Chapter 3

Parametric Curves

This chapter is concerned with the parametric approach to curves. The definition of a parametric curve is defined in Section 1 where several examples explaining how it differs from a geometric one are present. In Section 2 we introduce the arc-length for parametric curve and also the arc-length parametrization. In Section 3 two most common parametrization, namely, graphs and polar forms, are discussed. In Section 4 the signed curvature and curvature of a plane curve are defined using the arc-length parametrization. Finally, we illustrate the interplay between different parametrization, Kepler's First Law on planetary motion is derived In Section 5.

3.1 Parametric Curves

In this section we present an approach to geometric objects in $\mathbb{R}^n$ different from the zero set approach in the previous chapter. In this parametric approach, geometric objects such as curves and surfaces are regarded as mappings into $\mathbb{R}^n$. We will exclusively deal with curves here even though the idea extends to surfaces and beyond. In the case of curves it is partly motivated by physics where a curve is regarded as the trajectory of a particle in motion.

A continuous map $\boldsymbol{\gamma}$ from an interval I, open or closed, to $\mathbb{R}^n$ is called a **parametric curve.** By a continuous map we mean each component of the map

$$\boldsymbol{\gamma}(t) = (\gamma_1(t), \gamma_2(t), \cdots, \gamma_n(t)) : \ I \to \mathbb{R}^n \ ,$$

is continuous on I. It is called differentiable or continuously differentiable according to whether $\gamma_j', j = 1, 2, \cdots, n$, exist or are continuous. At this point a curve is not different from a continuous vector-valued function. An essential difference comes in when we define a regular curve. Indeed, a parametric curve $\boldsymbol{\gamma}$ is a **regular parametric curve** or simply a **regular curve** if it is continuously differentiable and $|\boldsymbol{\gamma}'(t)| > 0$ for all $t \in I$. Later we will encounter some parametric curve whose tangent does not vanish except at finitely many points. Strictly speaking these are not regular curves. However, they share many properties with regular curves. Note that

$$\boldsymbol{\gamma}'(t) = (\gamma_1'(t), \gamma_2'(t), \cdots, \gamma_n'(t)) ,$$

and

$$|\boldsymbol{\gamma}'(t)| = \sqrt{\gamma_1'^2(t) + \gamma_2'^2(t) + \cdots + \gamma_n'^2(t)} .$$

Example 3.1. Determine which of the following parametric curves defines a regular curve in the plane:

(a) $\boldsymbol{\gamma}(t) = (t^2+1, t\sin t) , \quad t \in (1, \pi) ,$

(b) $\boldsymbol{\eta}(t) = (t^2+1, t\sin t) , \quad t \in [-1, , \pi] ,$

(c) $\mathbf{c}(\theta) = (2\cos\theta, 3\sin\theta) , \quad \theta \in [0, 4\pi] ,$

(d) $\boldsymbol{\xi}(z) = (\sin z^{-1}, 3z+1) , \quad \boldsymbol{\xi}(0) = 0, \quad z \in [0, \infty) .$

(a) $\boldsymbol{\gamma}'(t) = (2t, \sin t + t\cos t)$ and

$$|\boldsymbol{\gamma}'(t)| = \sqrt{4t^2 + (\sin t + t\cos t)^2} \geq 2t > 0 , \quad \forall t \in (1, \pi) .$$

So $\boldsymbol{\gamma}$ is a regular curve.

(b) $\boldsymbol{\eta}$ and $\boldsymbol{\gamma}$ share the same formula but on different domains. They are different curves. At $t = 0$, $|\boldsymbol{\eta}'(0)| = 0$, so it is not a regular curve.

(c) $\mathbf{c}'(\theta) = (-2\sin\theta, 3\cos\theta)$, and

$$|\mathbf{c}'(\theta)| = \sqrt{4\sin^2\theta + 9\cos^2\theta} \geq \sqrt{4\sin^2\theta + 4\cos^2\theta} \geq 2,$$

so $\mathbf{c}$ is a regular curve. In fact, from

$$\frac{c_1^2(\theta)}{4} + \frac{c_2^2(\theta)}{9} = 1$$

we see that the image of the parametric curve is an ellipse. As θ runs from 0 to 4π, the trajectory covers the ellipse twice.

(d) Observe that the map $z \mapsto \sin z^{-1}$ does not have a limit at $z = 0$, $\boldsymbol{\xi}$ is not a continuous map. Therefore, it is not a parametric curve, let alone a regular one. Keep in mind that a parametric curve is always continuous in each of its components.

It is natural to call $\boldsymbol{\gamma}'(t)$ the **tangent** or **tangent vector** of the parametric curve $\boldsymbol{\gamma}$ at t and view it as a vector based at $\boldsymbol{\gamma}(t)$. The **tangent line** of $\boldsymbol{\gamma}$ at $\boldsymbol{\gamma}(t_0)$ is the straight line passing through $\boldsymbol{\gamma}(t_0)$ along the direction determined by the vector $\boldsymbol{\gamma}'(t_0)$, that is, it is given by

$$\{\boldsymbol{\gamma}(t_0) + t\boldsymbol{\gamma}'(t_0) : \ t \in \mathbb{R}\} \ .$$

(Here t_0 indicates the fixed time and t is the variable for the tangent line.)

When $\boldsymbol{\gamma}$ describes the motion of a particle in the duration from time a to b, $\boldsymbol{\gamma}(t)$ is the location of the particle at time t. It is usually called the **position vector**. Accordingly the tangent vector is called the **velocity** or velocity vector. Its magnitude $|\boldsymbol{\gamma}'(t)|$ is called the **speed** at time t. It describes how fast the particle is moving at this instant. The unit vector $\boldsymbol{\gamma}'(t)/|\boldsymbol{\gamma}'(t)|$ points to the direction of the motion at this instant. Finally, the second derivative $\boldsymbol{\gamma}''(t)$ is called the **acceleration** at t. To use t to denote the parameter was motivated from this context and is adopted elsewhere even it no longer carries the meaning of time.

Example 3.2. Consider the map $\boldsymbol{\gamma} : \mathbb{R} \to \mathbb{R}^n$ defined by $\boldsymbol{\gamma}(t) = \mathbf{p} + t\boldsymbol{\xi}$ where $\mathbf{p}$ and $\boldsymbol{\xi} \neq (0, \cdots, 0) \in \mathbb{R}^n$ are given. Each component is a linear function and clearly differentiable up to any order. Its velocity, speed and acceleration are all independent of time and given by $\boldsymbol{\xi}$, $|\boldsymbol{\xi}|$ and $\mathbf{0}$ respectively. One should be cautious about the subtle change of point of view : Now the straight line is regarded as a map from $\mathbb{R}$ to $\mathbb{R}^n$. The straight line previously defined is actually the image of this map.

The followings are some commonly used parametric equations for ellipses, hyperbolas and parabolas:

- $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$; $\quad x = a\cos\theta, \quad y = b\sin\theta, \quad \theta \in [0, 2\pi)$, $a, b > 0$.
- $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$; $\quad x = a\cosh\theta, \quad y = \pm b\sinh\theta, \quad \theta \in (-\infty, \infty)$, $a, b > 0$.
- $y = ax^2 + bx + c$; $\quad y = at^2 + bt + c, \quad x = t, \quad t \in (-\infty, \infty)$, $a, b, c \in \mathbb{R}$.

Other parametric equations are also available, for instance, for the hyperbola one may use

$$x = \frac{a}{2}\Big(t + \frac{1}{t}\Big), \quad y = \frac{b}{2}\Big(t - \frac{1}{t}\Big) , \quad t \in (0, \infty) \text{ or } t \in (-\infty, 0) \ .$$

Example 3.3. A bee is flying along the trajectory described by the helix

$$(x(t), y(t), z(t)) = (\cos t, \sin t, t) \ , \quad t \in [0, \infty) \ .$$

(a) Find its velocity, speed and acceleration at $t = \pi/2$ and at π.

(b) Find its tangent line at $t = \pi/2$.

(a) Let $\boldsymbol{\gamma}$ be this trajectory. Clearly it is a regular parametric curve in $\mathbb{R}^3$. We have

$$\boldsymbol{\gamma}'(t) = (-\sin t, \cos t, 1) \ , \quad \boldsymbol{\gamma}''(t) = (-\cos t, -\sin t, 0) \ .$$

So its velocity and acceleration at $t = \pi/2, \pi$ are given respectively by

$$\boldsymbol{\gamma}'\left(\frac{\pi}{2}\right) = (-1, 0, 1) \ , \quad \boldsymbol{\gamma}''\left(\frac{\pi}{2}\right) = (0, -1, 0) \ ,$$

and

$$\boldsymbol{\gamma}'(\pi) = (0, -1, 1) \ , \quad \boldsymbol{\gamma}''(\pi) = (1, 0, 0) \ .$$

We can also determine its speed

$$|\boldsymbol{\gamma}'(t)| = \sqrt{(-\sin t)^2 + \cos^2 t + 1} = \sqrt{2} \ ,$$

which is independent of time. The bee is flying at constant speed.

(b) The tangent line passing through $\boldsymbol{\gamma}(\pi/2) = (0, 1, \pi/2)$ is given by

$$\left(0, 1, \frac{\pi}{2}\right) + (-1, 0, 1)t \ , \quad t \in \mathbb{R}.$$

Example 3.4. Find all tangent lines of the ellipse

$$x^2 + \frac{y^2}{9} = 1$$

that passes through the point $(1, -6)$. Choose the parametrization of the ellipse to be

$$x = \cos t, \quad y = 3\sin t, \quad t \in [0, 2\pi) \ .$$

As t runs from 0 to 2π, the particle travels from $(1, 0)$ along the counterclockwise direction and back to $(1, 0)$ at $t = 2\pi$. The tangent vector at (x, y) is given by

$$(x', y') = (-\sin t, 3\cos t),$$

which shows this is a regular curve. The tangent line passing through (x, y) is given by

$$(\cos t_0, 3\sin t_0) + t(-\sin t_0, 3\cos t_0), \ t \in \mathbb{R} \ .$$

In order the line passing through $(1, -6)$, we require there is some t so that $(\cos t_0, 3\sin t_0)+ t(-\sin t_0, 3\cos t_0) = (1, -6)$, that is,

$$\cos t_0 - \sin t_0 \ t = 1, \quad 3\sin t_0 + 3\cos t_0 \ t = -6.$$

This system is readily solved to give

$$\cos t_0 - 2\sin t_0 = 1, \quad t = -2\cos t_0 - \sin t_0 \ .$$

Using the compound angle formula, the first equation can be written as

$$\cos(t_0+\tau) = \frac{\sqrt{5}}{5} \ ,$$

where τ satisfies

$$\cos\tau = \frac{\sqrt{5}}{5}, \quad \sin\tau = \frac{2\sqrt{5}}{5}, \ \tau \in (0, \frac{\pi}{4}) \ .$$

There are two t_1, t_2's lying respectively in $t_1 \in (0, \pi/2)$ and $t_2 \in (3\pi/2, 2\pi)$ satisfying $\cos t_1 = \cos t_2 = \frac{\sqrt{5}}{5}$. Indeed, we conclude that both $t_0 = 0$ and $t_0 = 2\pi - 2\tau$ satisfy the requirement. Then there are exactly two tangent lines of the ellipse passing through $(1, -6)$, one emitting from $(1, 0)$ and the other from $(\cos(2\tau), -3\sin(2\tau))$.

Newton's second law asserts that $m\mathbf{a} = \mathbf{F}$ where m is the mass of the particle, $\mathbf{a}$ its acceleration and $\mathbf{F}$ the external force. The last two are 3-vectors. In case the force only depends on the position of the particle which happens, for instance, under the influence of a gravitational field, the time t appears as the natural parameter of the motion. The law can be written as a second order differential equation

$$m\frac{d^2\mathbf{r}}{dt^2} = F(\mathbf{r}) \ ,$$

where $\mathbf{r}(t) = (x(t), y(t), z(t))$ is the position vector.

Example 3.5. Determine the motion of the projectile with initial position (x_0, y_0, z_0) and initial velocity (u_0, v_0, w_0). The motion is governed by the gravity only, so the second law reads as

$$m\frac{d^2\mathbf{r}}{dt^2} = -mg\mathbf{e}_3 \ , \quad \mathbf{r}(t) = (x(t), y(t), z(t)) \ .$$

Looking at the components separately, we have

$$\begin{cases} m\dfrac{d^2x}{dt^2} = 0 \ , \\[2ex] m\dfrac{d^2y}{dt^2} = 0 \ , \\[2ex] m\dfrac{d^2z}{dt^2} = -mg \ . \end{cases}$$

All three equations are of the form $f''(t) = c$ for some constant c. A single integration yields $f'(t) = ct + c_1$ for some constant c_1 and a further integration yields $f(t) = \frac{c}{2}t^2 + c_1 t +$

c_2 for another constant c_2. Therefore, the general solution of this system of differentiable equations is given by

$$\mathbf{r}(t) = (x(t), y(t), z(t)) = \left(a + bt, c + dt, e + ft - \frac{1}{2}gt^2 \right) ,$$

for arbitrary constants a, b, c, d, e, f. Plugging in $t = 0$, we get $(a, c, e) = (x_0, y_0, z_0)$. Plugging $t = 0$ in $r'(t)$ we get $(b, c, f) = (u_0, v_0, w_0)$. Therefore, the motion of the projectile is given by

$$\mathbf{r}(t) = (x(t), y(t), z(t)) = \left(x_0 + u_0 t, y_0 + v_0 t, z_0 + w_0 t - \frac{1}{2}gt^2 \right) .$$

Once the initial position and velocity are known, its position and velocity in any time is completely determined no matter it is in the future or back to the ancient times. This is the deterministic feature of Newton's world.

To end this section, let us once again point out that a parametric curve is different from a geometric curve. Different parametric curves could have the same image, that is, they represent the same geometric curve, so each one of them carries additional information.

Example 3.6. Describe the parametric curves defined by, for $a, b > 0$,

(a) $\boldsymbol{\gamma}_1(t) = (a \cos t, b \sin t)$, $\quad t \in [0, 2\pi]$,

(b) $\boldsymbol{\gamma}_2(t) = (a \cos t^2, b \sin t^2)$, $\quad t \in [0, \sqrt{2\pi}]$,

(c) $\boldsymbol{\gamma}_3(t) = (a \cos t, -b \sin t)$, $t \in [0, 2\pi]$,

(d) $\boldsymbol{\gamma}_4(t) = (a \cos t, b \sin t)$, $t \in [0, 4\pi]$

The image of all these four parametric curves have the same image, namely, the ellipse given by

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

It shows that a parametric curve contains more information such as orientation, velocity, and multiplicity than a geometric curve does. For instance, the curve in (a) describes the motion of a particle moving along the ellipse starting from $(a, 0)$ and going back to this point at $t = 2\pi$ in the counterclockwise direction. Its velocity at time t is given by $(a^2 \sin^2 t + b^2 \cos^2 t)^{1/2}$. The second curve moves similar to the first curve, but not uniform in speed: $2t(a^2 \sin^2 t + b^2 \cos^2 t)^{1/2}$. The third curve describes the motion going from $(a, 0)$ to itself in clockwise direction. In the fourth curve the particle travels the ellipse twice counterclockwisely as t goes from 0 to 4π. $\square$

3.2 The Arc-Length

In this section we define the length of a parametric curve. In fact, let $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ be a regular parametric curve. As t runs from a to b, $\boldsymbol{\gamma}(t)$ goes from $\boldsymbol{\gamma}(a)$ to $\boldsymbol{\gamma}(b)$ and never turns back. Introducing partition points $a = t_1 < t_2 < \cdots < t_m = b$, the sums

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

give a good approximation to the length of the curve as the partition getting refined. By the mean-value theorem the sums beocme

$$\sum_{j=1}^{m-1} |\boldsymbol{\gamma}'(t_j^*)|(t_{j+1} - t_j)$$

for some mean value $t_j^* \in (t_j, t_{j+1})$. As $t_{j+1} - t_j \to 0$, these sums tend to

$$\int_a^b |\boldsymbol{\gamma}'(t)|dt \ .$$

It is reasonable to define the **length** of the parametric curve $\boldsymbol{\gamma}$ to be

$$L = \int_a^b |\boldsymbol{\gamma}'(t)|dt \ .$$

Example 3.7. Find the length of the following two parametric curves:

(a) $\boldsymbol{\gamma}(t) = (\cos^3 t, \sin^3 t) \ , \quad t \in [0, \pi/2] \ .$

(b) $\boldsymbol{\xi}(t) = (\cos t, \sin t, t) \ , \quad t \in [0, T] \ .$

For (a), we have $\boldsymbol{\gamma}'(t) = (-3\cos^2 t \sin t, 3\sin^2 t \cos t)$ and

$$|\boldsymbol{\gamma}'(t)| = \sqrt{9\cos^4 t \sin^2 t + 9\sin^4 t \cos^2 t} = 3\cos t \sin t \ .$$

The length over $[0, \pi/2]$ is given by

$$\begin{aligned} \int_0^{\pi/2} |\boldsymbol{\gamma}'(t)|dt &= \int_0^{\pi/2} 3\cos t \sin t dt \\ &= \frac{3}{2} \ . \end{aligned}$$

We remark that the curve $\boldsymbol{\gamma}$ is not regular as its velocity vanishes at 0 and $\pi/2$. However, this does no harm as we could replace the endpoints by some $a, b, a < b$, in $(0, \pi/2)$ and then let $a \to 0^+$ and $b \to \pi/2^-$. For (b), we have

$$|\boldsymbol{\xi}'(t)| = \sqrt{(-\sin t)^2 + \cos^2 t + 1} = \sqrt{2} \ .$$

Therefore, the length over $[0, T]$ is

$$\int_0^T |\boldsymbol{\xi}'(t)| dt = \int_0^T \sqrt{2} dt = \sqrt{2} T \ .$$

In Section 3.1 we have seen that the same geometric curve can be described by many different parametric curves. In this section we will show that there is an intrinsic one that is particularly useful in many situations. This is the arc-length parametrization.

We need some preparation. Let $\boldsymbol{\gamma}$ be a regular parametric curve on $[a, b]$ and $\varphi : [c, d] \to [a, b]$ is continuously differentiable function satisfying $\varphi'(t) > 0$. Then $\tilde{\boldsymbol{\gamma}}(t) = \boldsymbol{\gamma}(\varphi(t))$ is a regular parametric curve from $[c, d]$ to $\mathbb{R}^n$ whose image is the same as the image of $\boldsymbol{\gamma}$. This new parametric curve is called a **reparametrization** of $\boldsymbol{\gamma}$. In the following we would like to show that every regular curve admits a special reparametrization whose speed is constant.

Note that $|\boldsymbol{\gamma}'(t)|$ is integrable as it is a continuous function on $[a, b]$ by assumption. To show that the notion of the length is a geometric one we will show that it is independent of reparametrization.

Let us recall the change of variables formula in calculus. Let $F : [a, b] \to \mathbb{R}$ and $\varphi : [c, d] \to [a, b]$ be continuously differentiable. We have the change of variable formula:

$$\int_c^t F'(\varphi(t))\varphi'(t) dt = \int_{\varphi(c)}^{\varphi(t)} F'(z) dz \ .$$

(Here is the proof: by the Chain Rule,

$$\frac{dF \circ \varphi}{dt} = F'(\varphi(t))\varphi'(t) \ .$$

Integrating this relation and using the fundamental theorem of calculus,

$$\int_c^t F'(\varphi(t))\varphi'(t) dt = F(\varphi(t)) - F(\varphi(c)) \ .$$

On the other hand, we have

$$\int_{\varphi(c)}^{\varphi(t)} F'(z)dz = F(\varphi(t)) - F(\varphi(c)) ,$$

and the formula follows by combining these two relations.)

Theorem 3.1. *Let $\tilde{\boldsymbol{\gamma}}(t) = \boldsymbol{\gamma}(\varphi(t))$ be a reparametrization of $\boldsymbol{\gamma}$ on $[c,d]$ satisfying $\varphi(c) = a, \varphi(d) = b$. Then*

$$\int_c^d |\tilde{\boldsymbol{\gamma}}'(t)|dt = \int_a^b |\boldsymbol{\gamma}'(t)|dt .$$

Proof. Taking F to be $\tilde{\boldsymbol{\gamma}}$ in the change of variables formula,

$$\begin{aligned}\int_c^d |\tilde{\boldsymbol{\gamma}}'(t)|dt &= \int_c^d |\boldsymbol{\gamma}'(\varphi(t))|\varphi'(t)dt \\ &= \int_{\varphi(c)}^{\varphi(d)} |\boldsymbol{\gamma}'(z)|dz \\ &= \int_a^b |\boldsymbol{\gamma}'(z)|dz ,\end{aligned}$$

where the change of variables formula has been applied in the second equality. □

The **length map** of a regular parametric curve $\boldsymbol{\gamma}$ is $\Lambda : [a,b] \to [0,L]$, where

$$\Lambda(t) = \int_a^t |\boldsymbol{\gamma}'(z)|dz.$$

It is strictly increasing and continuously differentiable. Its inverse map Φ is also continuously differentiable and maps $[0,L]$ to $[a,b]$ such that $\Phi(\Lambda(t)) = t$ for all $t \in [a,b]$ and $\Lambda(\Phi(s)) = s$ for all $s \in [0,L]$.

Theorem 3.2. *For any regular parametric curve $\boldsymbol{\gamma}$ on $[a,b]$, the reparametrization $\tilde{\boldsymbol{\gamma}}(s) \equiv \boldsymbol{\gamma}(\Phi(s))$ defined on $[0,L]$ satisfies*

$$|\tilde{\boldsymbol{\gamma}}'(s)| = 1 , \quad \forall s \in [0,L] .$$

Proof. Letting $t = \Phi(s)$, we differentiate the relation $\Phi(\Lambda(t)) = t$ to get

$$\frac{d\Phi}{ds}(s)\frac{d\Lambda}{dt}(t) = 1.$$

Now,

$$\begin{aligned}\tilde{\boldsymbol{\gamma}}'(s) &= \frac{d\boldsymbol{\gamma}\circ\Phi}{ds}(s) \\ &= \boldsymbol{\gamma}'(t)\frac{d\Phi}{ds}(s) \\ &= \boldsymbol{\gamma}'(t)\frac{1}{\Lambda'(t)} \\ &= \frac{\boldsymbol{\gamma}'(t)}{|\boldsymbol{\gamma}'(t)|} ,\end{aligned}$$

where in the last step we have used the definition of Λ. Therefore, $|\tilde{\boldsymbol{\gamma}}'(s)| = 1$ for all s.

□

Now, suppose we have two regular curves γ_1 and γ_2 on $[a, b]$ and $[c, d]$ respectively which satisfy $\gamma_1(a) = \gamma_2(c)$ and $\gamma_1(b) = \gamma_2(d)$, that is, their endpoints coincide. Since both curves map their intervals of definition bijectively into its image, the composite function $\varphi(t) = \gamma_2^{-1}(\gamma_1(t))$ maps $[a, b]$ to $[c, d]$ with $\varphi(a) = c$ and $\varphi(b) = d$. Using $|\gamma_i'| > 0, i = 1, 2$, it can be shown that $\varphi' > 0$ on $[a, b]$ and hence $\gamma_1(t) = \gamma_2(\varphi(t))$, that is, γ_1 is a reparametrization of γ_2. Using these two theorems we know that no matter γ_1 or γ_2 is used to obtain the arc-length parametrization, the result is the same.

Example 3.8. Consider two parametrizations of the circle of radius R centered at the origin

(a)
$$\boldsymbol{\gamma}_1(t) = (R\cos t^2, R\sin t^2) , \quad t \in [0, (2\pi)^{1/2}] ,$$

(b)
$$\boldsymbol{\gamma}_2(t) = (R\cos t^3, R\sin t^3) , \quad t \in [0, (2\pi)^{1/3}] .$$

Please find their arc-length reparametrization formula.

Both parametric curves describe the circle traveling once in the counterclockwise direction. Now,

$$\begin{aligned}\Lambda_1(t) &= \int_0^t |\boldsymbol{\gamma}_1'(z)|dz \\ &= \int_0^t \sqrt{4R^2z^2\sin^2 z^2 + 4R^2z^2\cos^2 z^2}dz \\ &= \int_0^t 2Rzdz \\ &= Rt^2 ,\end{aligned}$$

so

$$L_1 = \int_0^{\sqrt{2\pi}} |\boldsymbol{\gamma}_1'(z)|dz = 2\pi R \ .$$

The inverse to Λ_1 is given by $\Phi_1(s) = (s/R)^{1/2}$. Therefor, the arc-length parametric curve is given by

$$\tilde{\boldsymbol{\gamma}}_1(s) = \boldsymbol{\gamma}_1(\Phi_1(s)) = (R\cos\frac{s}{R}, R\sin\frac{s}{R}) \ , \quad s \in [0, 2\pi R] \ .$$

On the other hand,

$$\begin{aligned}
\Lambda_2(t) &= \int_0^t |\boldsymbol{\gamma}_2'(z)|dz \\
&= \int_0^t \sqrt{9Rz^4\sin^2 z^3 + 9Rz^4\cos^2 z^3}dz \\
&= \int_0^t 3Rz^2 dz \\
&= Rt^3 \ ,
\end{aligned}$$

so

$$L_2 = \int_0^{(2\pi)^{1/3}} |\boldsymbol{\gamma}_2'(z)|dz = 2\pi R \ .$$

The inverse to Λ_2 is given by $\Phi_2(s) = (s/R)^{1/3}$. Therefore, the arc-length parametric curve is given by

$$\tilde{\boldsymbol{\gamma}}_2(s) = \boldsymbol{\gamma}_2(\Phi_2(s)) = (R\cos\frac{s}{R}, R\sin\frac{s}{R}) \ , \quad s \in [0, 2\pi R] \ .$$

We see that $\tilde{\boldsymbol{\gamma}}_1$ and $\tilde{\boldsymbol{\gamma}}_2$ are equal, in other words, both parametric curves lead to the same arc-length parametric curve.

A side remark. In this example both $\boldsymbol{\gamma}_1'(0)$ and $\boldsymbol{\gamma}_2'(0)$ are equal to zero, so strictly speaking they are not regular curves. However, we could first restrict to $[t_0, 2\pi], t_0 > 0$, and then let $t_0 \downarrow 0$ to get the result.

Example 3.9. Consider the regular curve

$$\mathbf{c}(t) = (a\cos t, b\sin t) \ , \quad t \in [0, 2\pi] \ .$$

The curve describes an ellipse traveling in counterclockwise direction. Its length function is given by

$$\begin{aligned}
\Lambda(t) &= \int_0^t |\boldsymbol{\gamma}'(z)|dz \\
&= \int_0^t \sqrt{a^2\sin^2 z + b^2\cos^2 z}dz \\
&= \int_0^t \sqrt{a^2 + (b^2 - a^2)\cos^2 z}dz \ ,
\end{aligned}$$

which is an elliptic integral. It is well-known that it cannot be integrated explicitly for most a and b. This example demonstrates a drawback of the arc-length parametrization, namely, it cannot have an explicit expression in many cases.

Given a regular curve $\boldsymbol{\gamma}$ on $[a, b]$, it is always possible to define another regular curve whose domain and image coincide with those of $\boldsymbol{\gamma}$ but moves in opposite direction. In fact, the curve $\boldsymbol{\gamma}_1(t) = \boldsymbol{\gamma}(-t)$ defined on $[-b, -a]$ has the same image with $\boldsymbol{\gamma}$. Thus the curve $\boldsymbol{\gamma}_2(t) = \boldsymbol{\gamma}_1(t+b+a)$ defined on $[a, b]$ is what we are after. Although these two curves share the same image and their total length are equal, the length functions are different because $\boldsymbol{\gamma}$ is measured starting from $\boldsymbol{\gamma}(a)$ while $\boldsymbol{\gamma}_2$ is measured from $\boldsymbol{\gamma}_2(a) = \boldsymbol{\gamma}(b)$. Furthermore, letting $\mathbf{p} = \boldsymbol{\gamma}(t_1) = \boldsymbol{\gamma}_2(t_2)$, we have $\boldsymbol{\gamma}'(t_1) = -\boldsymbol{\gamma}'(t_2)$. That is, their tangents at the same point are opposite to each other.

All in all, we have succeeded in showing that every regular curve admits a parametrization in which the tangent is always a unit vector. With further effort, one may even to show there are exactly two non-equivalent arc-parametrization whose directions are opposite to each other. In any case, from now on it is free of cost to assume that a regular curve is parametrized by arc-length. The same conclusion applies to those parametric curves whose tangents do not change sign but could become zero at finitely many points.

3.3 Graphs and Polar Forms

In this section we examine two most common parametrization, namely, graphs and polar forms.

Let f be a continuous function defined on some interval I. Its graph $\{(x, f(x)) : x \in I\}$ becomes a curve over I. It is natural to use x as the parameter to describe this curve:

$$x \mapsto (x, f(x)) \ , \quad x \in I \ .$$

Such parametrization is usually called the **non-parametric form**. From

$$|(x, f(x))'| = |(1, f'(x))| = \sqrt{1 + f'^2(x)} > 0 \ ,$$

we see that it is always a regular curve whenever it is differentiable. In the non-parametric form the tangent vector, length function, unit tangent, and normal are given respectively by

$$(1, f'(x)) \ , \quad \text{(the tangent vector)}$$

$$L = \int_I \sqrt{1 + f'^2(z)}dz \ , \quad \text{(the length)}$$

and

$$\mathbf{t} = \frac{(1, f'(x))}{\sqrt{1+f'^2(x)}} \ , \quad \mathbf{n} = \frac{(-f'(x), 1)}{\sqrt{1+f'^2(z)}} \ , \qquad \text{(the unit tangent and unit normal).}$$

The tangent and the normal adapt to the right hand rule.

A disadvantage of non-parametric form is that in many cases the curve cannot be written as the graph of some function. For example, the unit circle defined by the equation $x^2 + y^2 = 1$ cannot be expressed as a single function over $[-1, 1]$, instead it is the union of two functions:

$$f_1(x) = \sqrt{1-x^2} \ , \quad f_2(x) = -\sqrt{1-x^2} \quad x \in [-1, 1] \ .$$

One may express in another way:

$$g_1(y) = \sqrt{1-y^2} \ , \quad g_2(y) = -\sqrt{1-y^2} \quad y \in [-1, 1] \ .$$

(In fact,

$$f_1'(x) = \frac{-x}{\sqrt{1-x^2}} \ ,$$

becomes unbounded at $x = \pm 1$. To obtain a regular non-parametric curve one must restrict the interval $[-1, 1]$ to a smaller one, resulting at least four functions are needed to describe the circle.)

Example 3.10. Consider the parabola in the non-parametric form

$$f(x) = 3x^2 - 5x + 11 \ , \quad x \in \mathbb{R} \ .$$

Find its explicit arc-length function on the interval $[\frac{5}{6}, x]$ with $x \geq \frac{5}{6}$.

Sol: The parametric curve describing the parabola is given by

$$\boldsymbol{\gamma}(x) = (x, 3x^2 - 5x + 11) \ .$$

We have

$$\boldsymbol{\gamma}'(x) = (1, 6x - 5)$$

Then we know that arc-length is given by

$$L(x) = \int_{\frac{5}{6}}^{x} \sqrt{1 + (6t-5)^2} dt = \frac{1}{6} \int_0^{6x-5} \sqrt{1+z^2} dz.$$

Noting

$$\int \sqrt{z^2 + a^2} dz = \frac{1}{2} \left\{ z\sqrt{z^2+a^2} + a^2 \ln \left| z + \sqrt{z^2+a^2} \right| \right\} + C,$$

then we have

$$
\begin{aligned}
L(x) &= \frac{1}{6}\int_0^{6x-5} \sqrt{1+z^2}dz \\
&= \frac{1}{12}\left\{ z\sqrt{z^2+1} + \ln\left| z+\sqrt{z^2+1}\right| \right\}\Bigg|_0^{6x-5} \\
&= \frac{1}{12}\left\{ (6x-5)\sqrt{(6x-5)^2+1} + \ln\left| 6x-5+\sqrt{(6x-5)^2+1}\right| \right\}.
\end{aligned}
$$

Another useful parametrization is the so-called polar forms. Recall that the polar coordinates provides an alternative way to describe points on the plane other than the cartesian coordinates. There are two parameters, the distance r and angle θ. The r-axis (or initial ray) is a semi-infinite axis starting from a point (the origin O) and extending to infinity. A point on the plane is specified by $P(r,\theta)$ where r is its distance from P to O and $\theta \in [0, 2\pi)$, is measure from the r-axis to $\overline{OP}$. To set up the conversion between the polar coordinates and cartesian coordinates we identify the r-axis with the positive x-axis. Then we have the relations

$$
r = \sqrt{x^2+y^2}, \quad \cos\theta = \frac{x}{\sqrt{x^2+y^2}}, \quad \sin\theta = \frac{y}{\sqrt{x^2+y^2}}\ , \text{ and } \tan^{-1}\theta = \frac{y}{x},
$$

and

$$
x = r\cos\theta\ , \quad y = r\sin\theta\ .
$$

Now a polar curve (or a curve in polar form) is described by a single polar equation $\rho = \rho(\theta)$ where θ ranges over some interval I. More precisely, it means the parametric curve is

$$
\theta \mapsto \rho(\theta)(\cos\theta, \sin\theta)\ .
$$

Hence $|\rho(\theta)|$ is the distance of the curve to the origin. The tangent vector, length function, unit tangent, normal and signed curvature are given respectively by

$$
\rho(-\sin\theta, \cos\theta) + \rho'(\cos\theta, \sin\theta)\ , \qquad \text{(the tangent vector)}
$$

$$
L = \int_I \sqrt{\rho^2(\alpha)+\rho'^2(\alpha)}\ d\alpha\ , \qquad \text{(the length)}
$$

$$
\mathbf{t} = \frac{\rho(-\sin\theta, \cos\theta) + \rho'(\cos\theta, \sin\theta)}{(\rho^2+\rho'^2)^{1/2}}, \qquad \text{(the unit tangent)}
$$

and

$$
\mathbf{n} = \frac{\rho(-\cos\theta, -\sin\theta) + \rho'(-\sin\theta, \cos\theta)}{(\rho^2+\rho'^2)^{1/2}}\ , \qquad \text{(the unit normal).}
$$

Using the right hand rule, you can verify that $\mathbf{n}$ is the inner unit normal when the curve is a simple closed one such as circles and ellipses.

The function $\rho(\theta)$ could be negative for some values of θ. For instance, $\rho = 1 - 2\cos\theta$ becomes negative on $(-\pi/3, \pi/3)$ and consequently the curve $(1-2\cos\theta)(\cos\theta, \sin\theta), \theta \in [0, 2\pi]$ develops an inner loop in the second and third quadrants, see the example below.

Example 3.11. The standard form of an ellipse is

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1 \ ,$$

where a and $b, a > b$, are respectively its major and minor axis length. It is known that its focus points are $(c, 0), (-c, 0)$ where $c^2 = a^2 - b^2, c > 0$. Now we would like to write down the polar equation for the ellipse using the focus $(-c, 0)$ as the origin. We have

$$x + c = \rho\cos\theta, \quad y = \rho\sin\theta \ ,$$

where ρ is the distance from (x, y) to $(-c, 0)$. Plugging these relations into the equation for the ellipse, after a lengthy but straightforward calculation we obtain

$$\rho(\theta) = \frac{a(1-e^2)}{1 - e\cos\theta} \ ,$$

where

$$e = \sqrt{1 - \frac{b^2}{a^2}} \in [0, 1)$$

is the eccentricity of the ellipse. When $(c, 0)$ is taken as the origin instead of $(-c, 0)$, the polar equation becomes

$$\rho(\theta) = \frac{a(1-e^2)}{1 + e\cos\theta} \ .$$

It is interesting to note that these equations describe a hyperbola when $e > 1$. I let you verify these facts.

Example 3.12. Consider the limacçn

$$\rho = 1 + a\cos\theta \ , \quad a > 0.$$

This regular curve $\theta \mapsto \rho(\theta)(\cos\theta, \sin\theta)$ is of 2π period, so its domain can be taken to be $[0, 2\pi]$. We compute

$$\rho' = -a\sin\theta, \quad \rho'' = a\cos\theta, \quad \rho^2 + \rho'^2 = 1 + a^2 + 2a\cos\theta > 0 \ .$$

In particular, this is a regular curve unless $a = 1$. When $a \in (0, 1)$, ρ is always positive. By plotting its graph, you can see that it is a simple, (locally) convex curve containing the origin. When $a = 1$, the curve is called the cardioid due to its shape. Now the magnitude of its tangent

$$\rho^2 + \rho'^2 = 2 + 2\cos\theta$$

which vanishes at $\theta = -\pi$. The cardioid has a cusp at this point. When $a > 1$, ρ becomes negative somewhere and the limaçọn develops an inner loop. To describe it let us examine the typical case $a = 2/\sqrt{3}$. Now ρ is positive for $\theta \in (-2\pi/3, 2\pi/3)$ and negative for $\theta \in (2\pi/3, 4\pi/3)$. Henceforth, as θ goes from $-2\pi/3$ to $2\pi/3$, the curve runs from the origin and back to it through the fourth quadrant and the first quadrant to make a loop L_1. After passing $2\pi/3$, $\rho(\theta)$ becomes negative, so the curve goes again from the fourth quadrant to the first quadrant when θ ends at $4\pi/3$. Let the loop be L_2. From

$$\rho(\theta) = 1 + \frac{2}{\sqrt{3}}\cos\theta < 1 + \frac{2}{\sqrt{3}}\cos(\theta+\pi), \quad \theta \in \left(\frac{2\pi}{3}, \frac{4\pi}{3}\right) ,$$

we see that L_1 is the outer loop and L_2 is the inner one.

It is curious to see what the limacọn looks like in cartesian coordinates. In fact, we write the polar equation as $\rho - a\cos\theta = 1$ and multiply it with ρ to get

$$\rho^2 - a\rho\cos\theta = \rho \ .$$

Taking square, we arrive at the equation

$$(x^2 + y^2 - ax)^2 = x^2 + y^2 \ .$$

Thus the limacọn is actually described by a quartic equation.

More examples of polar curves are found in the exercise.

3.4 The Curvature of a Curve*

Among all possible parametrizations of the same geometric curve, the parametrization by arc-length is a special one. As an application we use it to define the curvature of a curve. This is geometric notion that is independent of the choice of parametrization. It is natural to define it using the arc-length parametrization. We will be restricted to plane curves in this section even though it is possible to define it for any curve in $\mathbb{R}^n$.

Let $\boldsymbol{\gamma} : [0, L] \to \mathbb{R}^2$ be a regular plane curve parametrized by the arc-length. Its tangent $\boldsymbol{\gamma}_s(s) = (x_s(s), y_s(s))$ is a unit vector according to Theorem 3.2, the vector $\mathbf{n}(s) = (-y_s(s), x_s(s))$ is perpendicular to $\boldsymbol{\gamma}_s$. We define it to be the unit normal at $\boldsymbol{\gamma}(s)$. When one walks along the tangential direction, this normal points to his/her left. In following we will use the notions

$$\mathbf{t} = \boldsymbol{\gamma}_s = (x_s, y_s) \ , \quad \mathbf{n} = (-y_s, x_s) \ .$$

The **signed curvature** of the curve at $\boldsymbol{\gamma}(s)$ is defined to be

$$\begin{aligned} k(s) &= \begin{vmatrix} x_s(s) & y_s(s) \\ x_{ss}(s) & y_{ss}(s) \end{vmatrix} \\ &= x_s y_{ss} - y_s x_{ss} \ . \end{aligned}$$

In the definition second derivatives of the curve are involved. Therefore, a parametric curve has to be twice differentiable in order its signature curvature to make sense. It will implicitly assumed throughout our discussion. Here k is viewed as a function of the arc-length variable s, but one may also regard it as a function defined on the geometric curve, that is, a function of $\boldsymbol{\gamma}(s)$. The reason it is called the signed curvature is that whenever the direction of the curve is reversed, the sign of the curvature also changes. Indeed, we pointed out that once we reverse the direction of the curve, the tangent vector points in the opposite direction but the acceleration vectors, which is the second derivative, point to the same direction. Therefore, the signed curvature which is the determinant obtained by taking the tangent and acceleration vector as the rows changes its accordingly. In order to define a sense of curvature solely depends on the geometric curve and in particular independently of the direction of the parametrization, we define the **curvature** of the curve at $\boldsymbol{\gamma}(s)$ to be the absolute value of the signed curvature, that is,

$$\kappa = |x_s y_{ss} - y_s x_{ss}| \ .$$

To understand the idea behind the definition we write

$$\mathbf{t} = (\cos\theta(s), \sin\theta(s))$$

using the fact that $\mathbf{t}$ is a unit vector for some $\theta \in [0, 2\pi)$. Indeed, the angle θ is the angle between the vector $\boldsymbol{\gamma}_s$ and the x-axis. Then

$$\mathbf{t}_s = (-\sin\theta(s), \cos\theta(s))\ \theta'(s) = \theta'(s)\mathbf{n} \ ,$$

which implies

$$\begin{aligned} k(s) &= x_s y_{ss} - y_s x_{ss} \\ &= < \mathbf{t}_s, \mathbf{n} > \\ &= \theta_s(s) \ . \end{aligned}$$

In other words, the signed curvature $k(s) = \theta_s(s)$ measures the rate of the change of the angle at the point $\boldsymbol{\gamma}(s)$. The magnitude of $\theta_s(s)$ described how curved the curve at this point. It has positive or negative sign according to whether the angle increases or decreases at this point.

In the following we express the signed curvature in an arbitrary parametrization.

Theorem 3.3. *Let $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^2$ be a twice differentiable regular parametric curve. Its signed curvature at $\boldsymbol{\gamma}(t)$ is given by*

$$k(t) = \frac{x'(t)y''(t) - y'(t)x''(t)}{\left(x'^2(t) + y'^2(t)\right)^{3/2}} \ .$$

Proof. Let Φ be the inverse to the length function defined above in Section 2. For a function defined on $[a,b]$, we write

$$\tilde{f}(s) = f(t), \quad t = \Phi(s).$$

By the Chain Rule

$$\tilde{f}_s(s) = f_t(t)\Phi_s(s) = \frac{f_t(t)}{\sqrt{x_t^2 + y_t^2}} \ .$$

Applying it to the two components of $\tilde{\boldsymbol{\gamma}}(s) = \boldsymbol{\gamma}(t)$, we have

$$\tilde{\boldsymbol{\gamma}}_s(s) = \frac{\boldsymbol{\gamma}_t(t)}{\sqrt{x_t^2 + y_t^2}} \ .$$

Next,

$$\begin{aligned}
\tilde{\boldsymbol{\gamma}}_{ss}(s) &= \frac{d}{ds}\tilde{\boldsymbol{\gamma}}_s(s) \\
&= \frac{d}{ds}\frac{\boldsymbol{\gamma}_t(t)}{\sqrt{x_t^2(t) + y_t^2(t)}} \\
&= \frac{1}{\sqrt{x_t^2(t) + y_t^2(t)}}\left(\frac{x_{tt}}{(x_t^2 + y_t^2)^{1/2}} - \frac{x_t(x_t x_{tt} + y_t y_{tt})}{(x_t^2 + y_t^2)^{3/2}}, \frac{y_{tt}}{(x_t^2 + y_t^2)^{1/2}} - \frac{y_t(x_t x_{tt} + y_t y_{tt})}{(x_t^2 + y_t^2)^{3/2}}\right) \\
&= \frac{x_t(t)y_{tt}(t) - y_t(t)x_{tt}(t)}{\left(x_t^2(t) + y_t^2(t)\right)^{3/2}}\mathbf{n}(s) \\
&= k\mathbf{n}(s),
\end{aligned}$$

hence $< \tilde{\boldsymbol{\gamma}}_{ss}(s), \mathbf{n}(s) >= k(t) < \mathbf{n}, \mathbf{n} >= k(t)$.

□

Example 3.13. Find the signed curvature of the straight line $\mathbf{l}(t) = \mathbf{p} + t\boldsymbol{\xi}$, $\mathbf{p} \in \mathbb{R}^n$, $t \in \mathbb{R}$. We have $\mathbf{l}'(t) = \boldsymbol{\xi}$ and $\mathbf{l}''(t) = (0,0)$. From the formula for the curvature in the proposition above that the curvature of the straight line is equal to zero everywhere.

Example 3.14. Find the signed curvature of the circle given by the parametrization $\mathbf{c}(t) = (R\cos t, R\sin t)$, $t \in [0, 2\pi]$ which runs in the counterclockwise direction. We have $\mathbf{c}'(t) = (-R\sin t, R\cos t)$ and $\mathbf{c}''(t) = (-R\cos t, -R\sin t)$, so

$$k = \frac{(-R\sin t)(-R\sin t) - (R\cos t)(-R\cos t)}{((-R\sin t)^2 + (R\cos t)^2)^{3/2}} = \frac{1}{R} \ .$$

We see that the larger the radius of the circle, the less its curvature. You can also check that the signed curvature becomes $-1/R$ in clockwise direction. The curvature of a circle of radius R is $1/R$.

In Section 3.3 we introduced the non-parametric form. Let f be a twice differentiable function over some interval. Its signed curvature and curvature are given respectively by

$$k(x) = \frac{f''(x)}{(1+f'^2(x))^{\frac{3}{2}}} \ , \qquad \text{(the signed curvature)}$$

and

$$\kappa(x) = \frac{|f''(x)|}{(1+f'^2(x))^{\frac{3}{2}}} \ , \qquad \text{(the curvature)} \ .$$

Example 3.15. Find the curvature of the parabola given by the non-parametric form

$$y = 3x^2 - 5x + 11 \ , \quad x \in \mathbb{R} \ .$$

The parametric curve describing the parabola is given by

$$\boldsymbol{\gamma}(x) = (x, 3x^2 - 5x + 11) \ .$$

We have

$$\boldsymbol{\gamma}'(x) = (1, 6x-5), \quad \boldsymbol{\gamma}''(x) = (0, 6).$$

Therefore,

$$\begin{aligned} k(x) &= \frac{y''(x)}{(1+y'^2)^{3/2}} \\ &= \frac{6}{(1+(6x-5)^2)^{3/2}} \\ &= \frac{6}{(36x^2-60x+26)^{3/2}} \ . \end{aligned}$$

We see that the curvature tends to zero as $|x| \to \infty$.

The following result partly justifies the terminology of the curvature for a curve.

Theorem 3.4. *Let $\boldsymbol{\gamma}$ be the regular curve in non-parametric form for some twice differentiable function. It is a line segment if its curvature vanishes everywhere. It is a circular arc if its curvature is a non-zero constant.*

Proof. Here $\boldsymbol{\gamma}(x) = (x, f(x))$ for some twice differentiable function f. It is always a regular curve since $|\boldsymbol{\gamma}'(x)| = (1 + f'^2(x))^{1/2} > 0$. When $k \equiv 0$, $f''(x) = 0$ for all x. Integrating twice shows that $f(x) = c_1 x + c_2$ so $\boldsymbol{\gamma}(x) = (0, c_2) + x(1, c_1)$ is a straight line.

When $k(x) = k_0$ for some non-zero k_0. We have

$$f''(x) = k_0(1 + f'^2)^{3/2} \ .$$

Writing $g = f'$, we have

$$g'(x) = k_0(1 + g^2(x))^{3/2} \ ,$$

or

$$\frac{dg}{(1+g^2)^{3/2}} = k_0 dx \ .$$

To proceed further, let us assume $f'(a) = 0$. This can always achieved by rotating the graph which does not change the shape of the curve. Integrating both sides from a to x yields

$$\frac{g(x)}{\sqrt{1+g^2}} = k_0(x-a) \ ,$$

or

$$f'(x) = \pm \frac{k_0(x-a)}{\sqrt{1 - k_0^2(x-a)^2}} \ .$$

One more integration from a to x yields

$$f(x) = f(a) \pm \frac{1}{k_0}\sqrt{1 - k_0^2(x-a)^2} \ ,$$

that is,

$$(x-a)^2 + (f(x) - f(a))^2 = \frac{1}{k_0^2},$$

so the curve is an arc of a circle of radius $1/|k_0|$ centered at $(a, f(a))$.

□

In the polar form the signed curvature and curvature are given respectively by and

$$k(\theta) = \frac{-\rho\rho'' + 2\rho'^2 + \rho^2}{(\rho^2 + \rho'^2)^{3/2}} \quad \text{(the signed curvature)} \ ,$$

and

$$\kappa(\theta) = \frac{\left|-\rho\rho'' + 2\rho'^2 + \rho^2\right|}{(\rho^2 + \rho'^2)^{3/2}} \quad \text{(the curvature)} \ .$$

3.5 Kepler's First Law*

Consider the motion of a small planet around a large one governed by Newton's second law and the gravitational law of inverse square. Assuming the large planet does not move in time and using it as the origin, the equation of motion becomes

$$m\frac{d^2\mathbf{r}}{dt^2} = -\frac{mMg}{|\mathbf{r}|^2}\frac{\mathbf{r}}{|\mathbf{r}|} ,$$

where m, M, g are respectively the mass of the small planet, the mass of the large planet, gravitational constant, and $\mathbf{r}(t) = (x(t), y(t), z(t))$ is the position of the small planet at time t. Componentwise we have

$$\begin{cases} \dfrac{d^2x}{dt^2} = \dfrac{-kx}{(x^2+y^2+z^2)^{3/2}} , \\[2ex] \dfrac{d^2y}{dt^2} = \dfrac{-ky}{(x^2+y^2+z^2)^{3/2}} , \\[2ex] \dfrac{d^2z}{dt^2} = \dfrac{-kz}{(x^2+y^2+z^2)^{3/2}} , \quad k = Mg. \end{cases}$$

There is a plane passing through two points $\mathbf{r}(0)$ and $\mathbf{r}'(0)$ and we can take it to be the xy-plane. In this coordinate system $z(0) = z'(0) = 0$ and $\mathbf{r}(0) \times \mathbf{r}'(0) = (0, 0, a)$ for some number a. We have

$$\begin{aligned} \frac{d}{dt}\mathbf{r} \times \mathbf{r}' &= \mathbf{r}' \times \mathbf{r}' + \mathbf{r} \times \mathbf{r}'' \\ &= \mathbf{r} \times \ \mathbf{r}'' \\ &= -\frac{Mg}{|\mathbf{r}|^3}\ \mathbf{r} \times \mathbf{r} = 0 , \end{aligned}$$

which shows that the cross product $\mathbf{r} \times \mathbf{r}'$ is independent of time, so $\mathbf{r} \times \mathbf{r}' = (0, 0, a)$ for all t. It means that the motion is always restricted to the xy-plane. Letting $z = 0$ in the system above, we obtain a reduced system consisting of two equations with two unknowns

$$\begin{cases} \dfrac{d^2x}{dt^2} = \dfrac{-kx}{(x^2+y^2)^{3/2}} , \\[2ex] \dfrac{d^2y}{dt^2} = \dfrac{-ky}{(x^2+y^2)^{3/2}} . \end{cases} \tag{3.1}$$

Kepler's First Law is contained in the following statement.

Theorem 3.5. *Every periodic solution of (3.1) traces out an ellipse.*

Since it is expected that the orbit is an ellipse, it is advantageous to employ the polar form to describe the system. Set

$$x = \rho\cos\theta, \quad y = \rho\sin\theta \ ,$$

where both ρ and θ are functions of time t. We have

$$x' = \rho'\cos\theta - \rho\sin\theta\ \theta', \quad x'' = \rho''\cos\theta - 2\rho'\sin\theta\ \theta' - \rho\cos\theta\ \theta'^2 - \rho\sin\theta\ \theta'' \ ,$$

$$y' = \rho'\sin\theta\theta + \rho\cos\theta\ \theta', \quad y'' = \rho''\sin\theta + 2\rho'\cos\theta\ \theta' - \rho\sin\theta\ \theta'^2 + \rho\cos\theta\ \theta'' \ .$$

The equations become

$$\rho''\cos\theta - 2\rho'\sin\theta\ \theta' - \rho\cos\theta\ \theta'^2 - \rho\sin\theta\ \theta'' = \frac{-k\cos\theta}{\rho^2} \ ,$$

$$\rho''\sin\theta + 2\rho'\cos\theta\ \theta' - \rho\sin\theta\ \theta'^2 + \rho\cos\theta\ \theta'' = \frac{-k\sin\theta}{\rho^2} \ .$$

Multiplying the first equation by $\cos\theta$, the second equation by $\sin\theta$ and then summing up, we get

$$\rho'' - \rho\theta'^2 = \frac{-k}{\rho^2} \ .$$

Multiplying the first equation by $\sin\theta$, multiplying the second equation by $\cos\theta$ and then subtracting, we get

$$2\rho'\theta' + \rho\theta'' = 0 \ .$$

The last equation can be rewritten as

$$\frac{d}{dt}(\rho^2\theta') = 0 \ ,$$

which subsequently gives

$$\theta' = \frac{c_0}{\rho^2} \ , \quad c_0 \text{ some constant.}$$

Now we would like to change the independent variable from t to the polar angle θ. Regarding $\rho = \rho(\theta)$, we have

$$\rho' = \rho_\theta\theta', \quad \rho'' = \rho_{\theta\theta}\theta'^2 + \rho_\theta\theta'' \ .$$

The first equation is transformed into

$$\rho_{\theta\theta}\theta^2 + \rho_\theta\theta'' - \rho\theta'^2 = \frac{-k}{\rho^2} \ .$$

Substituting θ' and θ'', we arrive at

$$\frac{c_0^2}{\rho^4}\left(\rho_{\theta\theta} - \frac{2\rho_\theta^2}{\rho} - \rho\right) = \frac{-k}{\rho^2} \ .$$

Motivated by the polar form of the ellipse, we let $\rho = 1/u$. Using

$$\rho_\theta = \frac{-u_\theta}{u^2}\ , \quad \rho_{\theta\theta} = \frac{-u_{\theta\theta}}{u^2} + \frac{2u_\theta^2}{u^3}\ ,$$

we finally get

$$u_{\theta\theta} + u = \frac{Mg}{c_0^2}\ ,$$

which is readily solved to give

$$u(\theta) = \frac{Mg}{c_0^2} + \alpha\cos\theta + \beta\sin\theta, \quad \alpha, \beta \text{ some constants}\ .$$

Since the motion is a periodic one, there is some θ_0 at which ρ attains its extremum. If now we rotate the coordinates so that θ_0 becomes 0, we have

$$u'(0) = -\frac{\rho'(0)}{\rho^2(0)} = 0\ .$$

Setting $u_0 = u(0)$ and $\rho_0 = \rho(0)$ and plugging $\theta = 0$ in the above expression for u we readily get

$$\alpha = u_0 - \frac{gm}{c_0^2}\ .$$

On the other hand, by differentiating this expression and then using $u'(0) = 0$, we obtain $\beta = 0$. We have arrived at

$$\begin{aligned}\rho(\theta) &= \frac{1}{\dfrac{gm}{c_0^2} + \left(\dfrac{1}{\rho_0} - \dfrac{gm}{c_0^2}\right)\cos\theta} \\ &= \frac{gm}{c_0^2} \times \frac{1}{1 + \left(\dfrac{c_0^2}{gm\rho_0} - 1\right)\cos\theta}\ .\end{aligned}$$

Setting

$$e = \left|\frac{c_0^2}{gm\rho_0} - 1\right|\ ,$$

in case $e \geq 1$, the dominator in the expression of ρ would vanish at some θ and ρ would diverge to infinity approaching this angle. However, this is impossible for a periodic motion. We conclude that $e < 1$. Finally we arrive at

$$\rho(\theta) = \frac{a(1 - e^2)}{1 \pm e\cos\theta}\ ,$$

for some positive a. This is the polar equation for an ellipse. We have shown that the Earth moves along an ellipse around the sun. Kepler's First Law is established.

Comments on Chapter 3.

3.1 We summarize our discussion on the various descriptions of geometric objects using the ellipse as a prototype. Consider the ellipse with foci at $(-c, 0), (c, 0), c > 0$ and major axis length $a, a > c$. It can be described in the following six ways.

(a) The loci of all points whose sum of distance to the two foci is equal to $2a$ ("the Greek way").

(b) The zero set of

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad b = \sqrt{a^2 - c^2} \ .$$

(c) The nonparametric form: The union of the graphs of the two functions given respectively by

$$f_1(x) = b\sqrt{1 - \frac{x^2}{a^2}} \ , \quad f_2(x) = -b\sqrt{1 - \frac{x^2}{a^2}} \ , \quad x \in [-a, a] \ .$$

(d) The parametric curve

$$\boldsymbol{\gamma}(t) = (a \cos t, b \sin t) \ , \quad t \in [0, 2\pi] \ .$$

(e) * The curve in arc-length parametrization:

$$\tilde{\boldsymbol{\gamma}}(s) = (a \cos t(s), b \sin t(s)) \ , \quad s \in [0, L] \ ,$$

where $t(s)$ is the inverse function of

$$s(t) = \int_0^t \sqrt{b^2 + (a^2 - b^2) \cos^2 z} \ dz \ , \quad L = s(2\pi) \ .$$

(f) The polar equation

$$\rho(\theta) = \frac{a(1 - e^2)}{1 \pm e \cos \theta} \ , \quad e = \frac{c}{a}, \quad \theta \in [0, 2\pi] \ .$$

The first two descriptions were discussed in Chapter 2 and the others four are in this chapter. Each approach has its advantage and shortcoming. The proof of Kepler's First Law illustrates how different descriptions are used to solve the problem. Testing for more, you may determine which one of these descriptions is the best in calculating the tangent line or enclosed area of the ellipse.

3.2 Kepler's discovered three laws on the motion of planets through empirical data. It is Newton who proposed the gravitational inverse square law. To test the validity of his law, Newton derived Kepler's laws as his, and this is hailed as the triumph of Newton's mechanics. The second and the third laws are less important, and you may google for their statements and proofs.

3.3 The parametrization approach to curves has been extended to surfaces, hypersurfaces and other geometric objects. The discussion is beyond this course. You will learn them in MATH2020 Advanced Calculus II and MATH4030 Differential Geometry.

Supplementary Readings

1.5 and 2.1 in [Au]. 11.1, 11.3, 11.4, 11.6, 11.7, 13.1 and 13.2 in [Thomas] contain a nice discussion on many interesting parametric curves.