1. Euclidean Geometry

- Euclidean geometry describes the world well.
- It allows to measure lengths and angles.
- Length, angles, parallelism, orthogonality, and all other properties that are related via a linear/Euclidean transform are preserved.
- Euclidean coordinates of a point in a plane are given by a 2-tuple \([u, v]^T\).
- Ex: Consider the transformation that rotates 2 points, \(P_1, P_2\), in a plane counter-clockwise \(\theta\) with respect to the origin as shown in Fig. 1. The transformation can be represented by the linear equations,

\[
P_1' = R \cdot P_1 = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \cdot P_1
\]

\[
P_2' = R \cdot P_2 = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \cdot P_2
\]

Since the transformation is Euclidean, the length between the two points, and the angle subtended at the origin, before and after the transformation remains the same.

Q – Why do we need Projective geometry?

A – Because 3D objects are projected on to a 2D plane on capturing an image.

![Fig. 1. – Rotating a point about the origin is a Euclidean transformation.](image-url)
2. Projective Geometry

- Describes projection to lower dimensions well. For instance, parallel lines in 3D space are no longer parallel in a 2D image projection, and appear to meet. Such properties are captured well by projective geometry.
- The horizon has the same projection.
- Since parallelism between lines is not preserved, distances or angles are not preserved either.
- Projective geometry describes a larger class of transformations. It is an extension of Euclidean geometry and deals with the perspective projection of a camera.
- Projective coordinates of a point in a plane are homogenized and represented by a 3-tuple: \([u, v, 1]^T\).
- Rule: Scaling the projective coordinates by a non-zero factor does not change the Euclidean point it represents as it is homogenized. i.e., \([u, v, 1]^T \equiv [\lambda u, \lambda v, \lambda]^T\).

3. Projective Space

The Euclidean coordinates of a point in a plane can be represented by a 2-tuple: \([u, v]^T\). It's projective coordinates are obtained by appending a 1 to the vector: \([u, v, 1]^T\). By representing the point by this 3-tuple in projective coordinates, a one-to-one mapping is established between the 2D point in Euclidean coordinates and the corresponding point in projective coordinates. Thus, scaling the point by a non-zero zero factor does not change the Euclidean point it represents as it is homogenized. i.e., \([u, v, 1]^T \equiv [\lambda u, \lambda v, \lambda]^T\). Thus, projective coordinates represent naturally the operations performed by cameras.

**Definition:** The space of \((n + 1)\)-tuples of coordinates, with the rule that proportional (or scaled) \((n + 1)\)-tuples represent the same point, is called a *projective space* of dimension \(n\), and is denoted \(\mathbb{P}^n\).

In general, given coordinates in \(\mathbb{R}^n\), the corresponding projective coordinates are obtained as,

\[
[x_1, x_2, ..., x_n]^T \rightarrow [x_1, x_2, ..., x_n, 1]^T. \tag{3.1}
\]

To transform a point from projective coordinates back to Euclidean coordinates, we just need to divide by the last coordinate and the drop the last coordinate,

\[
[x_1, x_2, ..., x_n, x_{n+1}]^T \rightarrow [\frac{x_1}{x_{n+1}}, \frac{x_2}{x_{n+1}}, ..., \frac{x_n}{x_{n+1}}]^T. \tag{3.2}
\]

Points with last coordinate \(x_{n+1} \neq 0\) are usual points with representations in \(\mathbb{R}^n\), but points of the form \([x_1, x_2, ..., x_n, 0]^T\), do not have an equivalent representation in Euclidean coordinates. If we consider them as the limit of \([x_1, x_2, ..., x_n, \lambda]^T\), when \(\lambda \rightarrow 0\), i.e. the limit of \([\lambda x_1 / \lambda, x_2 / \lambda, ..., x_n / \lambda, 1]^T\) then they represent the limit of a point in \(\mathbb{R}^n\) going to infinity in the direction \([x_1, x_2, ..., x_n]^T\). Such points are called *points at infinity*.

Thus projective space contains more points than the Euclidean space of same dimensions, and is a union of the usual space \(\mathbb{R}^n\) and the set of points at infinity. i.e.,

\[
\mathbb{P}^n = \mathbb{R}^n \cup \{[x_1, x_2, ..., x_n, 0]^T\}. \tag{3.3}
\]

As a result of this formalism, points at infinity are represented without exceptions in projective space.
4. 3D to 2D Projection

Fig. 2. represents the pinhole model of projection of a point in 3D onto the 2D image plane, $R$. Point $C$ is the optical center, and does not belong to $R$. The projection, $m$, of a point in 3D space, $M$, is the intersection of the optical ray, $(C, M)$, with the image plane. The optical axis is the line through $C$ perpendicular to the image plane, pierces the image plane at the principal point, $c$.

The camera coordinate system is established by the orthonormal vectors, $x$, $y$ and $z$, centered at the optical center, $C$. Here the image coordinate system is aligned with the camera coordinate system. The distance between $C$ and $R$ is the focal length, and is chosen as unity without loss of generality. Thus the relationship between the camera-coordinates of point $M = [X, Y, Z]^T$, and the corresponding image-coordinates, $m = [u, v]^T$ is,

$$ u = \frac{X}{Z}, \quad v = \frac{Y}{Z} $$

(4.1)

Once the projection has been captured by the image, the true 3D depth of the point $M$, can no longer be inferred from a single image due to the inherent nature of 3D to 2D projection. Thus any other point, $M' = [\lambda X, \lambda Y, \lambda Z]^T$, that lies on the optical ray $(C, M)$, also has the same 2D-projection, $m$. This depth ambiguity cannot be inferred from a single image of the point using geometry alone, and the only information available from the single image projection is the vector along which the 3D point lies in space.