МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В.И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра МО ЭВМ

ОТЧЕТ

по практической работе №1

по дисциплине «Вычислительная математика»

Тема: Особенность машинной арифметики, точность вычисления на ЭВМ

Студент гр. 1304 ____________________ Шаврин А.П.

Преподаватель ____________________ Попова Е.В.

Санкт-Петербург

2022

Вариант 3.

Цель работы.

Изучить особенности вычислений с плавающей точкой.

Основные теоретические положения.

Множество F чисел с плавающей запятой характеризуется четырьмя параметрами: основанием системы счисления b, точностью t и интервалом показателей [L, M]. Каждое число x с плавающей запятой, принадлежащее F, имеет следующее значение (1)

$$x = \pm\left(\frac{d_1}{b} + \frac{d_2}{b^2} + \cdots + \frac{d_t}{b^t}\right) b^n, \quad (1)$$

где целые числа $d_1, d_2, \ldots, d_t$ удовлетворяют неравенствам $0 \le d_i \le b - 1$, $i \in \{1,2,\ldots,t\}$ и $L \le n \le M$. Целое число n называется показателем, а число $f = \left(\frac{d_1}{b} + \frac{d_2}{b^2} + \cdots + \frac{d_t}{b^t}\right)$ - дробной частью. Если принять, что $-N \le n < N$

и $N = 2^{m-1}$, то t называется разрядностью мантиссы, а m - разрядностью порядка. Определенная таким образом мантисса оказывается в диапазоне $0 \le f < 1$. Расположение представленных чисел на числовой оси уже не обладает свойством равномерности.

Действительная машинная реализация представлений чисел с плавающей точкой может отличатся в деталях от рассматриваемой идеальной, однако различия несущественны, и на практике их почти всегда можно игнорировать, анализируя основные проблемы ошибок округления. Величина b^{-t} является оценкой относительной точности плавающей арифметики, которая характеризуется посредством машинного эпсилон, т.е. наименьшего числа с плавающей точкой ε, такого, что 1+ε>1. Точное значение машинного эпсилон зависит не только то указанных выше параметров, но и от принятого способа округления. В вычислительных машинах используются различные системы чисел с плавающей точкой, причем в некоторых ЭВМ несколько систем. Так, для современных ПЭВМ характерно применение двух систем, которые называются обычной точностью и удвоенной точностью.

На множестве F определены арифметические операции в соответствии с тем, как они выполняются ЭВМ. Эти операции, в свою очередь моделируются в машине посредством приближений, называемых плавающими операциями. Для плавающих операций сложения, вычитания, умножения и деления существует возможность возникновения ошибок округления, переполнения и появления машинного нуля. Следует отметить, что операции плавающего сложения и умножения коммутативны, но не ассоциативны, и дистрибутивный закон для них также не выполняется. Невыполнение указанных алгебраических законов, имеющих фундаментальное значение для математического анализа, приводит к сложности анализа плавающих вычислений и возникающих при этом ошибок.

Постановка задачи.

Используя готовые программы, выполнить исследования машинной арифметики и точности вычислений на ПЭВМ. Порядок выполнения работы следующий:

1) Исследование распределения нормализованных чисел с плавающей точкой на вещественной оси для различных значений параметров b, m, t.

2) Вычисление значения величины машинного эпсилон при различных значениях константы c.

3) Исследование абсолютных и относительных ошибок округления при вычислениях с плавающей точкой сумм чисел при различных значениях шага суммирования

4) Исследование проявления ошибок округления, возникающих при вычислении показательной функции e x для чисел с плавающей точкой для двух вариантов алгоритма вычислений, а также скорости сходимости обоих вариантов

Выполнение работы.

1. Были проведены исследования распределения нормализованных чисел с плавающей точкой на вещественной оси для параметров b = 2, m = 4, t = 4. Результаты расчетов см. на рисунке 1.

```
x[ 0]=0.000000
x[ 1]=0.000015
x[ 2]=0.000017
x[ 3]=0.000019
x[ 4]=0.000021
x[ 5]=0.000023
x[ 6]=0.000025
x[ 7]=0.000027
x[ 8]=0.000029
x[ 9]=0.000031
x[10]=0.000034
x[11]=0.000038
x[12]=0.000042
x[13]=0.000046
x[14]=0.000050
x[15]=0.000053
x[16]=0.000057
```

```
x[16]=0.000057
x[17]=0.000061
x[18]=0.000069
x[19]=0.000076
x[20]=0.000084
x[21]=0.000092
x[22]=0.000099
x[23]=0.000107
x[24]=0.000114
x[25]=0.000122
x[26]=0.000137
x[27]=0.000153
x[28]=0.000168
x[29]=0.000183
x[30]=0.000198
x[31]=0.000214
x[32]=0.000229
```

```
x[32]=0.000229
x[33]=0.000244
x[34]=0.000275
x[35]=0.000305
x[36]=0.000336
x[37]=0.000366
x[38]=0.000397
x[39]=0.000427
x[40]=0.000458
x[41]=0.000488
x[42]=0.000549
x[43]=0.000610
x[44]=0.000671
x[45]=0.000732
x[46]=0.000793
x[47]=0.000854
x[48]=0.000916
```

```
x[48]=0.000916
x[49]=0.000977
x[50]=0.001099
x[51]=0.001221
x[52]=0.001343
x[53]=0.001465
x[54]=0.001587
x[55]=0.001709
x[56]=0.001831
x[57]=0.001953
x[58]=0.002197
x[59]=0.002441
x[60]=0.002686
x[61]=0.002930
x[62]=0.003174
x[63]=0.003418
x[64]=0.003662
```

```
x[64]=0.003662
x[65]=0.003906
x[66]=0.004395
x[67]=0.004883
x[68]=0.005371
x[69]=0.005859
x[70]=0.006348
x[71]=0.006836
x[72]=0.007324
x[73]=0.007812
x[74]=0.008789
x[75]=0.009766
x[76]=0.010742
x[77]=0.011719
x[78]=0.012695
x[79]=0.013672
x[80]=0.014648
```

```
x[80]=0.014648
x[81]=0.015625
x[82]=0.017578
x[83]=0.019531
x[84]=0.021484
x[85]=0.023437
x[86]=0.025391
x[87]=0.027344
x[88]=0.029297
x[89]=0.031250
x[90]=0.035156
x[91]=0.039062
x[92]=0.042969
x[93]=0.046875
x[94]=0.050781
x[95]=0.054687
x[96]=0.058594
```

```
x[96]=0.058594
x[97]=0.062500
x[98]=0.070312
x[99]=0.078125
x[100]=0.085937
x[101]=0.093750
x[102]=0.101562
x[103]=0.109375
x[104]=0.117187
x[105]=0.125000
x[106]=0.140625
x[107]=0.156250
x[108]=0.171875
x[109]=0.187500
x[110]=0.203125
x[111]=0.218750
x[112]=0.234375
```

```
x[112]=0.234375
x[113]=0.250000
x[114]=0.281250
x[115]=0.312500
x[116]=0.343750
x[117]=0.375000
x[118]=0.406250
x[119]=0.437500
x[120]=0.468750
x[121]=0.500000
x[122]=0.562500
x[123]=0.625000
x[124]=0.687500
x[125]=0.750000
x[126]=0.812500
x[127]=0.875000
x[128]=0.937500
```

```
x[128]=0.937500
x[129]=1.000000
x[130]=1.125000
x[131]=1.250000
x[132]=1.375000
x[133]=1.500000
x[134]=1.625000
x[135]=1.750000
x[136]=1.875000
x[137]=2.000000
x[138]=2.250000
x[139]=2.500000
x[140]=2.750000
x[141]=3.000000
x[142]=3.250000
x[143]=3.500000
x[144]=3.750000
```

```
x[144]=3.750000
x[145]=4.000000
x[146]=4.500000
x[147]=5.000000
x[148]=5.500000
x[149]=6.000000
x[150]=6.500000
x[151]=7.000000
x[152]=7.500000
x[153]=8.000000
x[154]=9.000000
x[155]=10.00000
x[156]=11.00000
x[157]=12.00000
x[158]=13.00000
x[159]=14.00000
x[160]=15.00000
```

```
x[160]=15.00000
x[161]=16.00000
x[162]=18.00000
x[163]=20.00000
x[164]=22.00000
x[165]=24.00000
x[166]=26.00000
x[167]=28.00000
x[168]=30.00000
x[169]=32.00000
x[170]=36.00000
x[171]=40.00000
x[172]=44.00000
x[173]=48.00000
x[174]=52.00000
x[175]=56.00000
x[176]=60.00000
```

```
x[176]=60.00000
x[177]=64.00000
x[178]=72.00000
x[179]=80.00000
x[180]=88.00000
x[181]=96.00000
x[182]=104.0000
x[183]=112.0000
x[184]=120.0000
x[185]=128.0000
x[186]=144.0000
x[187]=160.0000
x[188]=176.0000
x[189]=192.0000
x[190]=208.0000
x[191]=224.0000
x[192]=240.0000
```

```
x[192]=240.0000
x[193]=256.0000
x[194]=288.0000
x[195]=320.0000
x[196]=352.0000
x[197]=384.0000
x[198]=416.0000
x[199]=448.0000
x[200]=480.0000
x[201]=512.0000
x[202]=576.0000
x[203]=640.0000
x[204]=704.0000
x[205]=768.0000
x[206]=832.0000
x[207]=896.0000
x[208]=960.0000
```

```
x[208]=960.0000
x[209]=1024.000
x[210]=1152.000
x[211]=1280.000
x[212]=1408.000
x[213]=1536.000
x[214]=1664.000
x[215]=1792.000
x[216]=1920.000
x[217]=2048.000
x[218]=2304.000
x[219]=2560.000
x[220]=2816.000
x[221]=3072.000
x[222]=3328.000
x[223]=3584.000
x[224]=3840.000
```

```
x[224]=3840.000
x[225]=4096.000
x[226]=4608.000
x[227]=5120.000
x[228]=5632.000
x[229]=6144.000
x[230]=6656.000
x[231]=7168.000
x[232]=7680.000
x[233]=8192.000
x[234]=9216.000
x[235]=10240.00
x[236]=11264.00
x[237]=12288.00
x[238]=13312.00
x[239]=14336.00
x[240]=15360.00
```

```
x[233]=8192.000
x[234]=9216.000
x[235]=10240.00
x[236]=11264.00
x[237]=12288.00
x[238]=13312.00
x[239]=14336.00
x[240]=15360.00
x[241]=16384.00
x[242]=18432.00
x[243]=20480.00
x[244]=22528.00
x[245]=24576.00
x[246]=26624.00
x[247]=28672.00
x[248]=30720.00
Нажмите любую
 клавишу...
```

Рисунок 1 — Числа, сгенерированные программой с разными значениями параметров.

График распределения этих чисел представлен на рис. 2.

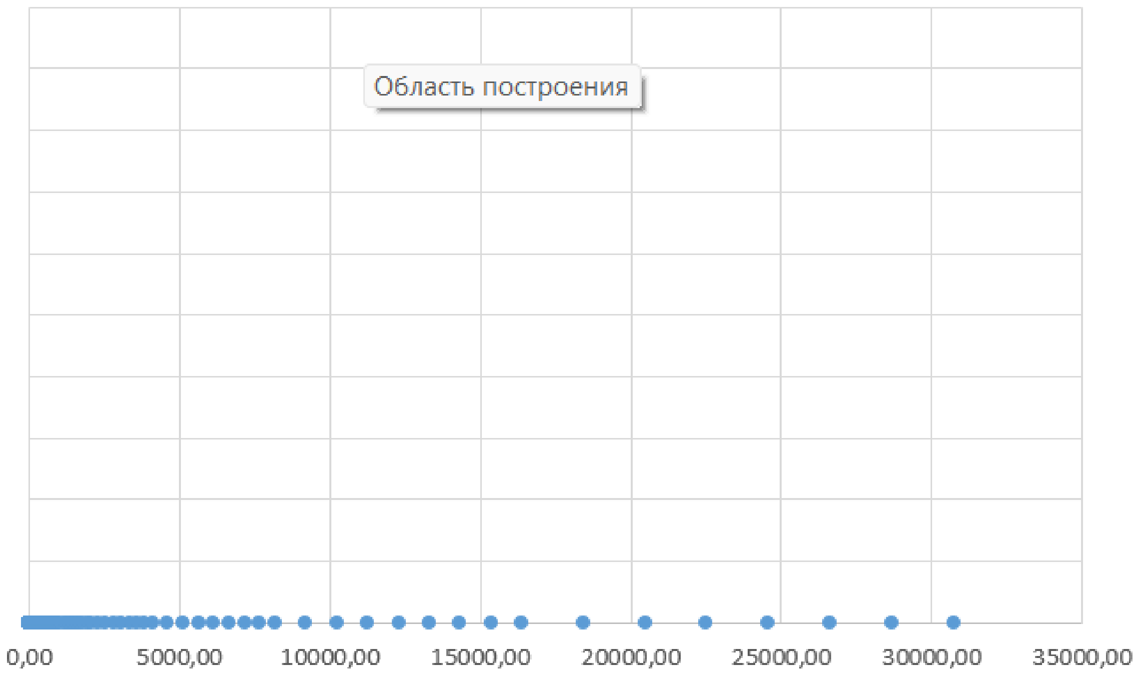

Рисунок 2 - график распределения нормализованных чисел с плавающей точкой.

Вывод: распределение нормализованных чисел с плавающей точкой на вещественной оси неравномерно.

2. Были вычислены значения ε, при разных значениях аргумента c.

Результаты вычислений см. в табл. 1.

Таблица 1 — Результаты вычисления ε при разных значениях c.

Значение c	Значение ε	Шаг итерации
5	4,3*10^(-19)	61
25	17,3*10^(-19)	59
125	69,3*10^(-19)	57
625	555,1*10^(-19)	54
3125	2220,4*10^(-19)	52

График зависимости ε от c показанна рис. 3.

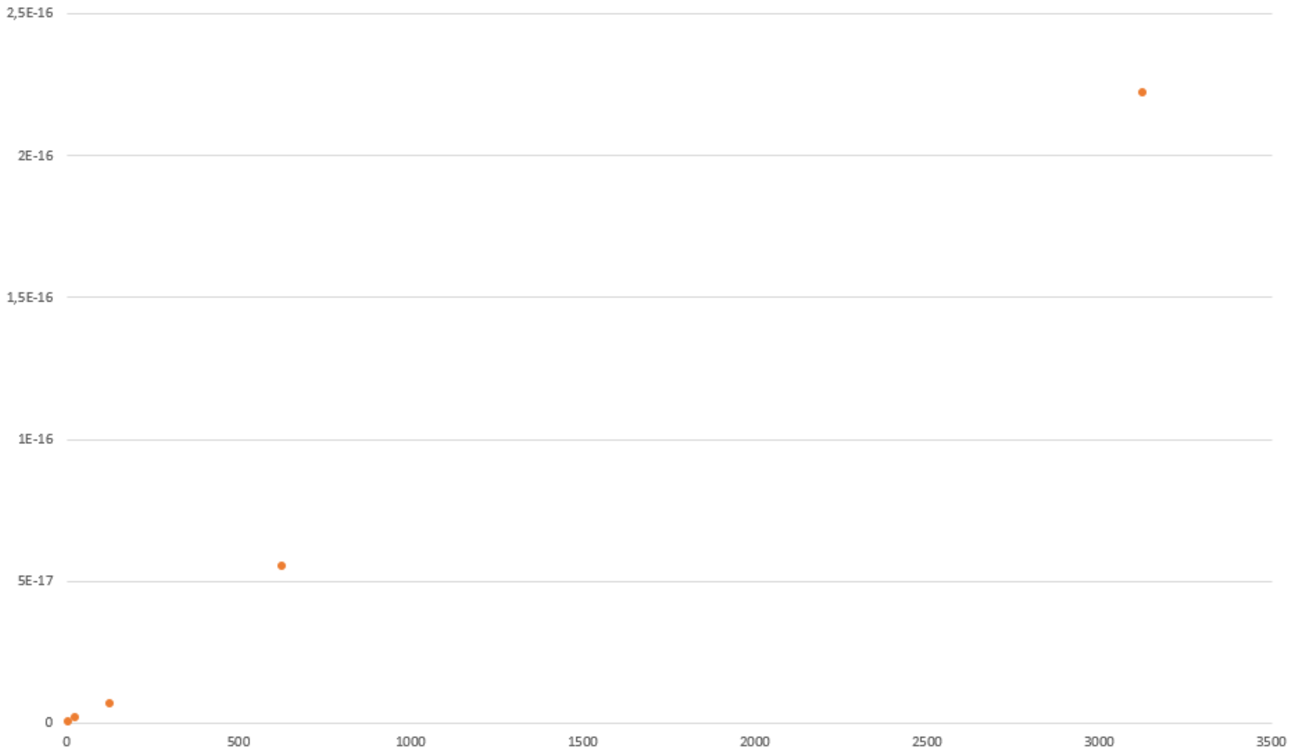

Рисунок 3 — График зависимости машинного ε от c.

Вывод: с увеличением значения c, значение ε также увеличивается.

3. Было проведено исследование абсолютных и относительных ошибок округления при вычислениях с плавающей точкой сумм чисел $\left(\sum_{i=1}^{N}\frac{1}{N}, x[i] = x[i-1] + \frac{1}{N}\right)$ при различных значениях шага суммирования.

   Результаты вычислений см. в табл. 2.

   Таблица 2 — Результаты исследования абсолютных и относительных ошибок округления (N – шаг суммирования x–dx – абсолютная погрешность, (x-dx)/x – относительная погрешность)

N	(x-dx)	(x-dx)/x
5	0,0000000596	0,000006 %
25	0,0000000224	0,000002 %

125	0,0000000689	0,000007 %
625	0,0000000253	0,000003 %
3125	0,0000000253	0,000003 %
15625	0,0000000025	0,000000 %

Вывод: абсолютная погрешность незначительно увеличивалась с шагами в ходе суммирования, что является показателем накопления ошибок. Абсолютная ошибка накапливается равномерно.

4. Было проведено исследование проявления ошибок округления, возникающих при вычислении показательной функции e^x для чисел с плавающей точкой для двух вариантов алгоритма вычислений:
   a. $e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}$
   b. $x = m + f;\ e^x = e^m * e^f,$ где $m -$ целая часть числа, $f -$ дробная часть числа.

Также были найдены скорости сходимости обоих вариантов.

Результаты обработки программой введенных данных см. в табл. 3.

Таблица 3 — Исследование проявления ошибок округления, возникающих при вычислении функции e^x для двух алгоритмов.

x	ε	Разложение Тейлора	Улучшенный алгоритм	Абсолютная погрешность	Относительная погрешность
0,01	0,001	1,01004999977 2 итерации	1,01004999977 2 итерации	0	0%
1.3	0,001	3.6692629675359 57 итераций 8 итераций	3.66923857270339 8 итераций 4 итерации	0.000024394822558	0.000665%

2.7	0,001	14.879648543953516 12 итераций	14.879596701154579 6 итераций	0.000051842798937	0.000348%
4,49	0,001	89,11 15 итераций	89,10 4 итерации	0,01221	0,013713%
16.9	0,001	21856296.7444100045 50 итераций	21855348.1673088297 6 итераций	948.5771011713259 21	0.004340%
20.02	0,001	494965464.996178389 59 итераций	494964815.003642976 2 итерации	649.9925354098959 36	0.000131%
30.31	0,001	14570200044657.209000000000000 87 итераций	14569931313698.548800000000000 4 итераций	268730958.659009933000000	0.001844%

Вывод: с увеличением x абсолютная погрешность возрастает. Данные таблицы показывают, что улучшенный алгоритм производит меньше итераций, чем обычное разложение Тейлора, и разрыв в количестве итераций лишь увеличивается с ростом x.

Выводы.

В ходе выполнения заданий лабораторной работы, были исследованы машинная арифметика, точность вычислений с плавающей точкой, распределение нормализованных чисел на вещественной оси. Были оценены абсолютные и относительные ошибки округления при вычислениях с плавающей точкой, зависимость машинного эпсилон от значения константы, с помощью которой он вычислялся. Таким образом, были изучены основные особенности вычислений с плавающей точкой.