Cutting Elliptical Pizza into Equal Slices

(12 November 2017)

Jim Stevenson

Having immersed myself in studying Kepler's discovery that the planetary orbits were ellipses ([1]), I was immediately aware of how the British mathematician, Katie Steckles, justified her technique to cut an elliptical pizza into equal slices. Here is her video of 14 March 2017:

Figure 1 Cutting an oval pizza into equal area slices (https://www.youtube.com/watch?v=K-VY2TsCano)

And here are snapshots of her solution:

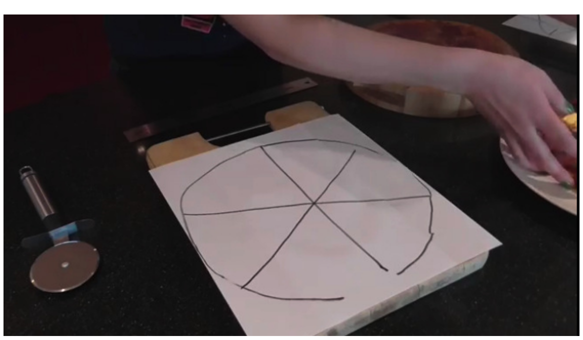


Figure 2 Equal Slices (Circle)

Figure 3 Equal Slices (Ellipse)

"Compressed Circle"

Katie was employing the "compressed circle" definition of an ellipse that Kepler used before the advent of analytic geometry in the 17^{th} century. Consider obtaining an ellipse with semimajor axis a and semiminor axis b. Suppose a circle of radius a is drawn and that the distance of every point on the circle perpendicular to a fixed diameter is shrunk by a factor b/a. Then the resulting curve will be an ellipse with semimajor axis a and semiminor axis b (See Figure 4).

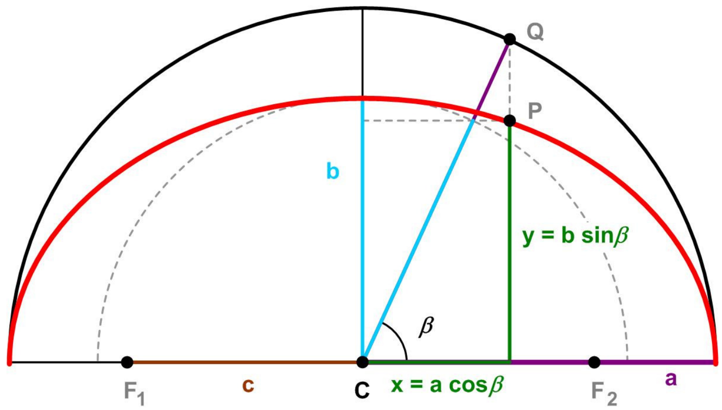


Figure 4 Parametric Definition of Ellipse

Using modern analytic geometry, we can show this property easily. From Figure 4 we see that the coordinates (x, y) on the ellipse are given by

$$x = a\cos\beta$$

$$y = b\sin\beta \tag{1}$$

where β is called the "eccentric anomaly" and describes a point Q moving around the circle with radius a. The distance from Q perpendicularly down to the diameter of the circle is given by a sin β. So the ratio of y, the ordinate of the ellipse, to this distance is y / a sin β = b sin β / a sin β = b/a, which is the designated compression factor.

Equal Areas

We now consider "slices" of the pizza beginning with sectors S of the circle (Figure 5).

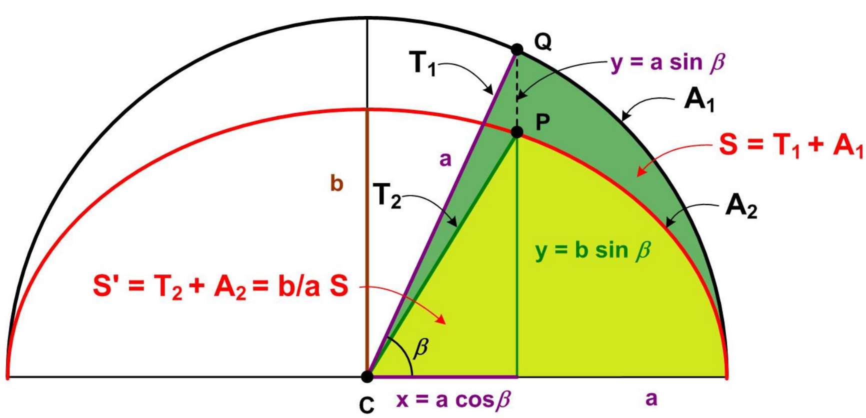


Figure 5 Area Compression

The sector S consists of the triangle T_1 and area A_1 under the circle up to point Q. The corresponding area under the ellipse S' is given by triangle T_2 and area A_2. Since the altitude of triangle T_2 is the altitude of T_1 shrunk by b/a and they share the same base, we have

$$T_2 = b/a\ T_1$$

Similarly, since all the ordinates of the ellipse are shrunk versions of the ordinates of the circle, where the shrink factor is a constant (b/a), then the areas are also shrunk by that factor:

$$A_2 = b/a\ A_1$$

Therefore,

$$S' = T_2 + A_2 = b/a\ T_1 + b/a\ A_1 = b/a\ (T_1 + A_1) = b/a\ S$$

This means if two slices (sectors) of the circle have the same areas, then so do the compressed versions, since they are a constant multiple of the circle sectors. Thus we get the equally-sliced result shown in Figure 6 and in Figure 3 of Katie's video.

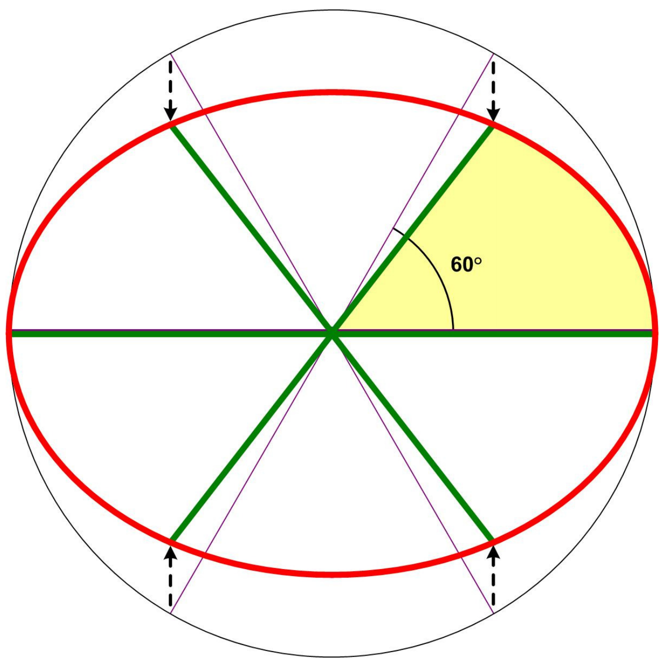


Figure 6 Equal Slices for an Ellipse

Affine Transformations

In her video Katie makes the claim that the result of any affine transformation of the circular pizza cut into equal sectors will also be a set of equal area slices. We need to explain what an affine transformation is (where the "compressed circle" is a special case) and then be more precise about the definition of area. Unfortunately, I can't see how to do this without involving calculus. I will try to keep it as simple (intuitive) as possible.

An affine transformation F sends points (x, y) to points (u, v) defined by the equations

$$u = a\,x + b\,y + e$$
$$v = c\,x + d\,y + f \tag{2}$$

where a, b, c, d, e, f are constants. In matrix notation, equation (2) can be written

$$\begin{pmatrix} u \\ v \end{pmatrix} = F\begin{pmatrix} x \\ y \end{pmatrix} = T\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix} \tag{3}$$

(See Figure 7.) So an affine transformation F consists of a linear transformation T and a translation by (e, f). In order not to have an area collapsed to a line, we assume the linear transformation T is non-singular, that is, its determinant, det T, is non-zero, where

$$\det T = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \tag{4}$$

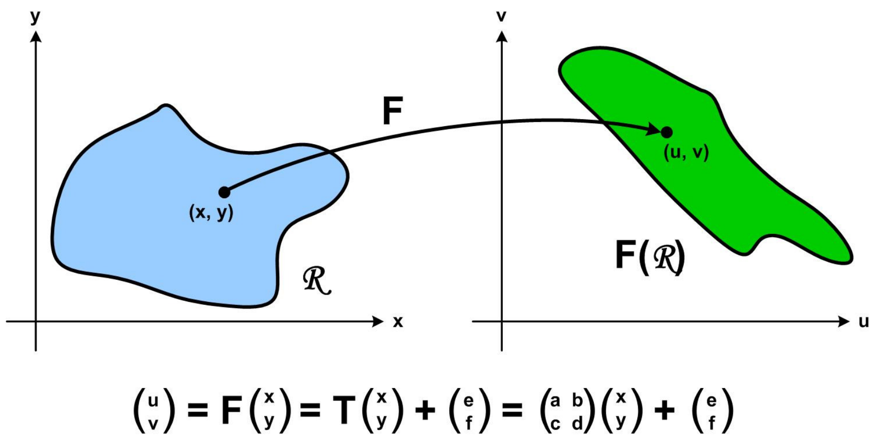


Figure 7 Affine Transformation

If the linear transformation T has determinant of absolute value 1, then the affine transformation F represents all rigid motions in a plane (rotation, translation, reflection). Otherwise, T can shrink, expand, shear, skew, flip, and rotate the original figure. Notice that the "compressed circle" transformation above is given by setting all coefficients equal to zero except a = 1 and d = semiminor/semimajor.

Areas

Now we wish to consider the effect of an affine transformation on the area of a region $\mathcal{R}$. Figure 8 indicates how we can define the area inside a curved figure. First we fill the area with a grid of squares where the k^{th} small square has an area $\Delta A_k = \Delta x_k \, \Delta y_k$. We then sum up all the squares inside the curve to get an approximation to the actual area: $P_n = \Sigma \, \Delta A_k = \Sigma \, \Delta x_k \, \Delta y_k$ (where the Greek S, Σ,

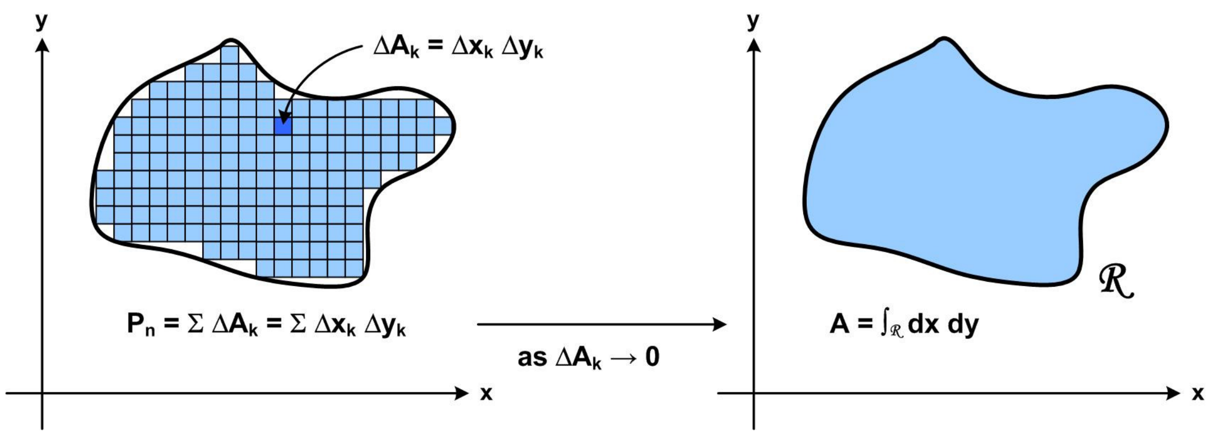


Figure 8 Limit Definition of Area (Integration)

represents summing). By taking smaller and smaller squares, we get a grid that approaches the actual area better and better. The limiting value of these approximations is defined to be the area inside the curve and is denoted by the integral ($\int$)

$$\text{Area of } \mathcal{R} = \int_{\mathcal{R}} dxdy.^{[1]} \qquad (5)$$

Differentials

The other side of calculus (besides integration) involves derivatives or rates of change. I will be able to provide even less guidance here. Basically the idea is to consider how one variable, say y, changes as another variable, say x, that it depends on, changes. An *average* rate of change would be $\Delta y/\Delta x$ where Δy represents how much y changed when x changed by an amount Δx. Now we consider the limit as the change in x, Δx, becomes shorter and shorter, that is, as $\Delta x \rightarrow 0$. Since $\Delta y \rightarrow 0$ also, we nominally have a limit of 0/0, which is nonsense. But in fact there are ample cases where a limit actually exists, and when that happens, we say this limit represents the *instantaneous* rate of change of y with respect to x, or the *derivative* of y with respect to x. We denote this limit by dy/dx. As a carry-over from the early history of calculus, we call dy and dx *differentials* and imagine them as infinitesimal entities (smallest, non-zero values – which again is nonsense).

When several variables are involved, such as a variable u depending on variables x and y, we denote the derivatives of u as partial derivatives $\partial u/\partial x$ and $\partial u/\partial y$, where $\partial u/\partial x$ means the instantaneous rate of change of u with respect to x when y is held constant and $\partial u/\partial y$ means the instantaneous rate of change of u with respect to y when x is held constant. In terms of differentials we have (without further explanation)

$$du = \partial u/\partial x \; dx + \partial u/\partial y \; dy$$

$$dv = \partial v/\partial x \; dx + \partial v/\partial y \; dy \qquad (6)$$

In matrix notation we can write all this as

$$\begin{pmatrix} u \\ v \end{pmatrix} = F\begin{pmatrix} x \\ y \end{pmatrix} \Rightarrow \begin{pmatrix} du \\ dv \end{pmatrix} = DF\begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} \end{pmatrix}\begin{pmatrix} dx \\ dy \end{pmatrix} \qquad (7)$$

DF is called the *Jacobian* matrix of the transformation F.

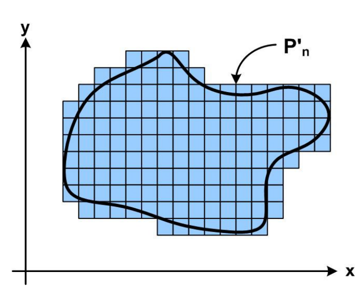
Back to the areas. Consider a transformation F (not necessarily affine, but still differentiable) that maps a region $\mathcal{R}$ onto a region $F(\mathcal{R})$. We would like to know how to represent the area of $F(\mathcal{R})$ in terms of the area of $\mathcal{R}$. Using the integral representation of area given by equation (5), we have (a partial explanation is given in the Appendix p.6):

---

[1] This is an extremely simplified presentation. To make sure a limit could exist, the Greeks also considered grids that circumscribed the curved region and shrunk down as smaller and smaller squares were considered. That is, if P'_n represented a decreasing sequence of circumscribing grid areas with successively smaller squares, and P_n represented an increasing sequence of inscribed grid areas, then

$$\ldots P_k < P_{k+1} < \ldots < P'_{k+1} < P'_k < \ldots$$

which trap a limiting value between them. This value is defined to be the area (integral). (This is still an oversimplification, but it suffices for now.)



$$\text{Area of } F(\mathcal{R}) = \int_{F(R)} du dv = \int_{R} (\det DF) dx dy \tag{8}$$

Now an affine transformation F is differentiable and its Jacobian matrix DF is just the linear transformation T. Therefore the det DF = Det T = ad – bc is a constant (because a, b, c, and d are), which we denote by D_0. From equation (8) for an affine transformation F, we have

$$\text{Area of } F(\mathcal{R}) = \int_{F(R)} du dv = \int_{R} D_0 dx dy = D_0 \int_{R} dx dy = D_0 \text{ Area of } \mathcal{R} \tag{9}$$

The key is that det DF is a constant that can be pulled out from under the integral. A non-singular affine transformation F maps non-overlapping regions $\mathcal{R}$ and $\mathcal{S}$ to non-overlapping regions $F(\mathcal{R})$ and $F(\mathcal{S})$. If the two regions $\mathcal{R}$ and $\mathcal{S}$ have the same area, then so do their images $F(\mathcal{R})$ and $F(\mathcal{S})$:

$$\text{Area of } F(\mathcal{R}) = D_0 \text{ Area of } \mathcal{R} = D_0 \text{ Area of } \mathcal{S} = \text{Area of } F(\mathcal{S})$$

And so we have the generalization that any affine transformation of a circular pizza cut into equal slices will produce a distorted pizza still cut into equal slices.

Appendix: Linear Transformation of Areas

We are going to try to motivate the result in Equation (8). Equation (7) above shows that DF represents a linear transformation of differentials. So we shall consider the behavior of a linear transformation T from the xy-plane $\mathbb{R}^2$ into the xy-plane $\mathbb{R}^2$. Points (x, y) in the plane can be represented by vectors from the origin, such as $\mathbf{z} = x\mathbf{i} + y\mathbf{j}$ where $\mathbf{i}$ is the unit basis vector of length 1 along the x-axis and $\mathbf{j}$ is the unit basis vector of length 1 along the y-axis (Figure 9). (We show vectors in boldface and scalars (real numbers) in non-boldface font.)

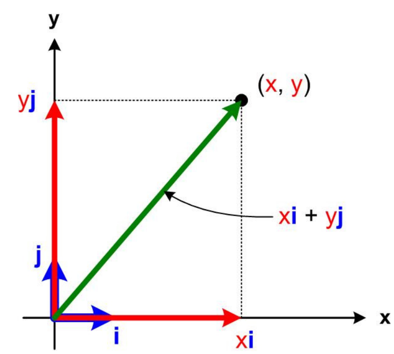


Figure 9 Vector Representation

Linear Transformation. Recall that a map T: $\mathbb{R}^2 \to \mathbb{R}^2$ is a *linear transformation* if for every vector **u** and **v** and scalar a and b,

$$T(a\mathbf{u} + b\mathbf{v}) = a\,T(\mathbf{u}) + b\,T(\mathbf{v}) \tag{10}$$

Vector (Cross) Product. Recall that the *vector* or *cross product* of two vectors **u** and **v**, $\mathbf{u} \times \mathbf{v}$, has the following properties.

1. For any two vectors **u** and **v**, $\mathbf{u} \times \mathbf{v}$ is a vector
2. For any scalar a, $a(\mathbf{u} \times \mathbf{v}) = (a\mathbf{u}) \times \mathbf{v} = \mathbf{u} \times (a\mathbf{v})$
3. For any three vectors **u**, **v**, and **w**,

$$\mathbf{u} \times (\mathbf{v} + \mathbf{w}) = \mathbf{u} \times \mathbf{v} + \mathbf{u} \times \mathbf{w} \quad \text{and} \quad (\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$$

4. For any two vectors **u** and **v**, $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$

Property 4 implies that $\mathbf{0} = \mathbf{u} \times \mathbf{v} + \mathbf{v} \times \mathbf{u}$, which in turn implies $\mathbf{0} = 2\,\mathbf{u} \times \mathbf{u}$, or $\mathbf{u} \times \mathbf{u} = \mathbf{0}$. If $\mathbf{u} = a\mathbf{i} + b\mathbf{j}$ and $\mathbf{v} = c\mathbf{i} + d\mathbf{j}$, then

$$\mathbf{u} \times \mathbf{v} = (ad - bc)\,\mathbf{i} \times \mathbf{j} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \mathbf{i} \times \mathbf{j} = \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mathbf{i} \times \mathbf{j} = \det\begin{pmatrix} a & c \\ b & d \end{pmatrix} \mathbf{i} \times \mathbf{j} \tag{11}$$

One final property, which we will not prove here, is that the magnitude of the vector $\mathbf{u} \times \mathbf{v}$ is given by

$$|\mathbf{u} \times \mathbf{v}| = |\mathbf{u}|\,|\mathbf{v}| \sin\theta \qquad (12)$$

where $|\mathbf{u}|$ and $|\mathbf{v}|$ are the magnitudes (lengths), respectively, of **u** and **v** and θ is the angle between them. As we can see in Figure 10, this implies that the magnitude of the cross product represents the area of the parallelogram spanned by the two vectors.

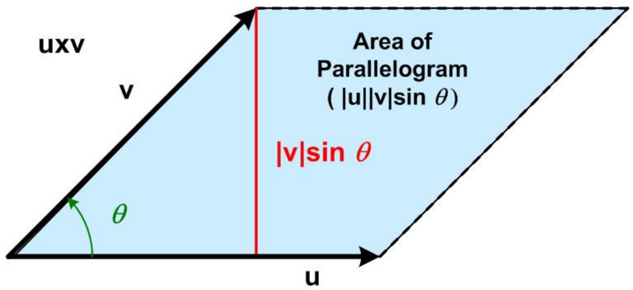


Figure 10   Magnitude of Cross Product is Area

We are now ready to show the behavior of a linear transformation T of $\mathbb{R}^2$ with respect to areas. All we need to compute are the values of T for basis vectors **i** and **j**, since by linearity we can determine how it maps any other vectors. Suppose we have

$$T(\mathbf{i}) = a\mathbf{i} + c\mathbf{j}$$

$$T(\mathbf{j}) = b\mathbf{i} + d\mathbf{j}$$

for some scalars a, b, c, and d. The matrix associated with T is then defined to be

$$\text{mat } T = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

(notice the transpose of b and c). From equation (11) we have

$$T(\mathbf{i}) \times T(\mathbf{j}) = \det T \; \mathbf{i} \times \mathbf{j}$$

Now let us consider a small rectangular region $\mathcal{R}$ with area $\Delta A = \Delta x\, \Delta y$ (Figure 11). Under the linear transformation T, $\mathcal{R}$ is mapped to the parallelogram region $T(\mathcal{R})$ spanned by the vectors $\Delta\mathbf{u}$ and $\Delta\mathbf{v}$ where

$$\Delta\mathbf{u} = T(\Delta x\mathbf{i}) = T(\mathbf{i})\, \Delta x$$

$$\Delta\mathbf{v} = T(\Delta y\mathbf{j}) = T(\mathbf{j})\, \Delta y$$

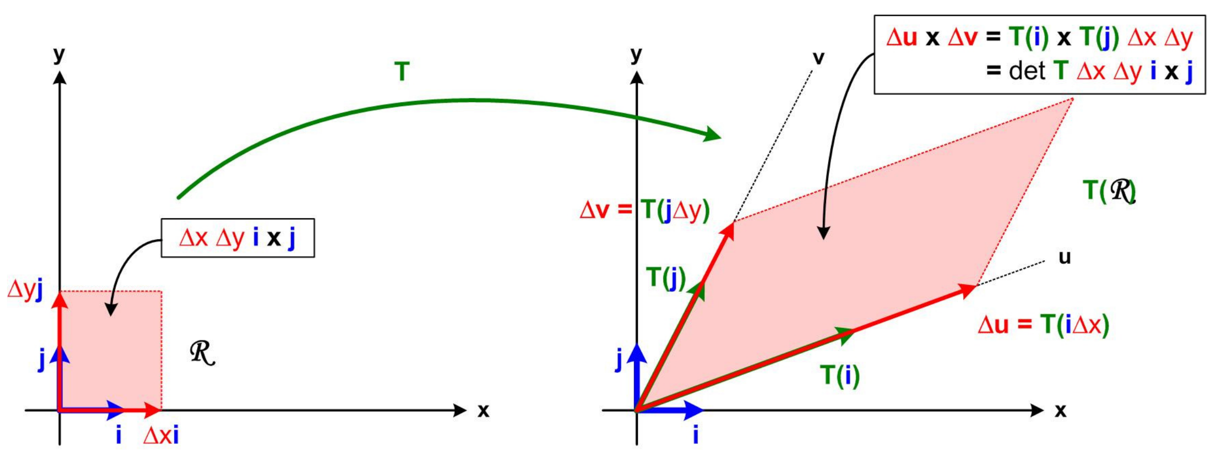


Figure 11   Linear Transformation of Areas in the Plane

Then the area of $T(\mathcal{R})$ is given as the magnitude of

$$\mathbf{\Delta u} \times \mathbf{\Delta v} = T(\mathbf{i}) \times T(\mathbf{j})\ \Delta x\ \Delta y = \det T\ \Delta x\ \Delta y\ \mathbf{i} \times \mathbf{j} = \det T \text{ Area of } \mathcal{R}\ \mathbf{i} \times \mathbf{j}$$

so that

$$\text{Area of } T(\mathcal{R}) = \det T \text{ Area of } \mathcal{R}$$

Equation (8) above (p.6) basically derives from this property where the incremental areas pass to the limit represented by the differentials.

References

1. Stevenson, James, "Kepler's Ellipse," in *Meditations on Mathematics*, 27 December 2016