *forward* direction of the curve is the direction followed as the parameter *increases*, and the *negative* or *backward* direction of the curve is the direction followed as the parameter *decreases*. (When we refer to "the direction" of a curve, we mean the positive direction.) Different parameterizations may yield different orientations.

If a plane curve is simple and closed, like a circle, then its direction can be classified as *clockwise* or *counter-clockwise*. For plane curves such as vertical parabolas, direction can be classified as *leftward* or *rightward*. For plane curves such as horizontal parabolas, direction can be classified as *upward* or *downward*. For complicated plane curves, none of these classifications may be applicable. For space curves, direction can be even more challenging to describe.

A curve may be represented by a set of parametric equations:

- For a plane curve, $x = x(t)$, $y = y(t)$.
- For a space curve, $x = x(t)$, $y = y(t)$, $z = z(t)$.

The domain of the parametric equations could be $(-\infty, \infty)$, or it could be some other interval, such as $[0, \infty)$ or $[0, 2\pi]$ or $(-\frac{\pi}{2}, \frac{\pi}{2})$.

We usually think of t as representing *elapsed time*, and we think of the curve as the **path** of a *moving particle*. The parametric equations give a unique **position** for the particle at each point in time. This is known as a **motion paradigm**. Any particular value of t is referred to as an *instant*.

Any real value of t generates a unique point on the curve, denoted P_t. In the two-dimensional case, $P_t = (x(t), y(t))$. In the three-dimensional case, $P_t = (x(t), y(t), z(t))$. Of particular interest are the points P_0 and P_1, which are referred to as the **initial** or **starting point** and the **unitary point**, respectively. We typically write P_0 as (x_0, y_0) or (x_0, y_0, z_0) instead of $(x(0), y(0))$ or $(x(0), y(0), z(0))$. Likewise, we typically write P_1 as (x_1, y_1) or (x_1, y_1, z_1) instead of $(x(1), y(1))$ or $(x(1), y(1), z(1))$. The initial and unitary points depend on the chosen parameterization–i.e., different parametrizations may have different initial and unitary points. We can say the forward direction of the curve is the direction *from P_0 to P_1*.

The plane curve $y = \tan x$, for $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$, can be parameterized as $x = t,\ y = \tan t$. The domain is $(-\frac{\pi}{2}, \frac{\pi}{2})$. (In general, if a plane curve is the graph of a function, $y = f(x)$, then we can let $x = t$ and $y = f(t)$.) Here, the initial point is $(0,0)$, and the direction is rightward and upward.

If the circle $x^2 + y^2 = 9$ is parameterized as $x = 3\cos t,\ y = 3\sin t$, the starting point is $(3,0)$ and the direction is counter-clockwise. If we use the parameterization $x = -3\cos t$, $y = 3\sin t$, the starting point is $(-3,0)$ and the direction is clockwise. If we use the parameterization $x = 3\cos t,\ y = -3\sin t$, the starting point is $(3,0)$ and the direction is clockwise. If we use the parameterization $x = 3\sin t,\ y = 3\cos t$, the starting point is $(0,3)$ and the direction is clockwise. All of these parameterizations have the same domain, $[0, 2\pi]$ (assuming we want our particle to make one complete revolution).

The parabola $y = -2x^2 + 5$ could be parameterized so that $x = t,\ \ y = -2t^2 + 5$, in which case $P_0 = (0,5)$ and $P_1 = (1,3)$. Or it could be parameterized so that $x = -t,\ \ y = -2t^2 + 5$, in which case $P_0 = (0,5)$ and $P_1 = (-1,3)$. Or it could be parameterized so that $x = t + 4$, $y = -2t^2 - 16t - 27$, in which case $P_0 = (4,-27)$ and $P_1 = (5,-45)$. The first and third parameterizations give us a rightward orientation, whereas the second gives us a leftward orientation. All of these parameterizations have the same domain, $(-\infty, \infty)$.

Suppose a particle in x, y, z space moves around and around a vertical circular cylinder in such a way that its height (i.e., its z coordinate) steadily increases. In other words, the particle follows a spiral path that looks like a "corkscrew" or a "slinky." This path is a space curve known as a **helix**. A simple example would be the curve $x = 3\cos t,\ y = 3\sin t,\ z = t$, which lies on the cylinder $x^2 + y^2 = 9$. Its starting point is $(3,0,0)$. Its domain is $(-\infty, \infty)$, assuming we want the helix to extend infinitely far both up and down. However, if we want the helix to go no lower than the x, y plane, then the domain would be $[0, \infty)$.

A **spherical curve** is a space curve that lies on a sphere. A simple example would be the curve $x = \frac{\cos t}{\sqrt{1+t^2}},\ y = \frac{\sin t}{\sqrt{1+t^2}},\ z = \frac{t}{\sqrt{1+t^2}}$, which lies on the sphere $x^2 + y^2 + z^2 = 1$. Its starting point is $(1,0,0)$. Its domain is $(-\infty, \infty)$.

A space curve may be defined as the intersection of two given surfaces. For example, the intersection of the circular cylinder $x^2 + y^2 = 1$ and the plane $y + z = 2$ gives us a **slanted ellipse**. The orthogonal projection of this ellipse onto the x, y plane is the circle $x^2 + y^2 = 1$. We already know that this circle can be parameterized as $x = \cos t,\ y = \sin t$. Since the equation of the plane can be written $z = 2 - y$, the ellipse can be parameterized as $x = \cos t$, $y = \sin t$, $z = 2 - \sin t$. The domain is $[0, 2\pi]$ (assuming we want our particle to make one complete revolution). The initial point is $(1, 0, 2)$. The direction is counter-clockwise when viewed from above (but clockwise when viewed from below). See the illustration on page 850 of your text.

The parametric equations for a curve can be combined into a single vector equation. Let $\mathbf{r}(t)$ be the position vector for the point P_t.

- In two-dimensional space, $\mathbf{r}(t) = < x(t), y(t) >$, or $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$.
- In three-dimensional space, $\mathbf{r}(t) = < x(t), y(t), z(t) >$, or $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j} + z(t)\mathbf{k}$.

The parametric functions $x(t)$, $y(t)$, and (in the three-dimensional case) $z(t)$ are referred to as the **component functions** of $\mathbf{r}(t)$.

The plane curve $y = \tan x$ would have the vector equation $\mathbf{r}(t) = < t, \tan t >$.

The circle $x^2 + y^2 = 9$ could have the vector equation $\mathbf{r}(t) = < 3\cos t, 3\sin t >$ or $\mathbf{r}(t) = < -3\cos t, 3\sin t >$ or $\mathbf{r}(t) = < 3\cos t, -3\sin t >$ or $\mathbf{r}(t) = < 3\sin t, 3\cos t >$, depending on which parameterization we adopt. We could also write these equations as follows:

- $\mathbf{r}(t) = 3 < \cos t, \sin t >$
- $\mathbf{r}(t) = 3 < -\cos t, \sin t >$ or $\mathbf{r}(t) = -3 < \cos t, -\sin t >$
- $\mathbf{r}(t) = 3 < \cos t, -\sin t >$ or $\mathbf{r}(t) = -3 < -\cos t, \sin t >$
- $\mathbf{r}(t) = 3 < \sin t, \cos t >$

The parabola $y = -2x^2 + 5$ could have the vector equation $\mathbf{r}(t) = t\,\mathbf{i} + (-2t^2 + 5)\mathbf{j}$ or $\mathbf{r}(t) = -t\,\mathbf{i} + (-2t^2 + 5)\mathbf{j}$ or $\mathbf{r}(t) = (t + 4)\mathbf{i} + (-2t^2 - 16t - 27)\mathbf{j}$, depending on which parameterization we adopt.

The helix discussed above would have the vector equation $\mathbf{r}(t) = < 3\cos t, 3\sin t, t >$.

The spherical curve discussed above would have the vector equation $\mathbf{r}(t) = \frac{\cos t}{\sqrt{1+t^2}}\mathbf{i} + \frac{\sin t}{\sqrt{1+t^2}}\mathbf{j} + \frac{t}{\sqrt{1+t^2}}\mathbf{k}$.

The slanted ellipse discussed above would have the vector equation $\mathbf{r}(t) = < \cos t, \sin t, 2 - \sin t >$.

The function $\mathbf{r}(t)$ is a **vector-valued function** of one real variable (or parameter), t. It is known as a **position function** for the curve, because it gives us the position of a moving particle at any instant. (Bear in mind, this function depends upon the chosen parameterization for the curve. A given curve has many possible parameterizations and hence many possible position functions.)

The magnitude of $\mathbf{r}(t)$, denoted $|\mathbf{r}(t)|$ or $r(t)$, is a **scalar-valued function**. It gives us the particle's **distance from the origin** at any instant.

- For a plane curve, $r(t) = \sqrt{x(t)^2 + y(t)^2}$

- For a space curve, $r(t) = \sqrt{x(t)^2 + y(t)^2 + z(t)^2}$

For a space curve, if the value of $r(t)$ is constant for all t, then the curve is a spherical curve centered at the origin, and the constant value of $r(t)$ is the radius of the sphere.

2. Limits And Continuity for Position Functions:

The following discussion will focus on two-dimensional position functions (i.e., position functions for plane curves), but all the concepts discussed can also be applied to three-dimensional position functions (i.e., position functions for space curves).

In Calculus I, we learned the basic concept of the limit. If we have a function $y = f(x)$, we can ask, does y approach any particular value as x approaches some specified value, such as a. If it does, we say the function has a **limiting value** (or just a **limit**, for short) as x approaches a. Suppose we have such a value. Call it L. We can say, "y approaches L as x approaches a," which can be written more compactly as follows: $y \to L$ as $x \to a$. We can also write $\lim_{x \to a} y = L$, or $\lim_{x \to a} f(x) = L$, which would be pronounced, "The limit to the function $f(x)$ as x approaches a is L." For example, $\lim_{x \to 0} \frac{1}{x} \sin x = 1$.

In Calculus I, we were dealing with functions that produced *numerical values* (i.e., for any given numerical value of x, the function $y = f(x)$ produces a numerical value of y). But now, in Calculus III, we are dealing with functions that produce *vector values*. In other words, the vector equation of a plane curve, $\mathbf{r}(t) = < x(t), y(t) >$, may be thought of as a function whose *input* is the *scalar* (or real number) t, and whose *output* is the *vector* $\mathbf{r}(t)$. For instance, given the numerical value $t = 2$, the function $\mathbf{r}(t) = < t^3, \frac{1}{t} >$ produces the vector $< 8, \frac{1}{2} >$. Can we apply the concept of the limit to such functions? We can! Here is how...

Given the vector-valued function $\mathbf{r}(t) = < x(t), y(t) >$, we can ask, does $\mathbf{r}(t)$ approach any particular *vector value* as t approaches some specified *numerical value*, such as a. If it does, we say the function has a **limiting** (**vector**) **value** (or just a **limit**, for short) as t approaches a. Suppose we have such a vector value. Call it $\mathbf{L}$. We can say, "$\mathbf{r}(t)$ approaches $\mathbf{L}$ as t approaches a," which can be written more compactly as follows: $\mathbf{r}(t) \to \mathbf{L}$ as $t \to a$. We can also write $\lim_{t \to a} \mathbf{r}(t) = \mathbf{L}$, which would be pronounced, "The limit to the function $\mathbf{r}(t)$ as t approaches a is $\mathbf{L}$."

That's the basic idea. Now how do we go about finding this sort of limit? We use the following principle...

- If $\mathbf{r}(t) = < x(t), y(t) >$, then $\lim_{t \to a} \mathbf{r}(t) = < \lim_{t \to a} x(t), \lim_{t \to a} y(t) >$
- Equivalently, if $\mathbf{r}(t) = x(t)\mathbf{i} + y(t)\mathbf{j}$, then
  $\lim_{t \to a} \mathbf{r}(t) = \lim_{t \to a} x(t)\mathbf{i} + \lim_{t \to a} y(t)\mathbf{j}$

Essentially, this says the limit distributes in the same way as scalar multiplication–recall that $c\mathbf{r}(t) = < cx(t), cy(t) >$.

Each of the limits on the right side of the equation can be evaluated using the methods of Calculus I.

Suppose $\mathbf{r}(t) = \frac{1}{t}\sin t\ \mathbf{i} + \frac{t^2+6t}{t^2-2t}\ \mathbf{j}$. Find $\lim_{t\to 0} \mathbf{r}(t)$.
Solution: $\lim_{t\to 0} \mathbf{r}(t) = \lim_{t\to 0} \frac{1}{t}\sin t\ \mathbf{i} + \lim_{t\to 0} \frac{t^2+6t}{t^2-2t}\ \mathbf{j} = \mathbf{i} - 3\mathbf{j}$

As we learned in Calculus I, some limits do not exist. For instance, $\lim_{x\to 0} \sin\frac{1}{x}$ does not exist, due to infinite oscillation. A limit "not existing" includes the possibility that the function could approach infinity. For instance, $\lim_{x\to 0} \frac{1}{x^2}$ does not exist, because the function approaches infinity as x approaches 0. We may write $\lim_{x\to 0} \frac{1}{x^2} = \infty$, but it is still the case that the limit does not exist! (Saying the limit "exists" means the function approaches a unique real number value, and ∞ is not a real number.)

Likewise, in this new situation, a limit may or may not exist. In order for $\lim_{t\to a} \mathbf{r}(t)$ to exist, *both* limits on the right side of the equation must exist. In other words, $\lim_{t\to a} \mathbf{r}(t)$ exists if and only if *both* of the following limits exist:
- $\lim_{t\to a} x(t)$
- $\lim_{t\to a} y(t)$

If *either* of these does not exist, then $\lim_{t\to a} \mathbf{r}(t)$ does not exist.

For instance, suppose $\mathbf{r}(t) = < 5t+2, \frac{1}{t-3} >$. $\lim_{t\to 3} \mathbf{r}(t)$ does not exist, because $\lim_{t\to 3} \frac{1}{t-3}$ does not exist.

In Calculus I, we learned the concept of **continuity**. If we have a function $y = f(x)$, and if we have a specified value of x, such as a, we can ask whether or not the function is **continuous** at a. In order for the function to be continuous at a, *all three* of the following conditions must be met:
1. $f(a)$ must be defined. In other words, a must be in the domain of f.
2. $\lim_{x\to a} f(x)$ must exist.
3. $\lim_{x\to a} f(x)$ must be equal to $f(a)$, i.e., $\lim_{x\to a} f(x) = f(a)$.

If *any* of these three conditions is not met, then the function is *not continous* (or is **discontinuous**) at a. In this case, we may say the function has a **discontinuity** at a.

Books or teachers may sometimes cite only the third condition listed above. Their thinking is that saying $\lim_{x\to a} f(x) = f(a)$ presupposes both condition #1 and condition #2. However, I believe it is best to think of it as three separate conditions, and to check them in the order I have specified. First check condition #1; if it fails, go no further. If condition #1 is met, then check condition #2; if it fails, go no further. If condition #2 is met, then check condition #3.

If we know in advance (based on some previously established theorem) that the function f is continuous at a value a, then we can evaluate $\lim_{x\to a} f(x)$ by simple "plug and chug," i.e., by simply evaluating $f(a)$. For instance, we have a theorem that says a polynomial function is continuous for all real values of x. Hence, to evaluate $\lim_{x\to 5} (3x^2 - 7x + 4)$, we just plug in 5 for x, giving us 44. But be careful. Plug and chug does not work when the function is not continuous! For instance, you cannot evaluate $\lim_{t\to 0} \frac{1}{t}\sin t$ by plug and chug.

The concept of continuity can be applied to vector-valued functions. The function $\mathbf{r}(t) = < x(t), y(t) >$ is **continuous** at $t = a$ if and only if *all three* of the following conditions are met:

1. $\mathbf{r}(a)$ must be defined, which means $x(a)$ and $y(a)$ must both be defined. In other words, a must be in the domain of each function.
2. $\lim_{t \to a} \mathbf{r}(t)$ must exist, which means $\lim_{t \to a} x(t)$ and $\lim_{t \to a} y(t)$ must both exist.
3. $\lim_{t \to a} \mathbf{r}(t)$ must equal $\mathbf{r}(a)$, which means $\lim_{t \to a} x(t) = x(a)$ and $\lim_{t \to a} y(t) = y(a)$.

If *any* of these three conditions is not met, then the function is *not continous* (or is **discontinuous**) at a. In this case, we may say the function has a **discontinuity** at a.

Here is a three-dimensional example: The function $\mathbf{r}(t) = (1 + t^3)\mathbf{i} + te^{-t}\,\mathbf{j} + \frac{\sin t}{t}\mathbf{k}$ has a discontinuity at 0. $\lim_{t \to 0} \mathbf{r}(t) = \mathbf{i} + \mathbf{k}$.