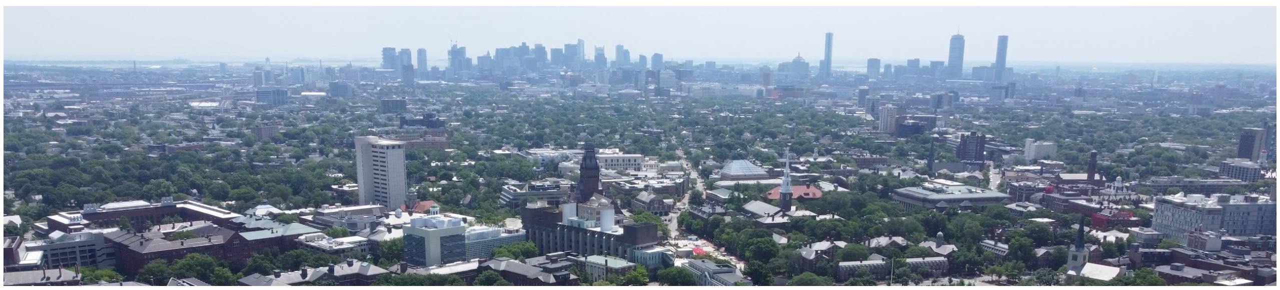
MULTIVARIABLE CALCULUS

MATH S-21A

Detective work with space data



1.1. In this first data project, we want to get to know each other and learn a bit about Harvard. This data project deals with spacial data in the image footage (https://www.youtube.com/watch?v=b5kkB7urSko June 20, 2021).
There are 25 small problems. Each can be answered in a single sentence. Sometimes, even a single word is enough. Split up the work up in your project group, then come together as a team and present the answers with 5 slides, each slide covering one page. The submit the PDF of the slides. Keep things tight and concise. For example, if the question would have been to find clues for the time of the year when the footage was taken, one could state simply "Trees and vegetation. Boats in the water, clothing of people, sun position".

Problem 1: a) What general clues are there to determine the wind direction during the time the video was taken?
b) What general clues are there to determine at which time of the day the pictures was taken?
c) Try to figure out the time when looking at 6:58.
d) There are other clue at clocks. Are they consistent?
e) Do you have an explanation for any discrepancies?

Problem 2: a) The movie has been time modified at time 7:36. We tell how much?
b) At what position and height was the camera at 3:00?
c) Can you use landmarks to determine the position and height of the camera at 5:01?
d) Can you figure out the name and height of the tall white building visible at 4:46?
e) Please reconstruct the picture at 5:01 in the application google earth and make a screen shot.

Problem 3: a) What is the name of the river house with the green dome at 3:50?
b) What is the name of the river house with the red dome at 1:46?
c) What is the name of the river house at 2:00.
d) What is the name of the river house with the blue dome at 4:10?
e) What is the name of the place square shape grass surrounded by a student dorms at 5:43?

Problem 4: a) The Boston skyline shows three large buildings at 2:00. What are the names of these buildings?
b) What is the large building near the river at 1:20. It is also visible at 2:10.
c) Why are there no cars on Memorial drive 1:55?
d) What is inside the gray half spheres on 8:25?
e) On 2:11, we see two bridges. Connect their centers and intersect with the horizon. Which town is this? The name starts with W.

Problem 5: a) At 5:45, there is some lake visible near the horizon. It appears in the picture as a triangular shape. What is the name of these lake. The lake is also visible at 5:36 but it is not the closer Fresh pond to the left.
b) At 4:52, you see at the horizon some chimneys, a bridge and an obelisk type shape which is not a chimney. Identify these points.
c) Oliver lives in the vicinity of the water tower visible at 5:43 in the middle of the picture at the horizon. What is the name of the town?
d) Around 6:05, one can see something from the Boston airport. Identify it and name it.
e) From 1:00 to 1:05, the camera flies towards large new building complex on the other side of the river. Identify it.

Data illustration 2: Acceleration data

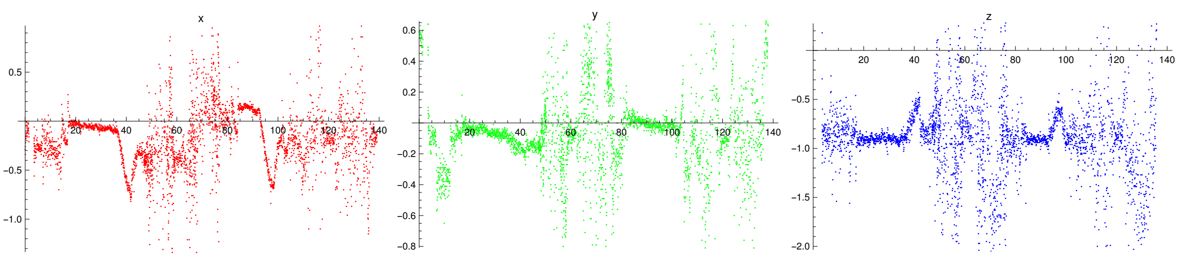
2.1. There are apps on your cellphone that can measure your acceleration data. Using these data, we can construct the curve $\vec{r}(t) = [x(t), y(t), z(t)]$ which describes the path which we have taken with the phone. What the data provide are lists of numbers which encode the **acceleration data** $x''(t), y''(t), z''(t)$. Summing up these data and fixing the initial velocities give lists of numbers which encode the **velocity data** $x'(t), y'(t), z'(t)$. Now one can do it again and get the **position data**.

2.2. In the blog

$$http://blog.robindeits.com/2013/11/11/roller-coaster-visualizations\ ,$$

Robin Deits has used a smartphone to measure the accelerations during a **roller coaster ride** on **Cedar point**, the foremost American roller coaster park.
Robin made the data available on the website *https://github.com/rdeits/coasters*.

2.3. Here are the **x-acceleration**, the **y-acceleration** and the **z-acceleration** data, plotted in a particular case, the Mine ride roller coaster:



2.4. One can get back the actual shape of the roller coaster from these data:

Image source: https://coasterforce.com/mine-train

2.5. Here are the 20 example data points. The first column is time, the second column is the x-acceleration, the third the y-acceleration and the fourth the z-acceleration.

```
A={
{ 0.002, −0.29, 0.59, 0.90},
{ 0.078, −0.26, 0.55, 0.85},
{ 0.080, −0.34, 0.67, 0.88},
{ 0.147, −0.22, 0.59, 0.88},
{ 0.192, −0.16, 0.55, 0.80},
{ 0.255, −0.15, 0.57, 0.85},
{ 0.257, −0.18, 0.57, 0.84},
{ 0.347, −0.18, 0.59, 0.72},
{ 0.383, −0.05, 0.58, 0.91},
{ 0.446, −0.13, 0.72, 0.81},
{ 0.485, −0.12, 0.63, 0.81},
{ 0.510, −0.01, 0.58, 0.93},
{ 0.581, −0.04, 0.53, 0.86},
{ 0.637, −0.01, 0.38, 0.93},
{ 0.684, −0.07, 0.54, 0.86},
{ 0.700, −0.15, 0.50, 0.79},
{ 0.777, −0.16, 0.57, 0.98},
{ 0.827, −0.12, 0.55, 0.98},
{ 0.902, −0.11, 0.56, 0.97},
{ 0.904, −0.05, 0.57, 0.97}}
};
```

Problem 1: How would you use this list of data to get a list of data representing the velocities assuming that the velocities at $t = 0$ are zero.

Problem 2: How would you use the velocity data to get the position data assuming that at $t = 0$ we have the position $(0, 0, 0)$.

Problem 3: On the data page, you have access to a Mathematica program and data files from that curve. Describe what the programs are doing.

Problem 4: Plot both graphs, the original curve and the reconstructed curve.

Problem 5: Explain why the two curves are translated to each other.

Acknowledment: Original idea for the project Summer 2019: Jun-Hou Fung.

Oliver Knill, knill@math.harvard.edu, Math S-21a, Harvard Summer School, 2021

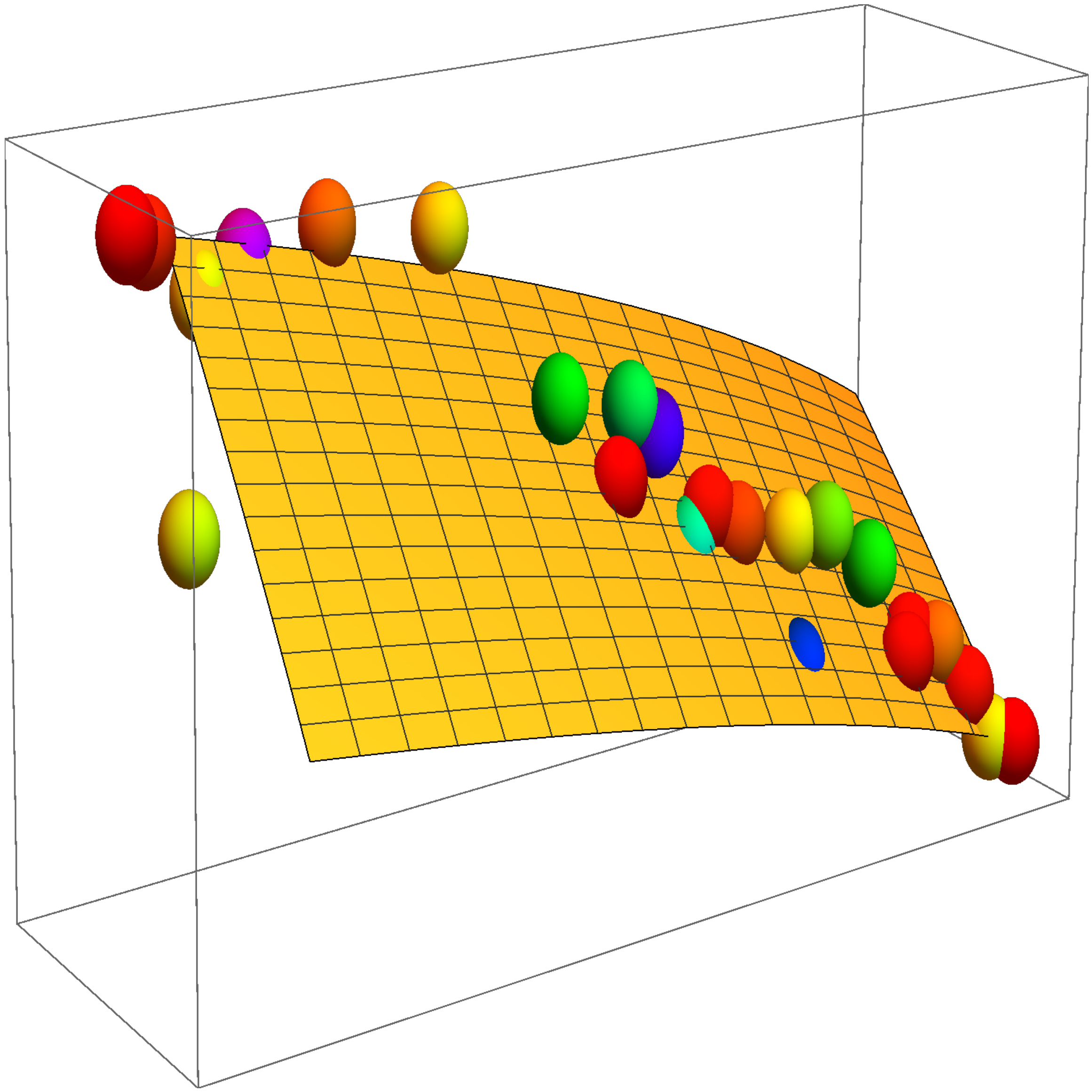
MULTIVARIABLE CALCULUS

MATH S-21A

Data illustration 3: Cobb Douglas

2.1. The mathematician and economist **Charles W. Cobb** at Amherst college and the economist and politician **Paul H. Douglas** who was also teaching at Amherst, found in 1928 empirically a formula $F(K,L) = L^{\alpha}K^{\beta}$ which fits the **total production** F of an economic system as a function of the **capital investment** K and the **labor** L. The two authors used logarithms variables and assumed linearity to find α, β. Below are the data normalized so that the date for year 1899 has the value 100.

$Year$	K	L	P
1899	100	100	100
1900	107	105	101
1901	114	110	112
1902	122	118	122
1903	131	123	124
1904	138	116	122
1905	149	125	143
1906	163	133	152
1907	176	138	151
1908	185	121	126
1909	198	140	155
1910	208	144	159
1911	216	145	153
1912	226	152	177
1913	236	154	184
1914	244	149	169
1915	266	154	189
1916	298	182	225
1917	335	196	227
1918	366	200	223
1919	387	193	218
1920	407	193	231
1921	417	147	179
1922	431	161	240



The graph of $F(L, K) = L^{3/4} K^{1/4}$ fits pretty well that data set.

Problem 1: Before we start with calculus, can you see in the data, which data point is an out-layer? You might see it by looking at differences of values from one year to the next, which is a time derivative derivative.

Problem 2: Assume that the labor and capital investment are bound by the additional constraint $G(L, K) = L^{3/4} + K^{1/4} = 50$. (This function G is unrelated to the function $F(L, K)$ as we are in a Lagrange problem.) Where is the production P maximal under this constraint? Plot the two functions $F(L, K)$ and $G(L, K)$.

MULTIVARIABLE CALCULUS

MATH S-21A

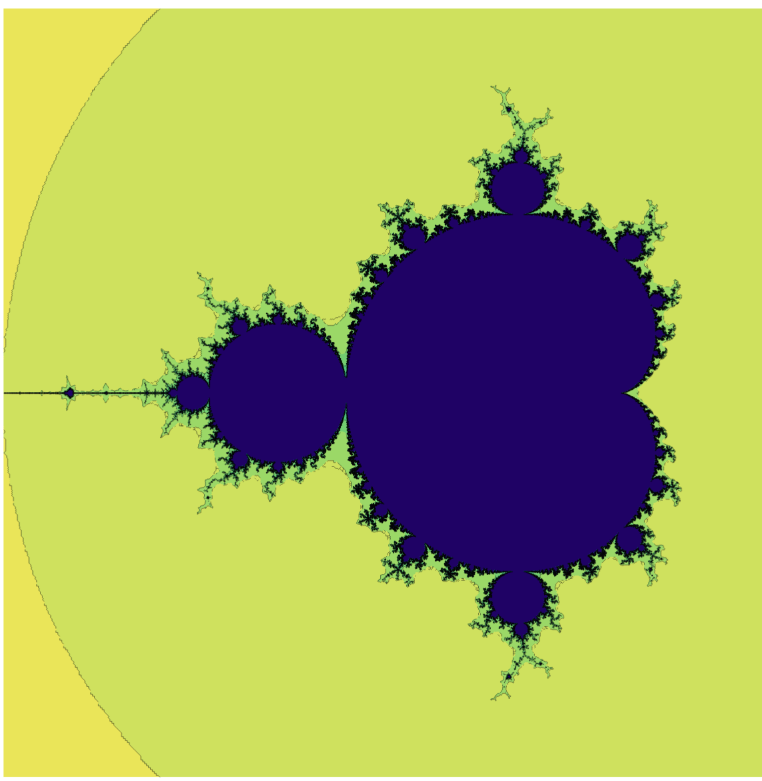
Data project Week 4: Monte Carlo computation

2.1. Often, when we deal with real data, we do not have analytic expressions for the region or function we want to integrate. We want to elaborate here on an example mentioned in the text for unit 15. It is the problem to find the area of Mandelbrot set

$$M = \{c = a + ib \in \mathbb{C} \in \mathbb{R}^2 \mid T_c(0)^n \text{ stays bounded } \} ,$$

where $T_c(z) = z^+c$ (as complex numbers, which is written out in real coordinates the map $T_c(x, y) = (x^2 - y^2 + a, 2xy + b)$.

2.2. Here is a picture: it can also be visualized as a function which is 1 on the Mandelbrot set and 0 else.



2.3. What is the area of the Mandelbrot set? We know it is contained in the rectangle $x \in [-2, 1]$ and $y \in [-3/2, 3/2]$. We now just randomly shoot into this rectangle and see whether we are in the Mandelbrot set or not after 1000 iterations. Here is some Mathematica code which allows you to compute things. When we ran it, it gave a value of about 1.515.... More accurate measurements reported hint for a slightly smaller value like 1.506.... Others have given bounds [1.50311, 1.5613027].

```
M=Compile[{x,y},Module[{z=x+I y,k=0},
     While[Abs[z]<2.&&k<1000,z=N[z^2+x+I y];++k];Floor[k/1000]]];
9*Sum[M[-2+3 Random[],-1.5+3 Random[]],{1000000}]/1000000
```

2.4. Here are the problems for Friday the 16th in your group. Please make again slides. I know many of you struggle with time. Split the work.

Problem 1: Sketch in words in one slide what the Monte Carlo method is. Who is credited for its invention. Why Monte Carlo? What are applications?

Problem 2: What is the Mandelbrot set? What is the significance in mathematics, in pop culture?

Problem 3: How accurately can you compute the area of the Mandelbrot set with the above code? Which parameters do you need to tune to make it more precise? Experiment.

Problem 4: What is the Mandelbulb set? Try to find out about the history.

Problem 5: How would you go about setting up a Monte Carlo computation for the volume of the Mandelbulb? This is a bit of a research task as we have not seen a serious estimate for the volume of the Mandelbulb.

(*) Some reference from 2013 *http://people.math.harvard.edu/~knill/slides/boston/sic.pdf*

MULTIVARIABLE CALCULUS

MATH S-21A

Data Visualization Week 5: Relief data

The following table shows elevation data near the town of Waltham in Massachusetts. Each grid is 4000 feet apart and elevations are in feet. Let $f(x, y)$ be the elevation data assigned to position cell (x, y). The bottom left cell corresponds to the position $(1, 1)$, the top right cell to position $(9, 7)$.

269	246	213	253	240	253	233	236	253
226	210	203	213	220	246	256	213	240
269	203	187	233	236	230	259	203	210
233	180	223	269	112	180	194	157	128
138	164	203	161	69	59	59	75	105
177	112	174	98	69	69	46	46	49
233	171	85	52	72	82	59	56	56

Problem 1: Sketch the contour $f(x, y) = 200$ by drawing out any division line between two cells where the value changes from below to above 200. You can do this on a printout of this worksheet.

Problem 2: If you move from cell $(1, 1)$ to cell $(7, 7)$ on a straight line, you see the elevation change in time. Draw a relief graph of that height.

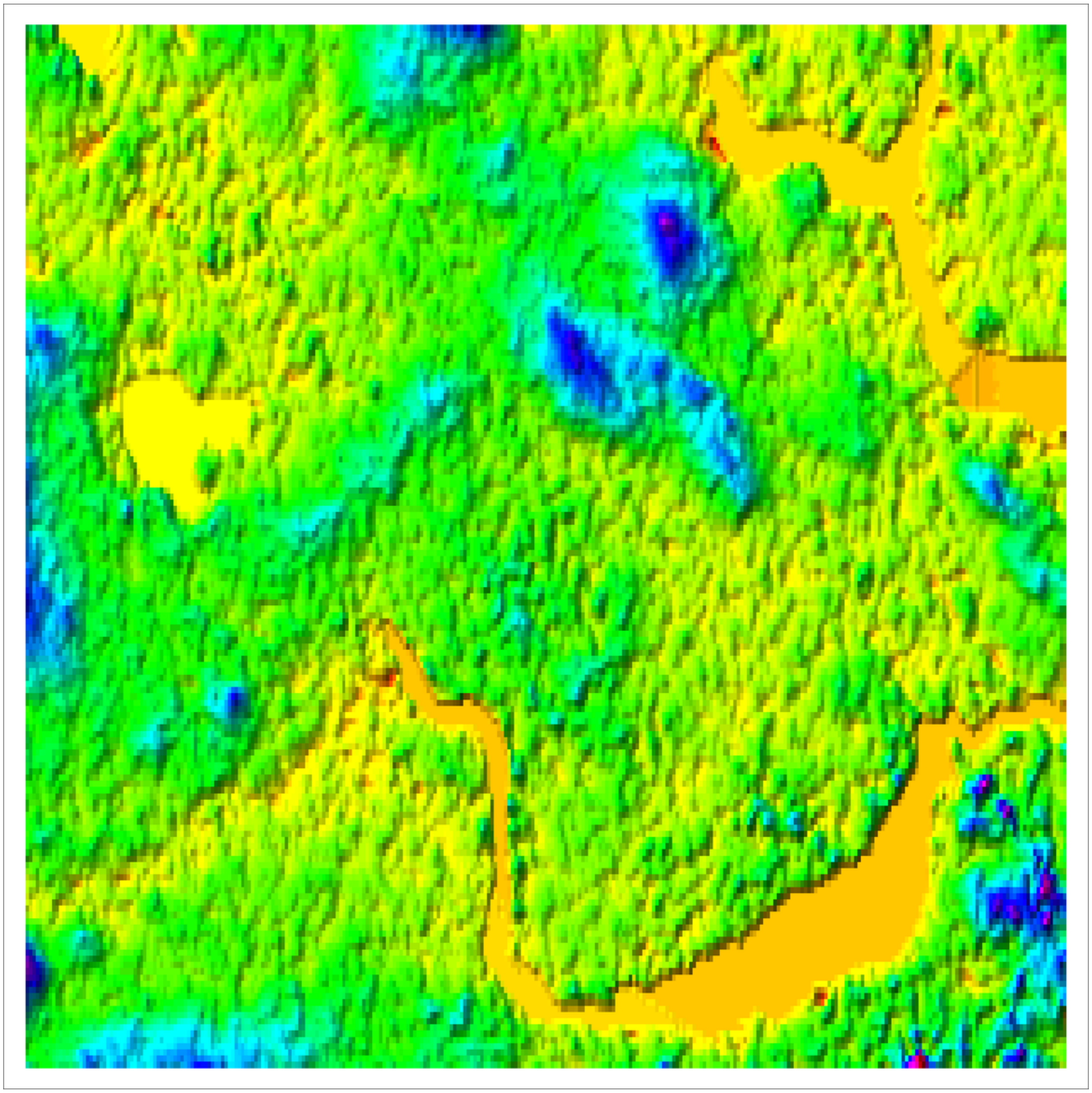
Problem 3: Between which points is the slope of the landscape described in the table maximal? How steep do you estimate it to be there? In what direction would one need to walk to go straight uphill? (you can restrict to up/down/left/right).

Problem 4: Mathematica which provided the table does not give details on how these numbers were obtained. Speculate about the choices which were made when collecting these elevation measurements? There is no right, nor wrong answer, but we want you to think of at least two different possibilities which were used to write down the data points.

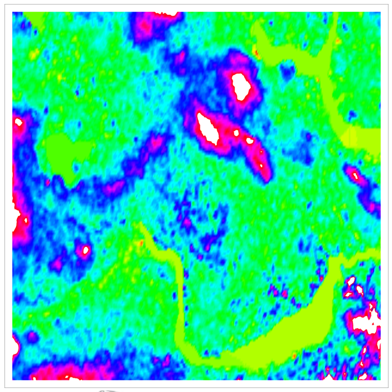
Problem 5: The Mathematica below gives example code to visualize elevation data. Most towns in the US are covered. Chose a town of your liking different from Cambridge visualize its geographic height data. You can get a Mathematica notebook from the website, run the modified notebook and print it out.

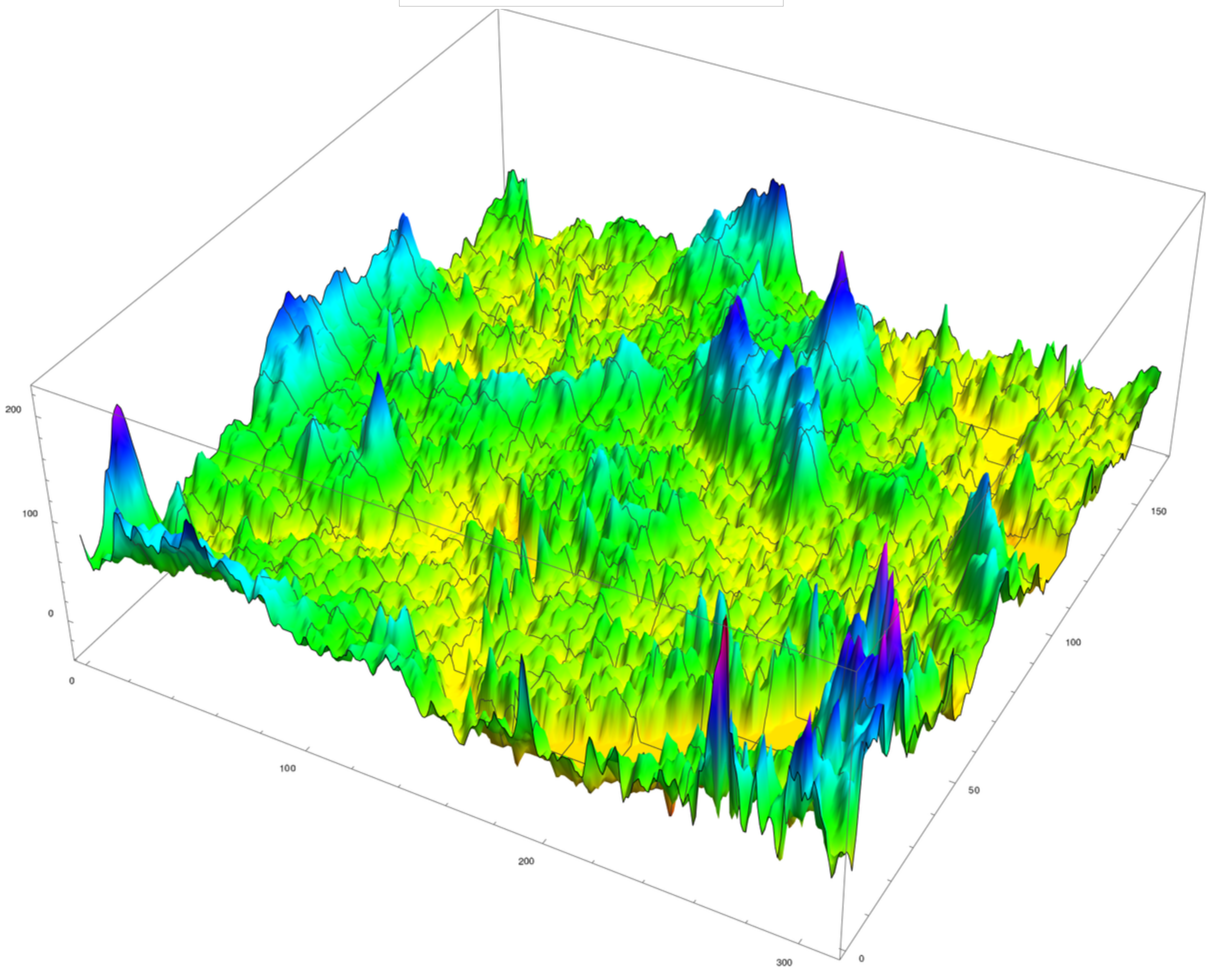
5.1. Mathematica example code with real Geo data:

```
A = Reverse[Normal[ GeoElevationData[
    Entity["City", {"Cambridge", "Massachusetts", "UnitedStates"}],
    GeoProjection -> Automatic]]];
ReliefPlot[A]
ListDensityPlot[A]
ListPlot3D[A]
ListContourPlot[A]
```



Multivariable Calculus





Oliver Knill, knill@math.harvard.edu, Math S-21a, Harvard Summer School, 2021

INTRODUCTION TO CALCULUS

MATH 1A

Data project 1: Primes

1.1. Define the **prime function** $f(x)$ as the function on positive integers which gives the x'th prime. So,

$$f(1) = 2, f(2) = 3, f(3) = 5, f(4) = 7 .$$

By definition, the prime function is **monotone** in the sense that $f(x+1) > f(x)$ and more generally $f(y) > f(x)$ if $y > x$. To start this project, make a list of the first 17 primes.

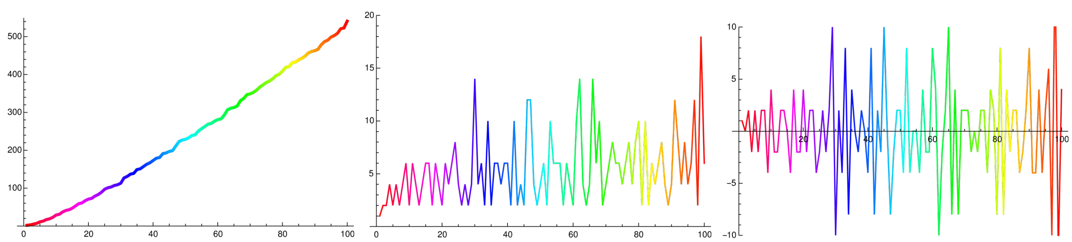


FIGURE 1. Prime function $f(x)$, velocity $f'(x) = f(x+1) - f(x)$ and acceleration $f''(x) = f(x+2) - 2f(x+1) - f(x)$.

1.2. Define the **derivative**

$$f'(x) = f(x+1) - f(x) .$$

This measures the increase or decrease between x and $x+1$ and is also known as **rate of change**. For any positive integer h, the expression

$$\frac{f(x+h) - f(x)}{h}$$

is called the **average rate of change** from x to $x+h$. The name average is justified because

$$\frac{f(x+h) - f(x)}{h} = \frac{f'(x) + f'(x+1) + \cdots + f'(x+h-1)}{h}$$

is the average between all the rate of changes. Question 2: write down the average rate of change equation for $x = 1$ and $h = 9$.

1.3. A point x for which $f'(x) = 2$ is called a **prime twin**. A big conjecture is that there are infinitely many prime twins. The problem appears too difficult for current mathematics to be solved; but we can find small prime twins, especially with a computer. Third question: find the first 10 prime twins

1.4. The second derivative of f is called the **acceleration**. It is defined as the derivative of the derivative function f'. We have

$$f''(x) = f'(x+1)-f'(x) = (f(x+2)-f(x))-(f(x+1)-f(x) = f(x+2)-2f(x+1)+f(x) .$$

We have $f''(2) = 0$ because $f(4) - 2f(3) + f(2) = 7 - 2 * 5 + 3 = 0$. Fourth problem: find the next two x for which $f''(x) = 0$.

1.5. The **fundamental theorem of calculus for data** tells that if $f'(x) = f(x + 1) - f(x)$, then

$$f'(a) + f'(a + 1) + f'(a + 2) + \cdots + f'(b - 1) = f(b) - f(a)$$

for any integers a, b. Verify this for the prime function $f(x)$ with $a = 1, b = 10$. The fundamental theorem will later in the course be written as

$$\int_a^b f'(x)dx = f(b) - f(a) .$$

1.6. Primes are one of the first things which have been considered in mathematics. Euclid was the first to prove that there are infinitely many primes. His argument was ingenious: assume there would be only finitely many prime data $p_1, \ldots, p_n$, then the number $m = p_1 \cdot p_2 \cdots p_n + 1$ has either to be prime or to contain a new prime factor different from the given list; indeed, m is not divisible by any of the primes. This is result is important as it justifies that the **prime function** f we have been talking about, really exists. In the last part, we want you to look up **Sophie Germain primes**. Fifth problem: Look up the definition Germain primes. Why was Sophie interested in these primes?

1.7. Finally, in problem 6, we want you to look up the largest known prime. There is a large collaborative project going on in which humanity tries to push the largest known prime further. What is the largest known prime today? What is the name of this prime hunting project?

FIGURE 2. An artistic rendering of Sophie Germain.

Data project 2: Polyhedra

2.1. In the context of **data science**, **graphs** are very important. **Graph databases** started to appear already with the **Incas** in the form of **Khipu** documents. Nature has used DNA as a database since millions of years. In modern information technology, graph databases replace more and more traditional relational databases. A **graph** is a collection of nodes called **vertices** joined by connections called **edges**. Street maps, electric grids, directories, family or collaboration trees, brains, social networks or polyhedra are examples of graphs.

2.2. A graph G comes with a **graph polynomial**. This brings us to calculus. If the graph has a_0 vertices and a_1 edges and a_2 triangles the polynomial is defined as $f_G(x) = 1 + a_0 x + a_1 x^2 + a_2 x^3$. More generally, if a_k is the number of k**-cliques**, groups of k vertices all connected to each other, then $f_G(x) = 1 + \sum_{k=0} a_k x^{k+1}$, where Σ is the summation sign.

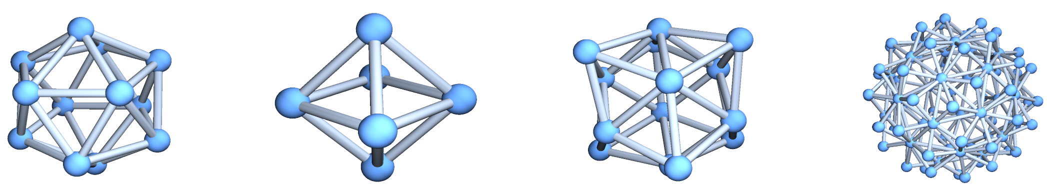


FIGURE 1. The Icosahedron, Octahedron, Tetraxishexahedron and Echidnahedron.

2.3. The **unit sphere** $S(v)$ of a vertex v consists of the subgraph build by all vertices connected to the node. If G is a friendship graph, and v is you, then $S(v)$ is the friend network of your friends without you. Let $f_{S(v)}(x)$ denote the polynomial function to $S(v)$.

Theorem: $f'_G(x) = \sum_{v \in V} f_{S(v)}(x)$

It is a **Gauss-Bonnet theorem** [1] because it allows to compute the Euler characteristic $\chi(G) = a_0 - a_1 + a_2 - \cdots$ of a graph by integrating and evaluating at -1: Indeed, we have from the **fundamental theorem of calculus** $\chi(G) = f_G(0) - f_G(-1) =$

[1] O. Knill, Dehn-Sommerville from Gauss-Bonnet, https://arxiv.org/abs/1905.04831, 2019

$\int_{-1}^{0} f'_G(x)\ dx$. Because the formula computes $f_G(x)$ recursively through unit spheres, allows to compute clique numbers of most graphs quickly. In general, it is a notoriously difficult NP complete problem.

2.4. Example: the icosahedron G has 12 vertices, 30 edges and 20 triangles. The polynomial is $f_G(x) = 1 + 12x + 30x^2 + 20x^3$. And the Euler characteristic is $\chi(G) = f(0) - f(-1) = 12 - 30 + 20 = 2$. René Descartes already noticed that Platonic solids lead to $\chi(G) = 2$ and wrote it down in an encrypted way in a secret notebook since mathematical theorems were at that time considered treasures. The unit sphere of a vertex is a circular graph with 5 vertices and 5 edges has the polynomial $f_{S(v)}(x) = 1 + 5x + 5x^2$. Summing over all 12 vertices gives $12 + 60x + 60x^2$. This agrees with $f'_G(x) = 12 + 60x + 60x^2$. Integrate this from -1 to 0 gives the already computed $\chi(G) = 2$.

2.5. Example: for the Echidnahedron with 92 vertices, 270 edges, 240 triangles and 60 tetrahedra, the polynomial is $f_G(x) = 1 + 92x + 270x^2 + 240x^3 + 60x^4$. The unit spheres are 60 triangles with $f_{S(v)}(x) = 1 + 3x + 3x^2 + x^3$ then 12 unit spheres with $f_{S(v)}(x) = 1+10x+15x^2+5x^3$ and 20 unit spheres with $f_{S(v)}(x) = 1+12x+18x^2+6x^3$. They all sum up to $92 + 540x + 720x^2 + 240x^3$ which agrees with $f'_G(x)$.

2.6. Problem 1: Repeat the computation for the octahedron G. So, compute $f_G(x)$, find the unit spheres $S(v)$ and compute $f_{S(v)}(x)$ the verify the formula from the theorem and compute the Euler characteristic.

2.7. Problem 2: Do the same for the **tetraxis hexahedron**: find $f_G(x)$ and $\chi(G)$. It is a cube on which we add a central node on each face and connect it to the 4 vertices of the square. So there are $8 + 6 = 14$ vertices and $6 * 4$ triangles.

2.8. The graphs of the polynomials are very interesting. One can study their roots, their maxima minima etc. When plotting the functions $f_G(x)$ for polynomials, we often notice some symmetry. This is called a **Dehn-Sommerville** property. Let us call a graph a **polyhedron** if all unit spheres $S(v)$ are circular graphs of 4 or more vertices.

2.9. Problem 3: Plot the graphs of $f_G(x)$ for the polyhedra: icosahedron, octahedron and tetraxis hexahedron.

Corollary: If G is a polyhedron, then $f_G(x)$ has a root at $x = -1/2$.

This follows from the theorem. All $g(x) = f_{S(v)}(x)$ have the property that $g(x + 1/2)$ even. Therefore the sum of such functions is even and so $f'_G(x+1/2)$ is even. But that means that $f_G(x + 1/2)$ is odd, so that $x = -1/2$ is a root.

2.10. If G is a d-dimensional Dehn-Sommerville space define the function $h_G(x) = (x - 1)^{d+1} f_G(1/(x - 1))$. In our case, if triangles are the largest cliques, then $d = 2$ and $h_G(x) = (x - 1)^3 f_G(1/(x - 1))$. The **Dehn-Sommerville relations** assert that the h-vector is **palindromic**, meaning that $h_i = h_{d+1-i}$ for all $i = 0, \ldots, d + 1$. These identities is what the Dehn and Sommerville have noticed first. If G satisfies Dehn-Sommerville, then $f(-1) = (-1)^d f(0) = (-1)^d$ and so $\chi(G) = 1 - f(-1) = 1 + (-1)^d$. In the case $d = 2$, we get $\chi(G) = 2$. For the icosahedron, $f_G(x) = 1+12x+30x^2+20x^3$, we have $h_G(x) = 1 + 9x + 9x^2 + x^3$ and indeed $(1, 9, 9, 1)$ is a polindrome. **Problem 4:** Compute $h_G(x)$ for the octahedron.

2.11. The theorem we have seen is very general and holds for any network. **Problem 5:** Draw your own graph (chose something you can actually do), then build all functions $f_{S(v)}(x)$ and compare their sum with the derivative $f'_G(x)$.

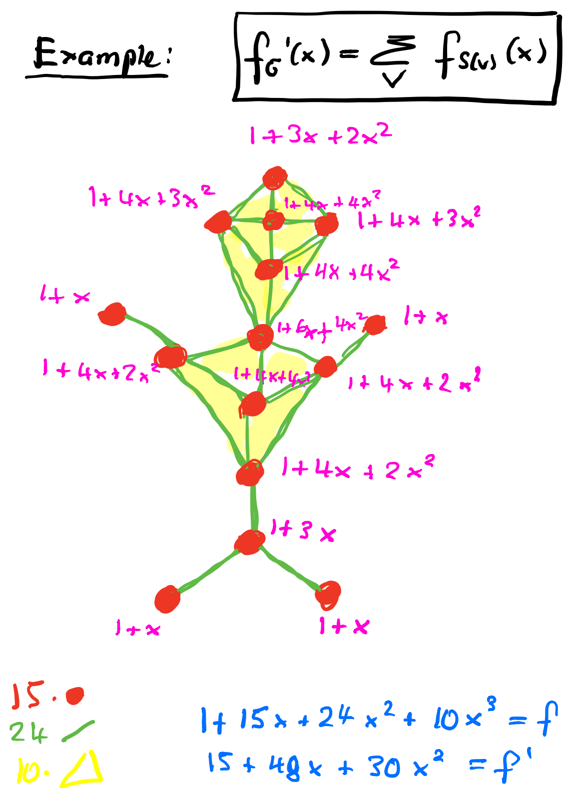


FIGURE 2. A hand-drawn graph. We compute at every vertex v the function $f_{S(v)}(x)$ and them up to get the derivative of $f_G(x)$. Pretty cool. The coefficients of the linear part produce the **Euler Handshake theorem** telling that 2 times the number of edges is the sum over all vertex degrees.

OLIVER KNILL, KNILL@MATH.HARVARD.EDU, MATH 1A, HARVARD COLLEGE, SPRING 2021

INTRODUCTION TO CALCULUS

MATH 1A

Data project 3: Chaos

3.1. Given a function, like $f(x) = 4x(1-x)$, we can look for points where $f(x) = x$. This is called a **fixed point**. You can see that in our case, $x = 0$ and $x = 3/4$ are the two solutions. The function $f(f(x))$ is a new function and in our example again a polynomial. Now, if $f(f(x)) = x$, then we have a periodic point of period 2 because if we apply f to such a point, we return after 2 steps, $x, f(x), f(f(x)) = x$.

3.2. Problem 1: when expanding $f(f(x))$ we get a polynomial of degree 4. $f(f(x)) = 4(4x(1-x))(1-(4x(1-x))) = 16x - 80x^2 + 128x^3 - 64x^4$. Find all the roots of this function. You can use that you already know two of them since if $f(x) = x$, then also $f(f(x)) = x$.

3.3. Let us start with $x = 0.4$ and apply $f(x) = 4x(1-x)$. We get 0.96. Now compute $f(0.96) = 1.536$. We do that 60 times. Here is the Mathematica code which does that.

```
f[x_] := 4x(1-x); NestList[f, 0.4, 60]
```

Running this code produces the following data 0.4, 0.96, 0.1536, 0.520028, 0.998395, 0.00640774, 0.0254667, 0.0992726, 0.35767, 0.918969, 0.29786, 0.836557, 0.546917, 0 0.991195, 0.034909, 0.134761, 0.466403, 0.995485, 0.0179785, 0.0706211, 0.262535, 0.774441, 0.698727, 0.84203, 0.532063, 0.995888, 0.0163812, 0.0644512, 0.241189, 0.732068, 0.784578, 0.676061, 0.87601, 0.434465, 0.982821, 0.067536, 0.251899, 0.753784, 0.742374, 0.76502, 0.719059, 0.808053, 0.620412, 0.942003, 0.218532, 0.683103, 0.865893, 0.464488, 0.994956, 0.0200758, 0.0786912, 0.289996, 0.823593, 0.581151, 0.973658, 0.102592, 0.368269, 0.930587, 0.258378, 0.766475, 0.715963.

3.4. We start again with $x = 0.4$ again but apply the function $g(x) = 4x - 4x^2$. We get 0.96. Now compute $f(0.96) = 1.536$ etc. We do that 60 times. Here is the Mathematica code which does that.

```
g[x_] := 4x-4x^2; NestList[g, 0.4, 60]
```

Running this line of code produces the following data 0.4, 0.96, 0.1536, 0.520028, 0.998395, 0.00640774, 0.0254667, 0.0992726, 0.35767, 0.918969, 0.29786, 0.836557, 0.546917, 0.991195, 0.034909, 0.134761, 0.466403, 0.995485, 0.0179785, 0.0706211, 0.262535, 0.774441, 0.698727, 0.84203, 0.532063, 0.995888, 0.0163812, 0.0644512, 0.241189, 0.732068, 0.784578, 0.676061, 0.87601, 0.434465, 0.982821, 0.0675375, 0.251905, 0.753795, 0.742353, 0.76506, 0.71 8973, 0.808203, 0.620042, 0.942359, 0.217273, 0.680263, 0.870022, 0.452336, 0.990913, 0.0360196, 0.138889, 0.478395, 0.998133, 0.00745456, 0.029596, 0.11488, 0.406731, 0.965204, 0.134342, 0.465178, 0.99515.

3.5. Problem 2) a) What is algebraically the difference between $f(x)$ and $g(x)$? Are you surprised when you look at the data? We have prepared a Desmos applet

https://www.desmos.com/calculator/xvh5ae0a1a

For you to explore that. b) What does Desmos give for $f^{(60)}(0.4)$? c) What result does Desmos give for $g^{(60)}(0.4)$? Note that we write $f^{(2)}(x) = f(f(x))$ and $f^{(3)}(x) = f(f(f(x)))$ etc.

3.6. As the title of this project shows, the reason for this strange behavior is "chaos". Chaos is defined as "sensitive dependence on initial conditions".

3.7. We now take a graphing calculator or the scientific calculator on your phone to experiment with iterating functions.

Problem 3) a) What happens if you push the cos button repetitively. b) What happens if you push alternating the tan and $1/x$ button? (This actually means that we are iterating the cot function but there is no cot button in general on calculators).
Problem 4) a) Who was Mitchell Feigenbaum? b) What did he discover? c) Where is his discovery applied?

Problem 5) a) What is the **Feigenbaum bifurcation diagram** for the logistic map $f_c(x) = cx(1-x)$ describe? b) What is the **Feigenbaum constant?**

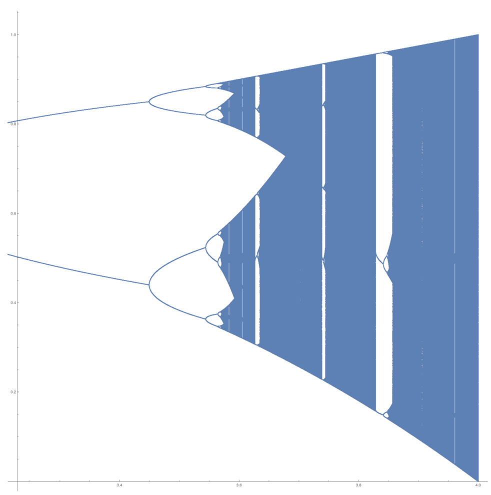


FIGURE 1. Feigenbaum bifurcation diagram

OLIVER KNILL, KNILL@MATH.HARVARD.EDU, MATH 1A, HARVARD COLLEGE, SPRING 2021

Data project 4: Monte Carlo

4.1. Instead of the Riemann integral $S_n = \sum_{a \leq \frac{k}{n} < b} f(\frac{k}{n})\frac{1}{n}$ approximating the integral $\int_a^b f(x)\, dx$, we can also chose random numbers x_k and produce $S_n = \sum_{k=1}^n f(x_k)\frac{b-a}{n}$. Computing the integral as such is called a **Monte Carlo** integration. It turns out to be equivalent to the **Lebesgue integral**, a more sophisticated integral which is required in real analysis or probability theory.

4.2. How do we generate random numbers? There are two fundamental types. The first uses a **dynamical system** to generate pseudo random numbers. An example is to start with a **seed** $x_0 = 0.4$ (which could also depend on the time or other metrics in the computer) and then compute $g^j(x)$ where $g(x)$ is a chaotic map like $g(x) = 4x(1-x)$. The random number generator can then be seeded like $k = 234$. The random numbers x_j are then $x_j = g^{k+j}$ where j goes from 1 to n. This is not bad as we have seen in the last project that it even matters whether we take $g(x) = 4x(1-x)$ or $g(x) = 4x - 4x^2$. Here is code:

```
x0=0.4; k=234; n=10; g[x_]:=4x(1-x);
x1=Last[NestList[g,x0,k-1]]; NestList[g,x1,n-1]
```

4.3. Pseudo random numbers could also be generated (a bit more costly of course) from the digits of pi.

```
n=1000; X[k_]:=Mod[N[Pi*10^k,2n],1];
f[x_]:=Sin[x]; a=0; b=Pi; Sum[f[Pi*X[k]],{k,1,n}]*(b-a)/n
```

Problem 1. Use at least 20 digits of π to find a numerical approximation of the integral $\int_0^1 x^2\, dx$. Make groups of 2 digits to form numbers. Start with the digits after the decimal point. The first number $x_1 = 14/100$, the second $x_2 = 15/100$. Now write down the Riemann sum and give the result.

4.4. Problem 2. We want you to compute the area of the Mandelbrot set. You can do that by hand, printing out a version in which the coordinate axes are given and splitting it up into a grid, then counting. You can also use a machine. Here is some Mathematica code:

```
M=Compile[{x,y},Module[{z=x+I y,k=0}, While[Less[Abs[z],2]
  &&Less[k,1000],z=N[z^2+x+I y];++k]; Floor[k/1000]]];
9*Sum[M[-2+3 Random[],-1.5+3 Random[]],{100000}]/100000
```

On an Online page like *https : //www.tutorialspoint.com/execute_python_online.php* you can also execute some Python code. If there is somebody in your group who knows how to run Python locally, you might get better results. The code should be pretty self-explanatory. If you are interested 1j denotes i. We shoot random points into a square of side length 3 (and area 9) and count the number of hits of the Mandelbrot set.

```
from numpy import random
import math

iterations = 200
samples = 1000

def M(x,y):
    z = x+y*1j
    k = 0
    while abs(z)<2 and k<iterations:
      z = z*z+ x+y*1j
      k += 1
    return math.floor(k/iterations)

sum = 0
n = 0
while n < samples:
   n += 1
   sum += M(-2 + 3*random.rand(), -1.5 +3*random.rand())

final = 9*sum/samples
print(final);
```

The Mandelbrot set can be generated quickest with the open source ray tracer Povray:

```
camera{location <-1,0,3> look_at <-1,0,-3> right <0,16/9,0>
up <0,0,1>} plane{z,0 pigment{mandel 200 color_map{[0 rgb 0]
[1/6 rgb <1,0,1>][1/3 rgb <1,0,0>][1 rgb 0]}} finish{ambient 1}}
```

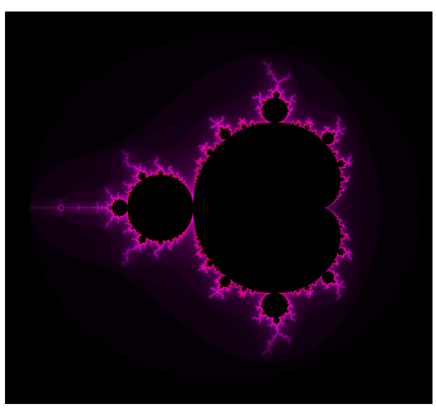


FIGURE 1. The Mandelbrot set.

4.5. Task: solve the two short problems. Author a 2 slide presentation or a document with two pages in which the two problems are solved, then submit this as a PDF. Please include the names of each of the group members who worked on the project.

Data project 5: Fibonacci

This data project was written in 2021 but not assigned.

1.1. The functions

$$J_n(x) = 2^{-n}\left(1-\sqrt{4x+1}\right)^n + 2^{-n}\left(\sqrt{4x+1}+1\right)^n$$

are called Lucas polynomials. You might not see immediately why these are polynomials. Let us compute a few examples $J_1(x) = 1, J_2(x) = 1+2x, J_3(x) = 1+3x, J_4(x) = 1+4x+2x^2$.

1.2. Problem 1: Evaluate the polynomials $J_5(x)$ and $J_6(x)$ to see that they are indeed polynomials and the square roots disappear.

1.3. There is a way to generate the polynomials recursively. Start with $J_0(x) = 2, J_1(x) = 1$ and then we have

$$J_n(x) = J_{n-1}(x) + xJ_{n-2}(x)$$

For example, $J_5(x)$ can be computed as $J_4(x) + xJ_3(x) = 1+4x+2x^2+x(1+3x)$.

1.4. Plugging in $x = 1$ gives us numbers $F_n = J_n(1)$. It satisfies the recursion

$$F_n = F_{n-1} + F_{n-2}$$

with initial condition $F_0 = 2, F_1 = 1$. This sequence is one of the **hyper Fibonacci type sequences** and also known under the name **Lucas sequence**. The usual Fibonacci sequence starts with $F_0 = 1, F_1 = 1$.

1.5. Problem 2: compute the first 10 numbers $F_0, F_1, \dots, F_{10}$. If we think of $F_{n+1}-F_n$ as a discrete derivative then $F_{n+1} - F_n = F_{n-1}$ essentially tells that the derivative of F is F itself. What classical function $f(x)$ do you know which has the property that

$$\frac{d}{dx}f(x) = f(x) \ ?$$

1.6. Here comes a bit of a surprise: integrate the functions $J_n(x)$ from -1 to 0. If you should tackle this data problem before having seen integrals in class, use the rule

$$\int_a^b cx^n \, dx = (cb^{n+1} - ca^{n+1})/(n+1) \ .$$

1.7. Problem 3. What do you get for the first 5 entries $\int_{-1}^0 J_n(x)\, dx$ using this rule? Find the definite integrals of the polynomials.

1.8. Problem 4 Verify the recursion relation

$$J_n(x) = J_{n-1}(x) + xJ_{n-2}(x)$$

with computer algebra system or by hand.

Hint when doing it by hand: Substitute $u = \sqrt{4x+1}$ so that $x = (u^2-1)/4$. Verify that $p_n(u) = (1+u)^n/2^n$ and verify $p_n(u) = p_{n-1}(u) + p(n-2)(u)(u^2-1)/4$. This shows that

$$P_n(x) = (1+\sqrt{4x+1})^n/2^n$$

satisfies

$$P_n(x) = P_{n-1}(x) + xP_{n-2}(x)$$

The same works also for

$$Q_n(x) = (1+\sqrt{4x+1})^n/2^n$$

which satisfies

$$Q_n(x) = Q_{n-1}(x) + xQ_{n-2}(x)$$

Finally, verify that $J_n(x) = P_n(x) + Q_n(x)$.

1.9. Problem 5. Who was the mathematician Fibonacci. Collect a few things and write it down.

FIGURE 1. Fibonacci statue

OLIVER KNILL, KNILL@MATH.HARVARD.EDU, MATH 1A, HARVARD COLLEGE, SPRING 2021