GPU Fast and Robust Computation for Barycentric Coordinates and Intersection of Planes Using Projective Representation

Vaclav Skala
Department of Computer Science and Engineering
University of West Bohemia
Plzen, Czech Republic
http://www.VaclavSkala.eu

Abstract—This paper describes algorithms for fast and robust GPU computation of barycentric coordinates and intersection of two planes. The presented algorithms are based on matrix-vector operations which make the algorithms convenient for GPU or SSE based architectures. Also a new formula for finding the closest point of two planes intersection to the given point is given.

Keywords—GPU, barycentric coordinates, two planes intersection, Plücker coordinates, closest point, homogeneous coordinates, projective space

I. INTRODUCTION

Intersection of two planes is frequently used in geometric computations, e.g. in set operations with geometric objects given by triangular meshes. In spite of the simplicity of the problem, the standard formulas for computations are not robust in cases when given planes are almost parallel or parallel. Also barycentric coordinates computation, which leads to a solution of linear equations, is very often used to solve geometrical problems.

In this paper standard approaches will be presented together with a new more robust approach especially convenient for GPU or SSE application based on projective representation. In the case of two triangles intersection computation, the vertices of triangles can be given in homogeneous coordinates and intersection of triangles can be computed directly without need to convert vertices coordinates to the Euclidean coordinates. This saves 18 division operations per one pair of triangles intersection computation, which leads to significant speed-up.

Similarly, barycentric coordinates can be computed directly in the projective space if coordinates of vertices are in homogeneous coordinates and the conversion to the Euclidean space is not needed, i.e. we save division operations as well.

In the following the notation will be used:

- \( X = (X, Y, Z) \in E^3 \) for coordinates or vectors in \( E^3 \)
- \( x = [x, y, z; 1]^T \) or \( x = [x, y, z; w]^T \) homogeneous coordinates in the projective space
- \( a \times b \) is the outer (cross) vector product
- \( a^T b = a \cdot b \) for dot (scalar) product

II. PROBLEM FORMULATION

Intersection of two planes is a seemingly simple problem, see Fig.1.

Figure 1: Intersection of two planes

Let us assume two planes \( \rho_1 \) and \( \rho_2 \), Fig.1, given by two vectors \( \rho_1 = [a_1, b_1, c_1; d_1]^T \) and \( \rho_2 = [a_2, b_2, c_2; d_2]^T \) as:

\[
\rho_1: a_1 x + b_1 y + c_1 z + d_1 = 0
\]

and

\[
\rho_2: a_2 x + b_2 y + c_2 z + d_2 = 0
\]

where: \( n_1 = [a_1, b_1, c_1]^T \) and \( n_2 = [a_2, b_2, c_2]^T \)

We need to determine a line \( p \in E^3 \) which is given as an intersection of those two planes if planes are “not parallel”. However due to the numerical precision in the floating point representation, the condition must be weaken to “not close to parallel”. There is the key problem, i.e. what does actually this condition mean and how it is defined from the algorithmic point of view.

It can be seen that the line \( p \) in the parametric form can be then determined as:

\[
s = n_1 \times n_2 = [a_3, b_3, c_3]^T \quad X(t) = X_0 + st
\]

As the computation of the vector \( s \) is simple and relatively precise, the “problem” is how to determine the point \( X_0 \) complying (1) and (2), reliably from the numerical
precision point of view. Unfortunately in some applications it can be found incorrect or non-robust solutions.

III. STANDARD SOLUTION

Standard formula for a line given as an intersection of two planes in the Euclidean space is given as:

\[ s = n_1 \times n_2 \equiv [a_2, b_2, c_2]^T \]

\[ X(t) = X_0 + s t \]

Then the “starting” point \( X_0 \) is given as

\[
X_0 = \begin{bmatrix}
\frac{a_2}{a_1} & \frac{b_2}{b_1} & \frac{c_2}{c_1} \\
\frac{a_3}{a_1} & \frac{b_3}{b_1} & \frac{c_3}{c_1} \\
\frac{a_1}{a_2} & \frac{b_1}{b_2} & \frac{c_1}{c_2}
\end{bmatrix} \frac{c_1}{\text{DET}}
\]

\[
Y_0 = \begin{bmatrix}
\frac{a_2}{a_1} & \frac{b_2}{b_1} & \frac{c_2}{c_1} \\
\frac{a_3}{a_1} & \frac{b_3}{b_1} & \frac{c_3}{c_1} \\
\frac{a_1}{a_3} & \frac{b_1}{b_3} & \frac{c_1}{c_3}
\end{bmatrix} \frac{c_1}{\text{DET}}
\]

\[
Z_0 = \begin{bmatrix}
\frac{a_2}{a_1} & \frac{b_2}{b_1} & \frac{c_2}{c_1} \\
\frac{a_3}{a_1} & \frac{b_3}{b_1} & \frac{c_3}{c_1} \\
\frac{a_1}{a_3} & \frac{b_1}{b_3} & \frac{c_1}{c_3}
\end{bmatrix} \frac{c_1}{\text{DET}}
\]

This formula is quite “horrible” one and for users is too complex to remember and they do not see from the formula comes from.

It can be seen that coordinates are computed using division operation and division by \( \text{DET} \) is required. There is a legitimate question about specification of cases when division by \( \text{DET} \) causes floating point overflow or underflow. In many cases the programmer uses a sequence:

\[ \text{if } \text{DET} < \text{eps } \text{then Error} \]

but it can happen that the expression \( \text{eps}/\text{DET} \) is too small for the \( x_0 \)-coordinate, and the expression is actually “well conditioned”. The formula for \( X_0 \) computation is not convenient for GPU, nor SSE application, too.

IV. PROJECTIVE SPACE REPRESENTATION

The projective extension of the Euclidean space is often used in geometry, sometimes is “hidden” in the known formulas.

Let us consider a point \( X = (X,Y) \in E^2 \) and its equivalent in the projective space \( [x,y:w]^T \in P^2 \). The mutual conversion is then given as [1],[3]:

\[
X = x/w \quad Y = y/w \quad w \neq 0
\]

It can be seen that it is actually one parametric set, i.e. a line passing the origin in the \( x,y,w \) coordinate system, but origin is excluded. The extension to the \( E^3 \) case is straightforward.

Homogeneous coordinates and the projective space representation is widely used in computer graphics to represent geometric transformations, projection operation etc.

Geometric transformations with points are described generally as:

\[
\mathbf{x'} = \mathbf{Qx}
\]

where \( \mathbf{Q} \) is a matrix \( 3 \times 3 \) in the \( E^2 \) case, or \( 4 \times 4 \) in the \( E^3 \) case.

The advantage of the projective representation is that fundamental transformations like translation, rotation, scaling, shearing and projections \( E^3 \to E^2 \) are represented by a cumulative transformation matrix multiplied by a vector containing point’s coordinates.

However, geometric transformations with lines in \( E^2 \) or planes in \( E^3 \) are given as:

\[
\mathbf{p'} = (\mathbf{Q}^{-1})^T \mathbf{p} \\
\mathbf{\rho'} = (\mathbf{Q}^{-1})^T \mathbf{\rho}
\]

as normal vectors are actually bivectors. It can be seen that the transformation matrices for lines or planes are different from the transformation matrix of points defining the given goe.

V. PRINCIPLE OF DUALITY

The principle of duality in \( P^2 \) states that any theorem remains true when we interchange the words “point” and “line”, “lie on” and “pass through”, “join” and “intersection”, “collinear” and “concurrent” and so on. In the case of \( P^3 \) dual identities are “point” and “plane” etc. Once the theorem has been established, the dual theorem is obtained as described above [2], [5].

In other words, the principle of duality says that in all theorems in \( E^2 \) it is possible to substitute the term “point” by the term “line” and the term “line” by the term “point” and the given theorem stays valid. In the case of \( E^3 \) dual terms are “point” and “plane”. This helps a lot to solve some geometrical problems.

A nice example of projective space representation and principle of duality application in the \( E^2 \) case is computation of a line \( \mathbf{p} \) given by two given points \( \mathbf{x_1}, \text{Eq.(8)} \)

\[
\mathbf{p} = \mathbf{x_1} \times \mathbf{x_2} = [x_1, y_1, w_1]^T \times [x_2, y_2, w_2]^T = [a, b, c]^T
\]

and an intersection point \( \mathbf{x} \) of two lines \( \mathbf{p_1}, \text{Eq.(9)} \)

\[
\mathbf{x} = \mathbf{p_1} \times \mathbf{p_2} = [a_1, b_1, c_1]^T \times [a_2, b_2, c_2]^T = [x, y, w]^T
\]

It can be shown that computation of an intersection of two lines in \( E^2 \) is dual to computation of a line given by two points (it is actually a join operation). It means that there should be the same programming sequence for solving both dual cases [7], i.e.

\[
\mathbf{x} = \mathbf{p_1} \times \mathbf{p_2} = \begin{vmatrix}
\mathbf{i} & \mathbf{j} & \mathbf{k} \\
\mathbf{a_1} & \mathbf{b_1} & \mathbf{d_1} \\
\mathbf{a_2} & \mathbf{b_2} & \mathbf{d_2}
\end{vmatrix}
\]

\[
\det \begin{vmatrix}
\mathbf{x} & \mathbf{y} & \mathbf{w} \\
\mathbf{a_1} & \mathbf{b_1} & \mathbf{d_1} \\
\mathbf{a_2} & \mathbf{b_2} & \mathbf{d_2}
\end{vmatrix} = 0
\]
In the case of $E^3$, a plane given as a join of three points and dual problem, i.e. an intersection of three planes, can be computed as \[8\]:

\[
p = x_1 \times x_2 = \begin{vmatrix} i & j & k \\ x_1 & y_1 & w_1 \\ x_2 & y_2 & w_2 \end{vmatrix}
\]

\[\text{det} \begin{vmatrix} a & b \\ x_1 & y_1 \end{vmatrix} = 0 \]

and

\[
x = q_1 \times q_2 \times q_3 = \begin{vmatrix} a_1 & b_1 \\ x_1 & y_1 \end{vmatrix} 
\begin{vmatrix} c_1 & d_1 \\ z_1 & w_1 \end{vmatrix} 
\begin{vmatrix} a_2 & b_2 \\ x_2 & y_2 \end{vmatrix} 
\begin{vmatrix} c_2 & d_2 \\ z_2 & w_2 \end{vmatrix} 
\begin{vmatrix} a_3 & b_3 \\ x_3 & y_3 \end{vmatrix} 
\begin{vmatrix} c_3 & d_3 \\ z_3 & w_3 \end{vmatrix}
\]

There are significant advantages of this approach:

- Natural support for GPU computation leading to significant speed up.
- There is no division operation used that is needed if an intersection is computed in the Euclidean space, like \( \text{det} x / \text{det} \), which leads to instability in principle.
- There are no special cases which a programmer has to take care of, i.e. detection of collinearity etc.
- If points are given in homogeneous coordinates, transformation to the Euclidean coordinates is not required. It means that 6, resp. 9 division operations are not needed in the case of $E^2$, resp. $E^3$, and higher precision can be expected as well.

As a direct consequence it can be proved \([9]\) that a solution of a linear system of equations is equivalent to an extended cross-product \([9]\). This is a significant result as instead of solving a linear system of equations $Ax = b$ we can use extended cross-product. It should be noted that replacing a solution of linear system of equations $Ax = b$ by the extended cross-product $x = \xi_1 \times \xi_2 \times \ldots \times \xi_n$ enables also further formal manipulation using linear algebra.

As a typical example of this approach is computation of barycentric coordinates which can be used also for the case, when points $x_1, \ldots, x_n$ are given in homogeneous coordinates.

VI. PLÜCKER COORDINATES

The Plücker coordinates are often used in robotics and geometrical problems solutions, e.g. in a ray-triangle intersection detection and intersection computation \([4]\).

Let us consider two points in the homogeneous coordinates:

\[
x_1 = [x_1, y_1, z_1; w_1]^{T} \quad x_2 = [x_2, y_2, z_2; w_2]^{T}
\]

(14) defining a line $p \in E^3$. The Plücker coordinates $l_{ij}$ of the anti-symmetric matrix $L$ are defined as:

\[
L = x_1 \times x_2 = [x_1, y_1, z_1; w_1]^{T} - [x_2, y_2, z_2; w_2]^{T}
\]

(15)

Let us define two vectors $\omega$ and $v$ as:

\[
\omega = [l_{41}, l_{42}, l_{43}]^{T} \quad v = [l_{23}, l_{31}, l_{12}, l_{13}]^{T}
\]

(16)

It means that $\omega$ represents the “directional vector”, while $v$ represents the “positional vector”. It can be seen that for the Euclidean space ($w = 1$) we get:

\[
X_2 - X_1 = \omega \quad X_1 \times X_2 = v
\]

(17)

where $X_1 = \left( \frac{x_1}{w_1}, \frac{y_1}{w_1}, \frac{z_1}{w_1} \right)$ are coordinates of points in the Euclidean space.

The line $p$ given by two points in the homogeneous coordinates is then given as:

\[
x(t) = \frac{v \times \omega}{\|\omega\|^2} + \omega t
\]

(18)

where $x = (X, Y, Z)$ are coordinates of points on the line $p$.

Due to the principle of duality in $E^3$ we can exchange “point” and “plane” in the Eq.(14). As the planes are given as:

\[
\rho_1 = [a_1, b_1, c_1; d_1]^{T} \quad [n_1^{T}; d_1]^{T}
\]

\[
\rho_2 = [a_2, b_2, c_2; d_2]^{T} \quad [n_2^{T}; d_2]^{T}
\]

(19)

Now, the matrix $L$ is defined as

\[
L = \rho_1 \rho_2^{T} - \rho_2 \rho_1^{T}
\]

(20)

and the rest is the same.

However, an intersection of two planes is the case also often solved in computer graphics and vision. Unfortunately in many cases available solutions are not robust or formulas used are neither simple, like above, nor convenient for GPU use. This approach also requires normalization of the $\omega$ vector and division operations.

In the following a new formulation of two plane intersection is presented and as the projective space is used for formulation and the solution is quite simple.
VII. INTERSECTION OF TWO PLANES

Let us consider again two planes given in implicit form, i.e. \( \mathbf{p}_i^T \mathbf{x} = 0 \), \( i = 1, 2 \) and given points \( \mathbf{x}_i = [x_i, y_i, z_i; w_i]^T \) in homogeneous coordinates, see Fig.1.

\[
\begin{align*}
\mathbf{p}_1 &= [a_1, b_1, c_1; d_1]^T \\
\mathbf{p}_2 &= [a_2, b_2, c_2; d_2]^T \\
&= [n_1^T; d_1]^T = [n_2^T; d_2]^T
\end{align*}
\]

(21)

The directional vector \( \mathbf{s} \) of the line \( p \) given as an intersection of those two planes is again given as:

\[
\mathbf{s} = \mathbf{n}_1 \times \mathbf{n}_2 = [a_3, b_3, c_3]^T
\]

(22)

An angle \( \varphi \) between two planes \( \rho_1 \) and \( \rho_2 \) is given as:

\[
\cos \varphi = \frac{n_1^T n_2}{\|n_1\| \|n_2\|}
\]

(23)

If the “standard” formula is used, the \( \cos \varphi \) value is to be used for detection of close to singular cases or:

\[
\cos^2 \varphi = \frac{(n_1^T n_2)^2}{(n_1^T n_1)(n_2^T n_2)}
\]

(24)

as this eliminates \( \sqrt{\text{function computation}} \) in the vector normalization as well.

Now, let us assume a plane \( \rho_0 \) passing the origin of the coordinate system having as a normal vector the vector \( \mathbf{s} \), see Fig.1. Therefore the point \( \mathbf{x}_0 \) is an intersection of three planes \( \rho_0, \rho_1, \rho_2 \) can be computed as a direct application of the principle duality in the \( \rho^3 \) case as a straightforward application of the Eq.(8) and Eq.(9):

\[
\mathbf{x}_0 = \mathbf{p}_0 \times \mathbf{p}_1 \times \mathbf{p}_2
\]

(25)

From the Eq.(25) we can write for the \( \mathbf{x}_0 \) point using homogeneous coordinates \( \mathbf{x}_0 = [x_0, y_0, z_0; w_0]^T \):

\[
\begin{align*}
x_0 &= +d_2 \begin{bmatrix} b_0 & c_0 & 0 \\
-1 & b_2 & 0 \\
-1 & c_2 & 0 \\
-1 & c_2 \\
-1 & c_2 & 1 \end{bmatrix} \\
y_0 &= -\left\{ +d_2 \begin{bmatrix} a_0 & c_0 & 0 \\
-1 & a_2 & 0 \\
-1 & a_2 & 0 \\
-1 & a_2 & 1 \end{bmatrix} \right\} \\
z_0 &= +d_2 \begin{bmatrix} a_0 & b_0 & 0 \\
a_1 & b_1 & c_1 \\
-1 & a_1 & b_1 & c_1 \end{bmatrix} \\
w_0 &= -\begin{bmatrix} a_0 & b_0 & c_0 \\
-1 & b_2 & c_2 \\
-1 & b_2 & c_2 \\
-1 & b_2 & c_2 \\
-1 & b_2 & c_2 \end{bmatrix}
\end{align*}
\]

(26)

The point \( \mathbf{x}_0 \) is the closest point on the line \( p \) to the origin of the coordinate system. As the computation is made in the projective space no division operation is needed.

The actual decision on “singularity” of computation is postponed. As a determinant is multilinear, if clear that normalization of plane’s equations to \( \|n_i\| = 1 \) is not needed and it is not a way how to increase stability of computations.

Even more we are getting a formula, which can be used for symbolic manipulation and computation as well.

Derivation of this formula is simple, fast and it is well suited for the GPU or SSE application as the cross product is an GPU instruction, see Appendix A

VIII. BARYCENTRIC COORDINATES

Barycentric coordinates are often used in many applications, not only in geometry. Barycentric coordinates computation leads to a solution of a system of linear equations. However it can be shown, that a solution of a linear system equations is equivalent to a cross product [7, 8, 10]. Therefore it is possible to compute barycentric coordinates using cross product for GPU oriented applications.

Let us consider the \( E^1 \) case, i.e. interpolation between two points, and vector:

\[
\mathbf{x} = [x_1, x_2; x]^T \quad \mathbf{w} = [w_1, w_2; w]^T
\]

(27)

Then the projective barycentric coordinates are given as:

\[
\lambda_1 = -\frac{\xi_1}{\xi_w} \quad \lambda_2 = -\frac{\xi_2}{\xi_w}
\]

(29)

The Euclidean barycentric coordinates are given as:

\[
\xi = \mathbf{x} \times \mathbf{w} = [\xi_1, \xi_2, \xi_3]^T
\]

(28)

Let us consider the \( E^2 \) case, i.e. interpolation between three points of the given triangle, and vectors:

\[
\mathbf{x} = [x_1, x_2, x_3; x]^T \quad \mathbf{y} = [y_1, y_2, y_3; y]^T \quad \mathbf{w} = [w_1, w_2, w_3; w]^T
\]

(30)

The projective barycentric coordinates are given as:

\[
\xi = \mathbf{x} \times \mathbf{y} \times \mathbf{w} = [\xi_1, \xi_2, \xi_3, \xi_w]^T
\]

(31)

The Euclidean barycentric coordinates are given as:

\[
\lambda_1 = -\frac{\xi_1}{\xi_w} \quad \lambda_2 = -\frac{\xi_2}{\xi_w} \quad \lambda_3 = -\frac{\xi_3}{\xi_w}
\]

(32)

Let us consider the \( E^3 \) case, i.e. interpolation between four points of the given tetrahedron, and vectors:

\[
\mathbf{x} = [x_1, x_2, x_3, x_4; x]^T \quad \mathbf{y} = [y_1, y_2, y_3, y_4; y]^T \quad \mathbf{z} = [z_1, z_2, z_3, z_4; z]^T \quad \mathbf{w} = [w_1, w_2, w_3, w_4; w]^T
\]

(33)

Then projective barycentric coordinates are given as:

\[
\xi = \mathbf{x} \times \mathbf{y} \times \mathbf{z} \times \mathbf{w} = [\xi_1, \xi_2, \xi_3, \xi_4, \xi_w]^T
\]

(34)

The Euclidean barycentric coordinates are given as:

\[
\lambda_1 = -\frac{\xi_1}{\xi_w} \quad \lambda_2 = -\frac{\xi_2}{\xi_w} \quad \lambda_3 = -\frac{\xi_3}{\xi_w}
\]

(35)

How simple and elegant solution!

It can be seen that the presented computation of barycentric coordinates is simple, convenient for GPU or SSE application. Even more, as we have assumed from the very beginning, there is no need to convert coordinates of points from the homogeneous coordinates to the Euclidean coordinates. As a direct consequence of that is that we save lot of division operations and also increase robustness of the computation.
IX. THE CLOSEST POINT TO AN INTERSECTION OF PLANES

Finding the closest point on the intersection of two planes to the given point $\xi$ is not an easy problem. The usual solution is based on application of the Lagrange multipliers [6] leading to a solution of a system of linear equations with a matrix $5 \times 5$. This solution is computationally costly and not convenient for the GPU applications. Using projective representation we can easily solve the problem as follows.

Let us assume given two planes:

$$\rho_1 = [a_1, b_1, c_1; d_1]^T \quad \rho_2 = [a_2, b_2, c_2; d_2]^T$$

with normal vectors:

$$n_1 = [a_1, b_1, c_1]^T \quad n_2 = [a_2, b_2, c_2]^T$$

Directional vector of a line given as an intersection of two planes $\rho_1$ and $\rho_2$ is determined as

$$s = n_1 \times n_2$$

and the “starting” point $x_0$ is determined as recently stated

$$x_0 = \rho_1 \times \rho_2 \times \rho_0$$

and the plane $\rho_0$ is passing the origin with a normal vector $s$, i.e. the plane is defined by $\rho_0 = [a_0, b_0, c_0; 0]^T$.

The solution is now quite simple applying the following steps:

- Translate planes $\rho_1$ and $\rho_2$ so the given point $\xi$ is in the origin, transformation matrix is $T$
- Compute intersection of two planes, i.e. determine the directional vector $s$ of the line and the point $x_0$
- Translate the computed point $x_0$ using $T^{-1}$

The point $x_0$ is the closest point of the line to the given point $\xi$. Even more, the parameter value $t$ to this point $x_0$ is $t = 0$.

Again – an elegant solution, simple formula supporting matrix-vector architectures like GPU and parallel processing.

X. CONCLUSION

In this paper we have presented a new robust approach for a solution of selected geometrical problems using projective representation. This approach offers much simpler formulation, algorithms based on matrix-vector multiplications. Resulting formulas are convenient for GPU and SSE applications with achievable significant speedup as well.

XI. ACKNOWLEDGMENT

The author thanks to anonymous reviewers for their critical comments and hints that improved this paper significantly. Thanks also belong to students and colleagues at the University of West Bohemia, Plzen for their suggestions.

The project was supported by the Ministry of Education of the Czech Republic, projects No.LG13047 and LH12181.

XII. REFERENCES


APPENDIX

The cross product in 4D is defined as

$$x_1 \times x_2 \times x_3 = \begin{vmatrix} i & j & k & l \\ x_1 & y_1 & z_1 & w_1 \\ x_2 & y_2 & z_2 & w_2 \\ x_3 & y_3 & z_3 & w_3 \end{vmatrix}$$

and can be implemented in Cg/HLSL on a GPU as follows:

```c
float4 cross_4D(float4 x1, float4 x2, float4 x3) {
    float4 a;
    a.x = dot(x1.yzw, cross(x2.yzw, x3.yzw));
    a.y = - dot(x1.xzw, cross(x2.xzw, x3.xzw));
    a.z = dot(x1.xyz, cross(x2.xyz, x3.xyz));
    a.w = - dot(x1.xyz, cross(x2.xyz, x3.xyz));
    return a;
}
```
or more compactly as

```c
float4 cross_4D(float4 x1, float4 x2, float4 x3) {
    return (dot(x1.yzw, cross(x2.yzw, x3.yzw)),
            - dot(x1.xzw, cross(x2.xzw, x3.xzw)),
            dot(x1.xyz, cross(x2.xyz, x3.xyz)),
            - dot(x1.xyz, cross(x2.xyz, x3.xyz)));
```