

---

# DOUBLE-PRECISION FLOATING-POINT DATA VISUALIZATIONS USING VULKAN API

---

Nezihe Sözen  
Siemens Turkiye  
Istanbul  
nezihe.sozen@siemens.com

## ABSTRACT

Proper representation of data in graphical visualizations becomes challenging when high accuracy in data types is required, especially in those situations where the difference between double-precision floating-point and single-precision floating-point values makes a significant difference. Some of the limitations of using single-precision over double-precision include lesser accuracy, which accumulates errors over time, and poor modeling of large or small numbers. In such scenarios, emulated double precision is often used as a solution. The proposed methodology uses a modern GPU pipeline and graphics library API specifications to use native double precision. In this research, the approach is implemented using the Vulkan API, C++, and GLSL. Experimental evaluation with a series of experiments on 2D and 3D point datasets is proposed to indicate the effectiveness of the approach. This evaluates performance comparisons between native double-precision implementations against their emulated double-precision approaches with respect to rendering performance and accuracy. This study provides insight into the benefits of using native double-precision in graphical applications, denoting limitations and problems with emulated double-precision usages. These results improve the general understanding of the precision involved in graphical visualizations and assist developers in making decisions about which precision methods to use during their applications.

**Keywords** IEEE-754 · Double-precision · Emulated-precision · Shaders · Shading Languages

## 1 Introduction

Nowadays, computer graphics tools play a significant role in high-precision data modeling and display for several themes within science and engineering. They make an impact on decision-making and consequent results in very diverse domains, from scientific research to gaming. High-precision molecular dynamic simulation [1], an attempt at film-game integration [2], next-generation AI-based games [3], the open-source dynamic simulation framework YADE [4], three-dimensional reconstruction using [5], and ray tracing of CAD-based simulation [6] are some of the studies which underpin the need.

In graphical models, precision and accuracy are vital and are mostly used in scientific computing, medical imaging, and engineering simulations. These models guarantee to convey complex and comprehensive information precisely and accurately. The ancillary problem here is obtaining correct data types, mainly distinguishing double-precision from single-precision floating-point types, which turn out to be very influential. This kind of technology is, however, vision-limited to some level of accuracy, hence less valuable if even higher levels of precision are called for. It is, hence, important to pick an alternative that offers precision and, importantly, the level of accuracy required by the application or the project.

Processing and visualizing double-precision data present several challenges, such as computational intensity [7], hardware and software support [8] [9], and rounding errors [10] [11], mainly due to the complexity and precision requirements of this type of data.

The choice of API in high-accuracy applications is based on each API's appropriateness. With changing application requirements, the suitability of each API changes: OpenGL [12], Vulkan [13], DirectX [14], OpenCL [15], and CUDA

[16]. Vulkan is a much newer API compared to OpenGL and gives a developer much more acute control over GPU processes, significantly improving rendering performance and speed. Direct3D is the most frequently used graphics API for game and multimedia applications that run on Windows systems. OpenCL is a portable programming language that can be run on various devices, hence the reason why OpenCL applications are highly portable across a wide variation of hardware like a GPU and CPU.

In the case of high-precision visual representation using Vulkan API and GLSL [17], complicated technical details shall be addressed in the analysis of double-precision and single-precision floating-point data. One major reason could be the accuracy differences that, in turn, compromise both performance and memory consumption. In old generations of GPU hardware, a representation of double-precision data often emulated single-precision floating-point numbers. Modern GPUs have, however, come with support for double-precision data in their design. This research is based on recent versions of GLSL supporting both single and double precision. The evaluations taken into consideration for this research are to frame the differences between single-precision and double-precision data representations. In this paper, the approaches are compared with respect to visualization performance, where a rendering operation differs in both. This comparison will provide important information about which graphics API is more suitable for applications requiring high precision and will provide guidance for current applications in the field of graphics programming. Additionally, current challenges in processing and visualizing double-precision floating-point data and methods to overcome these challenges will be discussed.

In view of the requirement of accuracy in graphics visualizations, primarily due to real-world applications in science and engineering, this research is done to find out how double-precision floating-point data visualizations realize performance gains using the Vulkan API. The specific research questions that guided this study were as follows:

- RQ 1. How does the performance of native double-precision floating-point implementations compare to emulated double-precision implementations in Vulkan for rendering 2D and 3D points datasets?
- RQ 2. How does the scalability of double-precision floating-point data visualization in Vulkan API hold up with increasing dataset sizes?

## 2 Related Work

This section reviews previous studies on the use of double- and single-precision floating-point values in graphics processing units (GPUs) and discusses the findings, similarities, and differences of these studies in the context of current research. Various floating-point approaches have been developed using graphics APIs and game engines. However, this section examines closely related approaches like emulated double-precision, hardware and software-based double-precision, and extended precision.

Da Graça & Defour [18] demonstrated a 44-bit solution emulating floating-point formats and corresponding operations, which increased the precision for applications that need more than the single precision, complemented with detailed performance and accuracy results. This implementation enabled straightforward and efficient operations for adding, multiplying, and storing floating-point numbers. It is shown that the research in compensated algorithms with float-float representation runs more efficiently for comparable accuracy, and adapting these algorithms to the GPU constitutes a significant part of future research.

Thall [19] introduced 'doublefloats' for extended precision representation of floating point numbers for GPU computation. This representation, while improving accuracy without performance degradation on a GPU, comes at the cost of resources by exploiting the inherent parallelism in it. Doublefloats are constructed from unevaluated sums of 32-bit floats and deliver a precision of 48 significant bits. This approach—crucial for very high precision applications—necessarily restricted the use of GPU hardware. This shows, with the help of the Mandelbrot Set Explorer, both the utility of doublefloats and some potential applications in simulation, scientific computing, and image analysis.

In the OpenSpace [20] study, the limitations of floating-point numbers in computers and the various methods developed to ensure accurate representation of large-scale astronomical data are discussed. With the method it offers, OpenSpace manages to solve precision problems by using Dynamic Scene Graph and rendering objects at the correct distances relative to the camera in cases where single-precision floating-point numbers are insufficient. This method enables accurate and efficient visualization of large-scale astronomical data and minimizes floating-point precision problems.

Dally et al. [21] review the progress of GPUs from special-purpose hardware for 3D graphics to powerful programmable processors applied in HPC and deep learning. It reflects all significant steps of this development, including the creation of CUDA, the introduction of double-precision floating-point arithmetic, and other innovations like Tensor Cores that have increased the performance and flexibility of contemporary GPUs manyfold. These results depict that in the future, a GPU will keep evolving to provide high performance and support many applications.

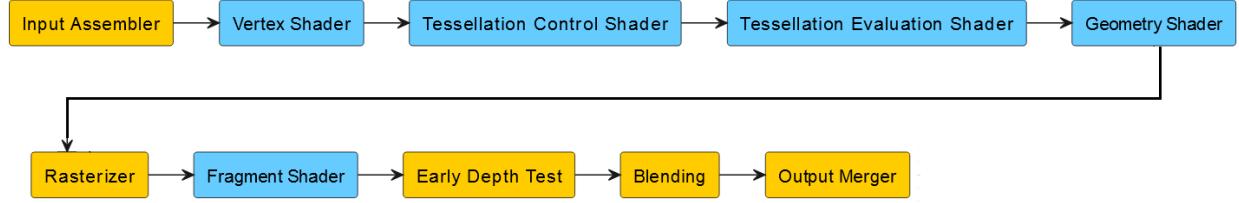


Figure 1: Beginning with the input assembler, which takes the vertex data to assemble vertices into primitives, the pipeline is followed by a vertex shader for geometric transformations. Then, there is a tessellation control shader that performs the division of the surface, followed by the tessellation evaluation shader, refining the vertices. Subsequently, there is a geometry shader for generating or modifying geometry. Thereafter, it proceeds to the rasterizer, which projects 3D primitives onto the 2D screen. Next up is a fragment shader to compute pixel attributes, an early depth test optimization by discarding occluded fragments, the blending stage, which combines fragment colors, among other things, for transparent effects, and an output merger that finally writes the image in the frame buffer for display.

Kaufmann et al. [22] addresses the challenges and limitations inherent in real-time physics simulations in large-scale environments, primarily due to the imprecision of single-precision floating-point calculations. The authors solve a limitation where traditional physics engines still rely on single-precision floating-point numbers by proposing a system in which the subdivision of the simulation world into independent sectors takes place, and these sectors are allocated dynamically. It drastically reduces the occurrence of precision errors through cloning at sector boundaries, ensuring very consistent and accurate interaction across these sectors. It has hugely improved the precision and efficiency of real-time physics simulation in large-scale virtual environments by dividing the world into independently simulated, dynamically allocated sectors and using a cloning system to maintain accurate interactions at sector boundaries.

The progress report [23] on the Godot Engine reviews challenges to render big worlds in games with single-precision floating-point numbers that lead to precision errors and jerky motions. The report explores solutions such as using double-precision in calculations but handling its impracticability again due to the limits of GPU. The final solution considered is emulating double precision by using two single precision floats for specific matrix calculations, where they preserve the accuracy and precision without much penalty in performance.

### 3 Background and Motivations

In this section, a number of the fundamental technologies and methodologies of computer graphics will be introduced, including the pipelines of modern Graphics Processing Units (GPUs), the functionality of key graphics APIs, for instance, OpenGL and Vulkan, the use of the Graphics Shader Language (GLSL), the key differences between single and double precision operations, and finally, common issues with single precision floating point computations.

#### 3.1 GPU Pipeline

The Graphics Processing Unit pipeline has become a critical component in modern computing, having extensive applications in graphics rendering and general-purpose computing. This paper explores several aspects of GPU pipeline advancements and applications as described in recent literature.

All pipeline stages concerning rasterization, pixel processing, and abstract geometry processing are implemented on the GPU. Internally, these are divided into a range of hardware stages with differing levels of programmability or configurability. The API provides an access method for the programmer to the logical model of a GPU; the actual implementation of this conceptual process in hardware is left to the manufacturer. The pipeline of the GPU processes graphical input through many phases, from the original vertices to the final pixel rendering, with differing degrees of programmability. The fully programmable vertex shader stage is responsible for perspective projection, lighting, model space to view space transformations and vertex shading. Another programmable stage is Geometry Shader, which deals with entire primitives to generate or to modify them and construct complex effects, including particle systems and shadow volumes.

Stream output reuses processed data and is thus useful for effects like hair rendering. Fixed-function stages are triangle preparation and traversal, where triangles are prepared and rasterized into pieces; screen mapping, which translates the vertices from clip space into screen space; and clipping, which cuts triangles beyond the viewing frustum. There may be removal of occluded pieces based on the early z-test step, which is varied across GPUs for increased efficiency. The

programmable pixel shader processes every fragment to provide texture and color effects. The final step, raster ops or blending, is where colors are combined, and other pixel tests, like depth and alpha testing, are managed. This pipeline consists of both fixed and programmable steps; both are required to efficiently render complex images. [24] [25].

### 3.2 Vulkan API

Vulkan is the next-generation, efficient, and cross-platform graphics and compute API for enabling access to modern GPUs in today's devices—PCs, consoles, mobile phones, and even embedded platforms. The Vulkan API has been designed to give much more direct control over the GPU, thus allowing finer-grained optimizations and efficient usage of the GPU. Vulkan significantly reduces driver overhead compared to older graphics APIs. Such overhead may yield great performance, particularly in CPU-bound applications. It also designs the API to be more predictable and with fewer errors, clear performance benefits from keeping the GPU busier, producing fewer bottlenecks than those caused by the CPU [26]. Vulkan is characterized by its verbosity and fragility, but it provides enhanced control, a streamlined threading architecture, and superior performance. It provides functionality for transport, computation, and graphics and may be chosen as an option [27]. The Vulkan Specification mandates a host environment with runtime support for 8-bit, 32, and 64-bit signed and unsigned twos-complement integers, 32- and 64-bit floating-point types satisfying range and precision constraints, and ensuring their representation and endianness match those on every supported physical device.

### 3.3 GLSL

GLSL, often known as the OpenGL Shading Language, has a crucial function in contemporary computer graphics by enabling programmatic control over the graphical processing pipeline. This programming language provides a whole set of tools with which developers can create very flexible shaders, improving the ability to develop complex, dynamic graphical effects vastly in any real-time application. The GLSL has become central to a modern graphics programmer due to its wide application in different domains, from game development and virtual reality to scientific visualization. Recent developments in this area include the integration of GLSL with all major graphics APIs and its application in parallel computing cases. Unlocking the doors of new frontiers in graphical rendering and visualization is possible with GLSL. Thereon, the further development of GLSL and the expansion of the practical spectrum of its implementation essentially volatily characterized the area of computer graphics when new challenges and prospects succeed one another. [28] [29]

### 3.4 An Explanation of the Distinction Between Single and Double Precision

IEEE 754 floating-point format represents all the standards; it includes 32-bit single precision, 64-bit double precision, and an extended precision format. Each format includes a sign bit, an exponent section, and the mantissa part (fraction) [30].

The single-precision floating-point format defined in IEEE Std 754-2019 utilizes a 32-bit (4-byte) structure. This format consists of a 1-bit sign bit, telling whether a number is positive or negative, an 8-bit exponent defining the scale of the number adjusted by a predefined "bias" value, and a 23-bit fraction representing the significant or mantissa part of the number. The single-precision format provides an accuracy of about 7 decimal digits while it covers a very wide range of values. This would normally be used in cases where speed is very essential and very fine precision is not required, for example, in the processing of graphics or audio [31] [32].

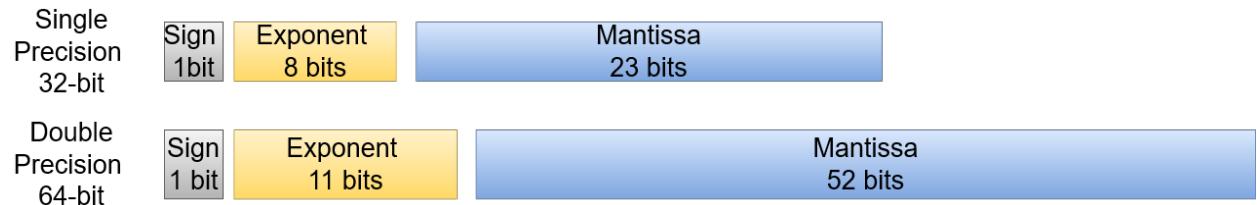


Figure 2: This image is a bit-layout of single and double-precision floating-point numbers, as represented in accordance with the IEEE 754 standard. Single precision number would be 32 bits long. Bits needed for this: 1 bit for the sign, 8 bits for exponent, and 23 bits for mantissa. Double precision number: it is 64 bits; 1 bit for the sign, 11 bits for exponent, and 52 bits for mantissa. It has a wider range and is more accurate in representing floating-point numbers.

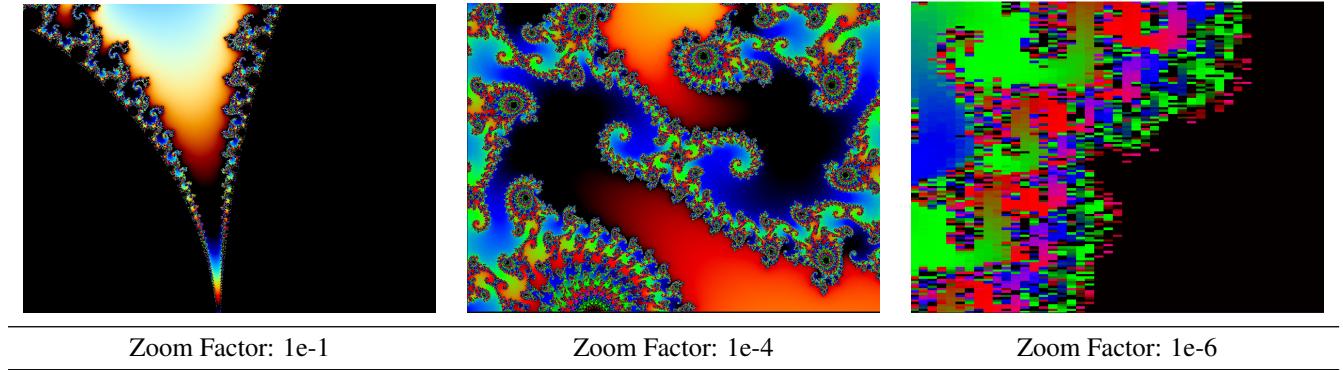


Table 1: Visualization of the Mandelbrot Set at the Point  $\text{Re}=-0.7436450$ ,  $\text{Im}=0.13182590$  with Different Zoom Factors. The images display the fractal structure at zoom levels of  $1\text{e}-1$ ,  $1\text{e}-4$ , and  $1\text{e}-6$ , showcasing the intricate details at progressively finer scales. Precision concerns and pixelization problems can be seen when the zoom is  $1\text{e}-6$

The format, otherwise known as double-precision floating point, is 64-bit (8-byte) in size. Much like the single precision, the format contains a 1-bit sign bit but reserves 11 bits for the exponent and 52 bits for the fraction. These give double precision a much greater range and much higher precision. This format provides an accuracy of about 15–16 decimal digits and is preferably used where high accuracy is required, like in scientific computations and precision engineering tasks.

The IEEE 754 standard standardizes the way of representation and processing of floating-point numbers within a computer system. This creates consistency and reliability for numerical computations. The standard provides for the accuracy of numerical operations across different systems and by different languages. This is very important in scenarios where different applications and cross-platform are required.

### 3.5 Problems with Single-precision floating point

The prevailing solution to obtain high precision in graphical visualizations is using double-precision floating-point values. Doubles give a maximum of about 15 to 16 decimal digits of precision and, at the same time, offer far greater range than single-precision, floating-point values. This rise in precision and range drastically reduces errors due to rounding, positioning, accumulation, overflow, underflow, and limitation [33]. In order to understand more easily the problems caused by single-precision floating point, the Mandelbrot Set [34] formula has been used and rendered. A Mandelbrot set is a mathematical set that repeats in a certain way in the complex plane.

#### 3.5.1 Rounding Issue

Inaccuracies can occur due to the rounding issues. The first rounding may be toward a midpoint which then gets rounded again, moving it further from the closest correct value [35]. Consider the number 3.1415926 that is represented base-10. The higher precision will round to three decimal places while the lower precision will round to the nearest integer. The higher precision rounds 3.1415926 to 3.142. When this result is then rounded to a lower precision it becomes 3. When 3.1415926 is rounded directly to the nearest integer, omitting the intermediate step the answer is again 3 so in this case there is no inaccuracy. A slight modification of the situation can make double rounding significant. For instance, if the exact value was 3.6515926 then rounding first to the higher precision gives 3.652 and further rounding to the lower precision gives 4. Rounding directly from 3.6515926 to the nearest integer gives 4 also so differences need not appear in every case, yet are of vital significance in the vicinities of some number values [36] [37].

#### 3.5.2 Limited Precision

In general, limited precision refers to the extent of precision that can be attained in any computation or measurement. For computational and scientific purposes, it is extremely important to be valid in domains as diverse as numerical analysis and engineering since the accuracy and reliability of a result are determined by its precision. Single-precision floating-point values provide an approximate 7 decimal digits of precision [38]. This causes insufficient precision for more complex visualizations, and it may introduce substantial rounding errors with very large or small-scale numbers and lead to loss of details.

### 3.5.3 Range Limitations

This refers to being restricted to some values within a range that the computer system or a model of computation is capable of representing. These may be the largest or smallest values a number takes, either positive or negative. Fundamentally, they are directly proportional to the size of the data type used and are bound by limitations in the handling of very large or very small numbers. One issue with floats is that they have a restricted range, which can result in overflows or underflows. This could be manifested in graphical contexts as visual anomalies or inaccuracies during rendering [30] [39].

## 4 Comparative Analysis between Single-precision Floating Point and Double-precision Floating-point Implementations

Most modern graphics and compute target applications rely on floating-point operation accuracy, which brings about dramatic performance and quality impacts. On the other side, single-precision versus double-precision floating-point implementations represent a compromise in both computational performance and accuracy for high-performance graphics rendering. This section describes a comparative study of both under the Vulkan API framework, pointing out each of their benefits and tradeoffs. In conjunction with the explanation, the diagram provided shows the Vulkan-based application — the full extent of which will initialize and manage GPU resources for rendering both 2D and 3D point datasets to cover in detail how the choice of precisions affects the resultant rendering and performance metrics.

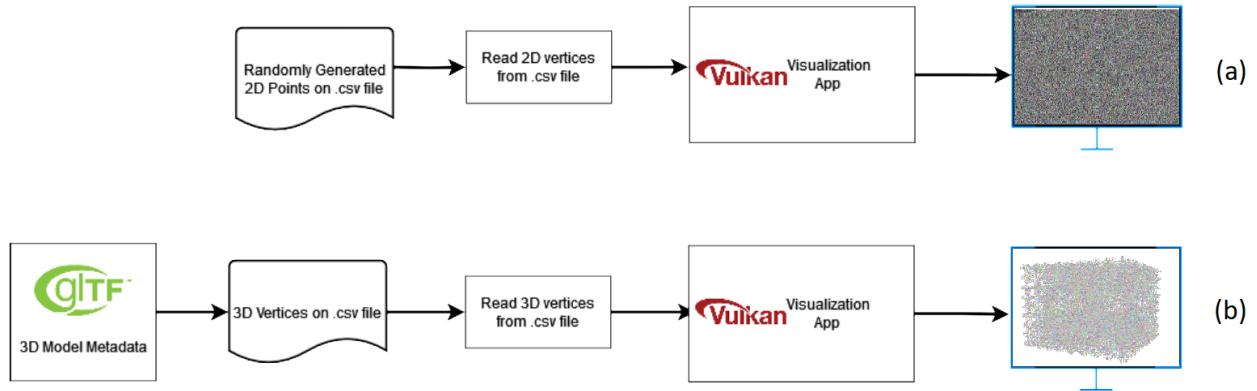


Figure 3: 2D and 3D Point Data Visualization Process: a) 2D Point Data Visualization: The Vulkan-based visualization application reads a .csv file consisting of randomly generated 2D points: in this data set, every point is represented by the x and y coordinates. The CSV file is then read by the Vulkan-based visualization application, which visualizes it on the screen. (b) 3D Point Data Visualization: Downloaded model data in glTF/GLB format and further converted it into a .csv file, including the x, y, and z coordinates for each point. Feeding this .csv file into the Vulkan-based visualization application would draw the 3D points onto the screen. As soon as three-dimensional points can be visualized, more complex structures and models represented with data will easily be analyzed and understood.

The first step is to create a Vulkan instance. A Vulkan instance is an instance that lets the application interact with the Vulkan API. Immediately following the creation of the Vulkan instance, a debug messenger is created for debugging during the development process. Following this, inter-functioning the windowing system with the surface creation is done. All core Vulkan API elements have been created at this stage, which would now allow an application to use the GPU.

Another core stage in managing the rendering of graphics using Vulkan is logical device setup. This logical device provides an interface with the GPU from the application. It enables a number of commands to be executed. At this stage, Command Buffers are allocated for the storage of rendering commands. Synchronization objects, like Semaphores and Fences, are created to control the synchronicity in the execution of commands. A Vertex Buffer is also allocated for vertex data. These parts ensure that all rendering processes go through smooth and concurrently.

This can do high-performance graphics operations with the Vulkan API by creating a Graphics Pipeline. It describes the process of rendering—a pipeline starting from processing vertices up to fragment shading. Shader Modules are loaded into the pipeline to handle certain parts of the rendering process. Rendering output is controlled by the arrangement of

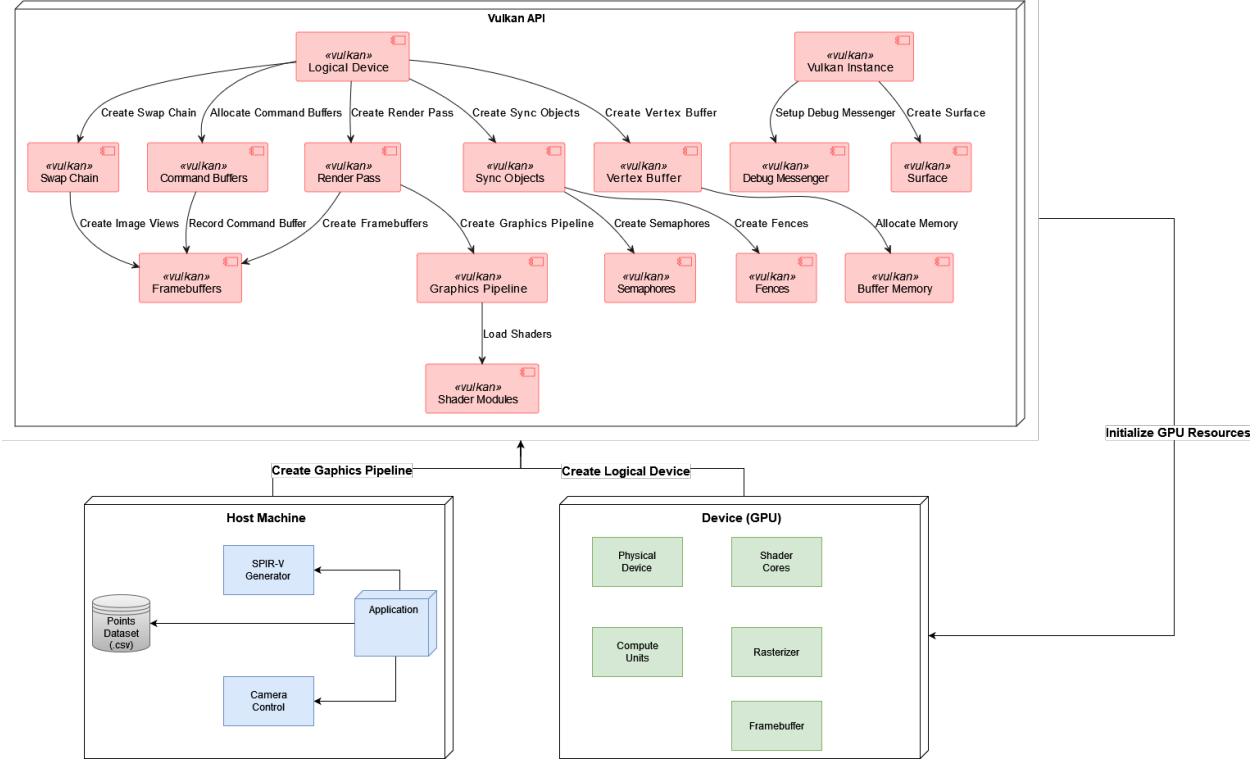


Figure 4: This diagram shows initializing and managing GPU resources using the Vulkan API to visualize 2D and 3D point datasets. Vulkan exposes a low-level, general-purpose graphics API that is conceived to offer direct control of the GPU resources to ensure both high performance and flexibility. It details all the processes, from the initialization of the GPU resources to the creation of the graphics pipeline. The graphics pipeline and the initialization of GPU resources are showcased. This diagram provides a comprehensive overview of the stages and interactions involved in setting up and utilizing Vulkan for rendering within the application.

Render Passes and Framebuffers. Therefore with such structure, it is possible to execute efficiently complex graphics operations.

It involves all rendering commands, once drawing 2D and 3D points, updating the Framebuffers, and, finally, memory allocation to manage one buffer—the so-called Buffer Memory—which contains vertex data and other related information regarding the rendering process. In this step, it will be finalized how an application is going to handle its rendering process to present the final output. Finally, memory is allocated to manage one buffer—the so-called Buffer Memory—which contains vertex data and other related information regarding the rendering process. Efficient memory management has been taken into consideration to ensure that GUI resources are used effectively.

Control and data flow between Host Machine and GPU Host machine components: Points Dataset, SPIR-V Generator, Application, Camera Control. A Points Dataset, prepared, typically, in CSV format, contains 2D and 3D points. A SPIR-V Generator generates a low-level representation from an input high-level source representation, such as high-level shaders, for execution on a GPU. The Application is in charge of the entire rendering process and communicates with the Vulkan API. Camera Control deals with how the visualization is to be viewed.

The device components include the GPU, Physical Device, Shader Cores, Compute Units, Rasterizer, and Framebuffer. Each of these is one of the key elements in a rendering pipeline, and together they allow high performance in graphics-oriented operations.

Once the shaders are written in GLSL, they need to be compiled into SPIR-V, which is the intermediate representation used by Vulkan. The compilation process ensures that the shaders are optimized and can be executed efficiently on the GPU. The compilation can be done using tools like ‘glslangValidator’. The compiled SPIR-V shaders are then integrated into the Vulkan pipeline [40] [41]. The steps include creating shader modules, setting up the pipeline, and binding the necessary resources. Below is an example of how the shader modules are created and integrated:

The created shader module is then used in the graphics pipeline to execute the vertex and fragment shaders. By providing a detailed explanation of the shader code, its compilation, and integration into the Vulkan pipeline, this section offers a comprehensive understanding of how native double-precision floating-point operations are utilized in Vulkan applications.

#### 4.1 Emulated Double-precision Floating-point

To compare the findings, case studies exemplifying emulated double-precision and native double-precision are presented. Development for emulated precision was based on David H. Bailey's approach in the DSFUN90 library [42].

---

**Algorithm 1** Emulate Double Precision

---

**Require:**  $value \in \mathbb{R}$  ▷ A double precision floating point number  
**Ensure:**  $low, high \in \mathbb{R}$  ▷ Two single precision floating point numbers representing the double precision input

- 1: **procedure** EMULATEDOUBLEPRECISION( $value, low, high$ )
- 2:      $high \leftarrow (\text{float})value$
- 3:      $highDouble \leftarrow (\text{double})high$
- 4:      $low \leftarrow (\text{float})(value - highDouble)$
- 5: **end procedure**
- 6: **return** ( $low, high$ )

---

This algorithm aims to store a double precision floating point number by dividing it into two single precision floating point numbers. First, the variable value is converted to a single precision floating point number and assigned to the variable high. This step is to obtain a lower precision representation of value. The high value is then converted back to a double precision number and assigned to the highDouble variable. This conversion is necessary for comparison with the original value. Finally, the difference is calculated by subtracting highDouble from the value, and this difference is assigned to the low variable as a single precision number. Thus, the variables high and low are stored as a two-part representation of the original double-precision value. This method is especially useful when precision is essential and memory savings are required.

Listing 1: Emulated Double-precision Floating-point Vertex Shader Code

```
#version 450
layout(push_constant) uniform PushConstants {
    mat4 mvp;
} pushConstants;
layout(location = 0) in vec3 highPos;
layout(location = 1) in vec3 lowPos;
layout(location = 2) in vec3 highColor;
layout(location = 3) in vec3 lowColor;
layout(location = 0) out vec3 outHighColor;
layout(location = 1) out vec3 outLowColor;
void main() {
    outHighColor = highColor;
    outLowColor = lowColor;
    gl_Position = pc.mvp * vec4(highPos + lowPos, 1.0);
}
```

Listing 2: Emulated Double-precision Floating-point Fragment Shader Code

```
#version 450
layout(location = 0) in vec3 outHighColor;
layout(location = 1) in vec3 outLowColor;
layout(location = 0) out vec4 fragColor;
void main() {
    vec3 fullColor = outHighColor + outLowColor;
    fragColor = vec4(fullColor, 1.0);
}
```

## 4.2 Native Double-precision Floating-point

Modern GPUs support double-precision floating-point natively, meaning that high-precision computation can be carried out without emulation. Vulkan API strongly supports native double-precision operations due to its shader language, GLSL (OpenGL Shading Language). In this work, double precision data types have been used, such as double and dvec2, performing vertex and fragment shaders in order to compute the exact values for the requested operations. Later, the shaders were compiled into SPIR-V, which represents the Vulkan Intermediate Representation.

Listing 3: Native Double-precision Floating-point Vertex Shader Code

```
#version 450
#extension GL_ARB_gpu_shader_fp64 : enable
layout(push_constant) uniform PushConstants {
    dmat4 mvp;
} pushConstants;
layout(location = 0) in dvec3 pos;
layout(location = 1) in dvec3 color;
layout(location = 0) out flat dvec3 fragColor;
void main() {
    gl_Position = vec4(pc.mvp * dvec4(pos, 1.0));
    fragColor = color;
}
```

This is a vertex shader written in GLSL, version 4.5.0, with the GL\_ARB\_gpu\_shader\_fp64 extension. The program uses a push constant block called PushConstants, including a member of type dmat4 for an MVP matrix. It also has an input of type dvec3 for position and color, and finally defines a variable named fragColor of data type flat dvec3 to create its output. In the main function, declare a dvec4 vector with a pos data and 1.0. Multiply the vector by the MVP matrix and assign it to gl\_Position. Finally, assign the input color to the fragColor variable.

## 5 Experimental Results

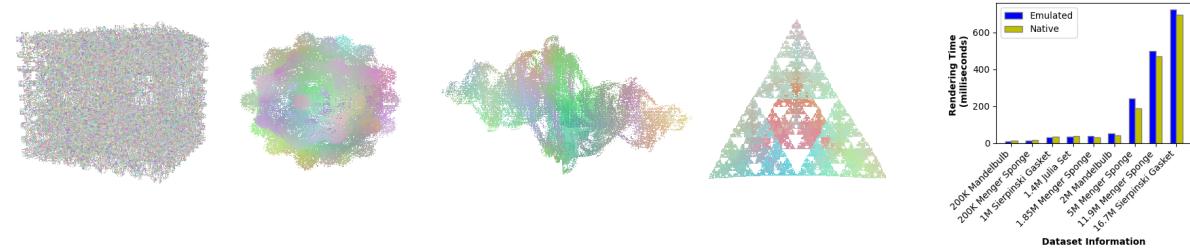


Figure 5: A few of the visualized objects from 3D datasets with double-precision floating-points are included, and the column chart on the right-hand side shows a comparison of rendering times in milliseconds for emulated double-precision and native double-precision.

The proposed method has been implemented using Vulkan API, C++ 20, and GLSL, using only vertex shaders and fragment shaders compiled into SPIR-V code. To be used for testing, 2D point data consisting of 10,000, 100,000, 1,000,000, and 10,000,000 (x, y) coordinates were created randomly and uniquely in the range (-1.0 and 1.0) and then saved in .csv files. 3D point data consisting of 200,000 to 16,700,000 (x,y,z) coordinates were acquired 3D fractal models that created with the open sources libraries. The algorithms and open-source libraries used for dataset production are explained in Appendix A. The same datasets were used for both emulated double-precision experiments and native double-precision experiments. The experiments were conducted using NVIDIA RTX 3090 GPU, Intel(R) Core(TM) i7-6850K @ 3.60GHz CPU, and 44 GB DDR4 RAM hardware components. The framerate was recorded using the RenderDoc [43] v1.33.

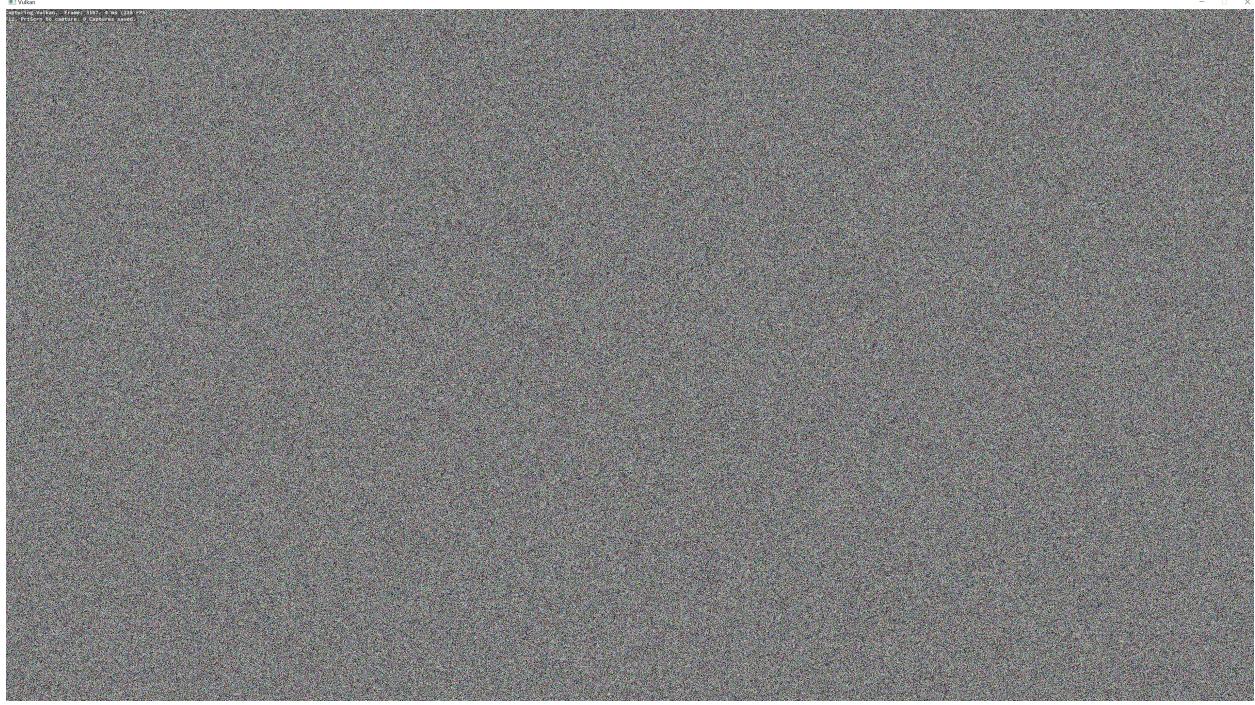


Figure 6: This figure is a screenshot of the application of an example 2D dataset. This dataset has 10 million 2D point data double-precision and the framerate value is calculated as 218 fps.

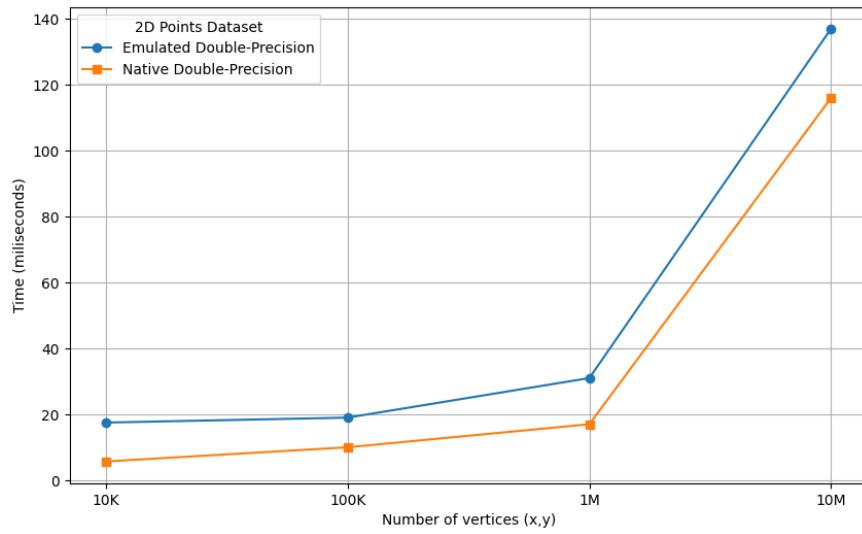


Figure 7: The graph shows the performance of both approaches in time (seconds) for different vertex numbers (10K, 100K, 1M, and 10M). The results reveal that native double-precision calculations are faster overall. The emulated double precision shows a significant performance degradation, especially at high vertex counts (10M). This finding indicates that local double-precision calculations are a more efficient option for calculations requiring high precision. While the emulation performs reasonably well at low vertex counts, local calculations are significantly faster on larger-scale datasets. This is critical for developers who want to perform high-precision and high-performance graphics operations using the Vulkan API.

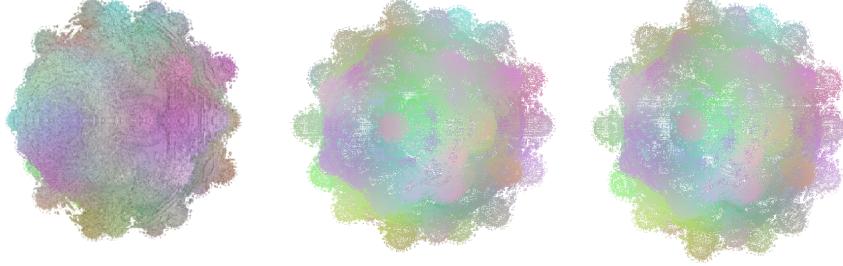


Figure 8: The images prove that three renderings are realized from the same dataset. The left shape represents a triangulated Mandelbulb mesh in GLB/GLTF format; the middle image offers a rendering of those vertices in native double-precision; and the rightmost image, the output of the emulated double-precision implementation run against the same mesh vertices. Comparisons of the performance of these two were done, and because the dataset being fed is the same, the generated renderings are the same. This study compares rendering performance and accuracy for both the native double precision and the emulated double precision methods for rendering.

Experiments demonstrate the advantages of using native double-precision arithmetic within Vulkan. Performance measurements indicate that rendering time for a large dataset improves significantly. The results for native double precision are summarized in the tables below.

Dataset Information	Rendering Time (milliseconds)	Framerate (fps)
200K vertices of 3D Mandelbulb	9.71	729
200K vertices of 3D Menger Sponge	12.9	703
1M vertices of 3D Sierpinski Gasket	32.19	745
1.4M vertices of 3D Julia Set	35.26	717
1.85M vertices of 3D Menger Sponge	36.69	612
2M vertices of 3D Mandelbulb	53.19	687
5M vertices of 3D Menger Sponge	239.73	698
11.9M vertices of 3D Menger Sponge	499.12	468
16.7M vertices of 3D Sierpinski Gasket	722.89	115

Table 2: Performance Results for Emulated Double-precision Floating-point Implementations.

The data in Table 2 indicates that, in general, the emulated double-precision floating point calculations are worse when compared to native calculations. Generally, the render times are longer and frame rates are lower. For example, the render time for a 3D Menger Sponge with 11.9 million vertices was up to 499.12 milliseconds. The highest frame rate observed was 729 fps for 3D Mandelbulb with 200K vertex. Native Double Precision Floating Point Calculations

Dataset Information	Rendering Time (milliseconds)	Framerate (fps)
200K vertices of 3D Mandelbulb	11.87	842
200K vertices of 3D Menger Sponge	14.88	804
1M vertices of 3D Sierpinski Gasket	35.26	789
1.4M vertices of 3D Julia Set	39.26	802
1.85M vertices of 3D Menger Sponge	31.12, 16	763
2M vertices of 3D Mandelbulb	42.68	854
5M vertices of 3D Menger Sponge	189.25	802
11.9M vertices of 3D Menger Sponge	471.29	593
16.7M vertices of 3D Sierpinski Gasket	695.76	330

Table 3: Performance Results for Native Double-precision Floating-point Implementations.

Data in Table 3 clearly indicates that natively provided double precision excels in performance compared to emulated calculations. In most cases, render times are shorter, along with frame rates that become higher. For instance, 3D Menger Sponge containing 11.9 million vertices dropped the render time to as short as 471.29 milliseconds, beating emulated calculations. The highest frame rate among the datasets was achieved with 3D Mandelbulb 200K vertex at 854 fps.

Generally, the performance of natively conducted double-precision computations is better on a large dataset. Though there is droppings performance in emulated computations, the native ones are much more stable. This supports that the local computations are more appropriate when working on large data sets.

## 6 Limitations

The research demonstrates significant improvements in graphic visualization under the Vulkan API using double-precision floating-point data. However, several limitations should be realized. The native double-precision implementation is highly dependent upon modern GPU availability and capability, thereby limiting such an approach to older/less powerful hardware. Although the proposed method has much better scalability with large data sets than emulated double-precision, challenges to efficiently processing extremely large datasets may exist that decrease performance gains with dataset size. Also, the experimental results were obtained on specific hardware and software configurations and may not generalize to other systems; additional benchmarking on a diversity of configurations is thus required. Still, however illuminating the controlled setting with 2D and 3D point datasets was for this method, further testing on real-world data is required to confirm its applicability and performance in practical scenarios since different kinds of data will raise unique issues and possibly differing performance characteristics.

## 7 Conclusion and Future Works

This study was conducted to compare the performance and accuracy of Vulkan API, explicitly focusing on mutable and double-precision data types. The results show enormous implications for data type choices, more so concerning calculation speed and processing time. Double-precision solutions have been executed on the GPU at the moment using double-float and double-double techniques. Test applications consist of 2D points and 3D points; each of them contains double-precision vertex coordinates.

Necessary information is given in the paper hand, and experiments on using double-precision directly have been performed in the case of supporting GPU hardware. The work straight followed a method of double precision ordered by the Khronos Group in OpenGL Shading Language specification 4.5 and Vulkan specification 1.3. This is a different approach compared to traditional emulated precision methods, with no extra processing required. Although this method requires both advanced hardware and software, this is an example of the studies in scientific visualization where precision and performance are desirable together. In this work, the fundamental layer of the visualization of 2D points by uniquely generating random x and y values for these points has been addressed, and 3D points by generating from 3D mesh models with double precision x,y, and z values.

In the future, several experiments can be performed—the ones using actual data to make the application more applicable and accurate. These datasets can be used in tests that verify the visualizing methods. In the case of successful implementation of the most straightforward building block of graphical visualization, that is, point rendering, additional visualization components can be integrated to compile a full-featured visualization library. This will provide the ability to create more complex and informative visualizations, therefore increasing the application’s functionality beyond simple point rendering.

## A Datasets

To provide a more challenging and realistic evaluation, 3D point data were prepared using fractal algorithms, including Mandelbulb, Menger Sponge, Sierpinski Gasket and Julia fractals. The use of fractal models provides a complex and intricate dataset to test the rendering capabilities and performance of the proposed method in three-dimensional space. The diverse point counts ensure that the method’s efficiency and effectiveness can be evaluated under different levels of complexity.

Advanced libraries and algorithms were employed in this research:

- NumPy (<https://numpy.org/>) and Pandas (<https://pandas.pydata.org/>): Used for data handling and manipulation.

- Skimage (<https://scikit-image.org/>): Applied for surface extraction and mesh generation.
- Pygltflib (<https://pypi.org/project/pygltflib/>): Utilized to store outputs in GLTF2 format.
- Numba (<https://numba.pydata.org/>): Provided computation acceleration with jit and prange functions.
- Open3D (<https://www.open3d.org/>): Used for 3D data visualization.
- Noise (<https://pypi.org/project/noise/>) library's pnoise3 function: Generated noise data for fractal colors.

All this has been applied to the fine development of such important complex fractal structures as Mandelbulb, Menger Sponge, Julia, and Sierpinski Gasket, with double precision x, y, z local coordinates, and RGB color values. The meshes that are generated manifest high visual quality and generally accurate details for scientific analyses. All datasets and their source codes can be accessible on this repository: <https://github.com/NeziheSozen/3d-fractal-generators>

The Mandelbulb is a three-dimensional, mathematical object of fractal nature [44] [45] [46]. This paper uses the Daniel White and Paul Nylander approach using spherical coordinates. The following algorithm was used to generate the 3D Mandelbulb datasets:

---

**Algorithm 2** Mandelbulb 3D Fractal Creation

---

```

Input: max_iterations, bailout, power, resolution, x_min, x_max, y_min, y_max, z_min, z_max
Output: Mandelbulb 3D fractal
for x  $\leftarrow$  x_min to x_max step resolution do
  for y  $\leftarrow$  y_min to y_max step resolution do
    for z  $\leftarrow$  z_min to z_max step resolution do
       $z_x \leftarrow x, z_y \leftarrow y, z_z \leftarrow z$ 
      iteration  $\leftarrow$  0
      while iteration  $<$  max_iterations and  $(z_x^2 + z_y^2 + z_z^2) < bailout$  do
         $r \leftarrow \sqrt{z_x^2 + z_y^2 + z_z^2}$ 
         $\theta \leftarrow \arctan 2(\sqrt{z_x^2 + z_y^2}, z_z)$ 
         $\phi \leftarrow \arctan 2(z_y, z_x)$ 
         $new\_r \leftarrow r^{\text{power}}$ 
         $new\_theta \leftarrow \theta \cdot \text{power}$ 
         $new\_phi \leftarrow \phi \cdot \text{power}$ 
         $z_x \leftarrow new\_r \cdot \sin(new\_theta) \cdot \cos(new\_phi) + x$ 
         $z_y \leftarrow new\_r \cdot \sin(new\_theta) \cdot \sin(new\_phi) + y$ 
         $z_z \leftarrow new\_r \cdot \cos(new\_theta) + z$ 
        iteration  $\leftarrow$  iteration + 1
      end while
      if iteration = max_iterations then
        Point (x, y, z) is part of the fractal.
      else
        Point (x, y, z) is not part of the fractal.
      end if
    end for
  end for
end for

```

---

This algorithm creates points in a 3-dimensional space by calculating the Julia set with a given quaternion and parameters [47] [48]:

**Algorithm 3** Quaternion Based Julia Set Algorithm

---

```

function JULIASETQUATERNION(c, max_iter, threshold, resolution)
    for x  $\leftarrow$  -resolution to resolution do
        for y  $\leftarrow$  -resolution to resolution do
            for z  $\leftarrow$  -resolution to resolution do
                q  $\leftarrow$  quaternion(x, y, z, 0)
                n  $\leftarrow$  0
                while n < max_iter and norm(q) < threshold do
                    q  $\leftarrow$  q2 + c
                    n  $\leftarrow$  n + 1
                end while
                if norm(q)  $\geq$  threshold then
                    plot_point(x, y, z, n)
                end if
            end for
        end for
    end for
end function

```

---

The Menger sponge is a three-dimensional fractal geometric shape defined by Karl Menger. A structure such as this one is developed by removing smaller cubes from the center and each face of an initial cube, therefore making a structure of almost no volume and infinite surface area through its infinite iterations. [49] [50]

**Algorithm 4** Menger Sponge Creation

---

Input: *max\_iterations*, *cube\_size*  
 Output: Menger Sponge

```

function MENGERSPONGE(x, y, z, size, iteration)
    if iteration = max_iterations then
        Draw cube at (x, y, z) with size size
    else
        new_size  $\leftarrow$  size/3
        for i, j, k  $\in$  {0, 1, 2} do
            if not (i = 1  $\wedge$  j = 1)  $\vee$  (i = 1  $\wedge$  k = 1)  $\vee$  (j = 1  $\wedge$  k = 1) then
                MENGERSPONGE(x + i · new_size, y + j · new_size, z + k · new_size, new_size, iteration + 1)
            end if
        end for
    end if
end function
MENGERSPONGE(0, 0, 0, cube_size, 0)

```

---

The Sierpinski Gasket also referred to as the Sierpinski Triangle, is named after by the name of Wacław Sierpiński, who described this fractal [51]. The creation of this fractal involves recursively cutting an equilateral triangle into three smaller equilateral triangles and leaving the central triangle of each such division empty. This is done as many times as possible. Thus, it goes on to yield an extremely elaborate and self-replicating pattern [52]. The below algorithm shows the procedures to create The Sierpinski Tetrahedron:

**Algorithm 5** Sierpinski Gasket Tetrahedron

---

```

1: procedure GENERATESIERPINSKITETRAHEDRON
2:   Input: Number of iterations  $n$ 
3:   Output: Sierpinski gasket tetrahedron
4:   Initialize tetrahedron with vertices  $A, B, C, D$ 
5:   for  $i = 1$  to  $n$  do
6:     Divide each face of the tetrahedron into 3 equal parts
7:     Calculate the midpoints of each face:
8:      $M_{AB} = \frac{A+B}{2}$ 
9:      $M_{BC} = \frac{B+C}{2}$ 
10:     $M_{CA} = \frac{C+A}{2}$ 
11:    Create new tetrahedrons using these midpoints
12:    Replace the original tetrahedron with the new tetrahedrons
13:   end for
14: end procedure

```

---

**References**

- [1] Weile Jia, Han Wang, Mohan Chen, Denghui Lu, Lin Lin, Roberto Car, E Weinan, and Linfeng Zhang. Pushing the limit of molecular dynamics with ab initio accuracy to 100 million atoms with machine learning. In *SC20: International conference for high performance computing, networking, storage and analysis*, pages 1–14. IEEE, 2020.
- [2] Qingpu Lou. The lost: An attempt to combine film and game. *Highlights in Science, Engineering and Technology*, 72:446–452, 2023.
- [3] Yue Zhao et al. Deep learning of 3d high-precision model digital engraving of next-generation games based on artificial intelligence. *Advances in Multimedia*, 2022, 2022.
- [4] Janek Kozicki, Anton Gladky, and Klaus Thoeni. Implementation of high-precision computation capabilities into the open-source dynamic simulation framework yade. *Computer Physics Communications*, 270:108167, 2022.
- [5] Bin Cui, Wei Tao, and Hui Zhao. High-precision 3d reconstruction for small-to-medium-sized objects utilizing line-structured light scanning: A review. *Remote Sensing*, 13(21):4457, 2021.
- [6] Patrick C Shriwise, Paul PH Wilson, Andrew Davis, and Paul K Romano. Hardware-accelerated ray tracing of cad-based geometry for monte carlo radiation transport. *Computing in Science & Engineering*, 24(2):52–61, 2022.
- [7] Greg Henry, Ping Tak Peter Tang, and Alexander Heinecke. Leveraging the bfloat16 artificial intelligence datatype for higher-precision computations. In *2019 IEEE 26th Symposium on Computer Arithmetic (ARITH)*, pages 69–76. IEEE, 2019.
- [8] Riccardo Alidori. *Hardware support for a novel variable precision floating point format in a scientific computing environment*. PhD thesis, Politecnico di Torino, 2020.
- [9] Noureddine Ait Said, Mounir Benabdenbi, and Katell Morin-Allory. Fpu reduced variable precision in time: Application to the jacobi iterative method. In *2021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)*, pages 170–175. IEEE, 2021.
- [10] Fabienne Jézéquel, Stef Graillat, Daichi Mukunoki, Toshiyuki Imamura, and Roman Iakymchuk. Can we avoid rounding-error estimation in hpc codes and still get trustworthy results? In *Software Verification: 12th International Conference, VSTTE 2020, and 13th International Workshop, NSV 2020, Los Angeles, CA, USA, July 20–21, 2020, Revised Selected Papers 13*, pages 163–177. Springer, 2020.
- [11] Matteo Croci, Massimiliano Fasi, Nicholas J Higham, Theo Mary, and Mantas Mikaitis. Stochastic rounding: implementation, error analysis and applications. *Royal Society Open Science*, 9(3):211631, 2022.
- [12] OpenGL. Opengl api, 2024. Available at: <https://www.opengl.org>.
- [13] Vulkan. Vulkan api, 2024. Available at: <https://www.vulkan.org>.
- [14] DirectX. Directx api, 2024. Available at: <https://www.microsoft.com/directx>.
- [15] OpenCL. Opencl api, 2024. Available at: <https://www.khronos.org/opencl>.
- [16] CUDA. Cuda api, 2024. Available at: <https://developer.nvidia.com/cuda-zone>.

- [17] OpenGL Wiki. Core language glsl, 2024.
- [18] Guillaume Da Graça and David Defour. Implementation of float-float operators on graphics hardware. *CoRR*, abs/cs/0603115, 2006.
- [19] Andrew Thall. Extended-precision floating-point numbers for gpu computation. In *ACM SIGGRAPH 2006 Research Posters*, SIGGRAPH '06, page 52–es, New York, NY, USA, 2006. Association for Computing Machinery.
- [20] Alexander Bock, Emil Axelsson, Carter Emmart, Masha Kuznetsova, Charles Hansen, and Anders Ynnerman. Openspace: Changing the narrative of public dissemination in astronomical visualization from what to how. *IEEE Computer Graphics and Applications*, 38(3):44–57, 2018.
- [21] William J Dally, Stephen W Keckler, and David B Kirk. Evolution of the graphics processing unit (gpu). *IEEE Micro*, 41(6):42–51, 2021.
- [22] Lorenzo Schwertner Kaufmann, Flavio Paulus Franzin, Roberto Menegais, and Cesar Tadeu Pozzer. Accurate real-time physics simulation for large worlds. In *VISIGRAPP (1: GRAPP)*, pages 135–142, 2021.
- [23] Clay John. Emulating double precision on the gpu to render large worlds, 2022.
- [24] Jason Gregory. *Game engine architecture*. AK Peters/CRC Press, 2018.
- [25] Tomas Akenine-Moller, Eric Haines, and Naty Hoffman. *Real-time rendering*. AK Peters/crc Press, 2019.
- [26] Mike Bailey. The vulkan computer graphics api. In *ACM SIGGRAPH 2023 Courses*, pages 1–158. bul, 2023.
- [27] Graham Sellers and John Kessenich. *Vulkan programming guide: The official guide to learning vulkan*. Addison-Wesley Professional, 2016.
- [28] Ed Angel and Dave Shreiner. Modern opengl programming. In *SIGGRAPH Asia 2011 Courses*, SA '11, New York, NY, USA, 2011. Association for Computing Machinery.
- [29] Mike Bailey and Steve Cunningham. *Graphics shaders: theory and practice*. AK Peters/CRC Press, 2009.
- [30] IEEE-754. Ieee approved draft standard for floating-point arithmetic. *IEEE P754/D2.50, April 2019*, pages 1–83, 2019.
- [31] Robert Johns. Float vs double data types: What is the difference [updated], 2023.
- [32] Floating point visually explained, 2017.
- [33] Double precision coordinates, April 2024.
- [34] Benoit B Mandelbrot, Carl JG Evertsz, and Martin C Gutzwiller. *Fractals and chaos: the Mandelbrot set and beyond*, volume 3. Springer, 2004.
- [35] Sylvie Boldo and Guillaume Melquiond. When double rounding is odd. *bul*, 07 2005.
- [36] Vincent Lafage. Revisiting "what every computer scientist should know about floating-point arithmetic", 2020.
- [37] Dina Tsarapkina and David J Jeffrey. Exploring rounding errors in matlab using extended precision. *Procedia Computer Science*, 29:1423–1432, 2014.
- [38] David A. Patterson and John L. Hennessy. *Computer Organization and Design, Revised Fourth Edition, Fourth Edition: The Hardware/Software Interface*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 4th edition, 2011.
- [39] Simon Klarmann and Jens Wackerfuß. Possibilities and drawbacks using arbitrary precision numbers for structural analysis. *PAMM*, 20(1):e202000079, 2021.
- [40] Spir-v toolchain, July 2024.
- [41] What is spir-v, July 2024.
- [42] David H. Bailey. Dsfun90 (fortran-90 double-single package), 2005.
- [43] Baldur Karlsson. Renderdoc. URL <https://renderdoc.org>, 2019.
- [44] Daniel White. The unravelling of the real 3d mandelbulb. *On line*, 2009.
- [45] Javier Barrallo. Expanding the mandelbrot set into higher dimensions. In *Proceedings of Bridges 2010: Mathematics, Music, Art, Architecture, Culture*, pages 247–254, 2010.
- [46] Oliver Knill. Mandelbulb, mandelbrot, mandelring and hopfbrot. *arXiv preprint arXiv:2305.17848*, 2023.
- [47] Alan Norton. Julia sets in the quaternions. *Computers & graphics*, 13(2):267–278, 1989.
- [48] John C Hart, Louis H Kauffman, and Daniel J Sandin. Interactive visualization of quaternion julia sets. In *Proceedings of the First IEEE Conference on Visualization: Visualization90*, pages 209–218. IEEE, 1990.

- [49] Gerald A Edgar. *Classics on fractals*. CRC Press, 2019.
- [50] Heinz-Otto Peitgen, Hartmut Jürgens, Dietmar Saupe, and Mitchell J Feigenbaum. *Chaos and fractals: new frontiers of science*, volume 106. Springer, 2004.
- [51] Waclaw Sierpinski. *General topology*. Courier Dover Publications, 2020.
- [52] Robert Whitrow. *Curves, Surfaces and Patterns*, pages 249–311. Springer London, London, 2008.