APPLICATIONS OF DYNAMICAL IDEAS IN NUMBER THEORY

COMPUTER ASSISTED HOME ASSIGNMENT MMA421/TMA014

1. Format of this assignment

The computer assignments described in this document are compulsory. The reports must be delivered in the form of a pdf-document sent by e-mail to `micbjo@chalmers.se` no later than the date of the written exam. The reports will be graded (with grades pass (G) and pass with distinction (VG) for students registered with the University of Gothenburg, and (3), which is equivalent to pass, (4) or (5) for students registered with Chalmers), and the grade will be used when the final grade for the course is determined. To pass the course, the student must obtain pass on the computer assignment.

You are allowed (and even encouraged) to work in pairs while doing the assignments. However, every student must write his/her own report, and hence must participate in all parts of the work. If you worked with somebody, you should state that in the beginning of the report. You must write your name and personal identification number in the report.

The report does not need to be a complete technical report, and in particular you should not spend a lot of time making it look very nice typographically. It should be sufficiently detailed to be read without the assignment description (this document) in hand. You should present the results of your computer experiments, explain how they were carried out (you should put the the Matlab code as an appendix to the report, but it should be possible to read the report without looking at the code), and comment on them, and of course, answer as many as possible of the explicit exercises given in the text.

The exercises are worth 20 points in total. In order to pass (G at GU and 3 at Chalmers), 10 points are required. For VG (at GU), 14 points are required. For 4 (at Chalmers), you will need 13 points, and 16 points for the grade 5.

Please report any misprints or other inaccuracies to micbjo@chalmers.se.

1.1. **Matlab**

You are encouraged to solve the computer assignments by using Matlab, which is available on the computer system. It is allowed to use other types of mathematical software, but a person who chooses a different kind of software must be particularly careful to respect the intention of the assignment. Some familiarity with Matlab is assumed. Students who are new to Matlab should consult a book to learn the basics. A good example is "Learning MATLAB" by T. A. Driscoll, that can be obtained from SIAM (Society for Industrial and Applied Mathematics; http://www.ec-securehost.com/SIAM/OT115.html). Many universities give a course called "Crash Course in Matlab" or something similar, and some also provide good material on their web pages. Also the online Matlab manual is very useful.

2. Leading digits in large powers of 2

Here is a curious question due to the Russian school around Gelfand in the sixties: Does 9 ever appear as the first digit of 2^n for some n? With a powerful enough computer one can verify that 2^{54} indeed starts with a 9. Is this the only power of two with leading digit 9? Are the infinitely many such numbers? How about 7 as a leading digit? Let us not make more of these questions than what they are- mere curiosities. However, they will serve as first motivations for what you will do in this assignment.

How do we attack the problems above? Let us take a more general view on things. Denote by a_k the first digit in 2^k for $k \geq 0$ (written out in base 10). For $q = 1, 2, \ldots, 9$ and $n \geq 0$, we define

$$m_q(n) = |\{k = 0, 1, \ldots, n : a_k = q\}|.$$

The first question above asks whether $m_9(n) > 0$ for some n, the third question whether $m_9(n) \to \infty$, and the fourth question whether $m_7(n) > 0$ for some n, or even $m_7(n) \to \infty$. Your first task will be to show that, in fact,

$$\lim_n \frac{m_q(n)}{n+1} = \log_{10} \frac{q+1}{q}, \quad \text{for all } q = 1, \ldots, 9. \tag{2.1}$$

In particular, $m_q(n) \to \infty$ for every such q. Here is an outline of how to prove it.

Exercise 1 (2p)**.** Let I_q denote the half-open interval $[\log_{10} q, \log_{10}(q+1)) \subset [0,1)$. Show that the first digit of 2^k is q if and only if $k \log_{10} 2$ modulo one belongs to I_q.

Exercise 2 (1p)**.** Show that $\alpha = \log_{10} 2$ is irrational.

In particular, this exercise shows that

$$m_q(n) = \sum_{k=0}^{n} \chi_{I_q}(k \log_{10} 2 \mod 1), \quad \text{for all } n,$$

where χ_{I_q} denotes the indicator function of I_q. Note that the length $|I_q|$ equals $\log_{10}(q+1) - \log_{10} q = \log_{10} \frac{q+1}{q}$, which is precisely the number that appears on the right hand side in (2.1). Hence, in view of the second exercise above, we would have solved our problem if we could prove that for any interval $I \subset [0,1)$ and any irrational α

$$\lim_n \frac{1}{n+1} \sum_{k=0}^{n} \chi_I(k\alpha \mod 1) = |I|. \tag{2.2}$$

Exercise 3 (3p)**.** Try to convince yourself (numerically - using say Matlab) of the validity of (2.2), by computing the averages for large n (say 100, 1000, ...), and for a few different intervals (say with $\alpha = \log_{10} 2$).

In order to prove (2.2) in general it will be useful to replace χ_I with a continuous function f on the compact interval $[0,1]$ with $f(0) = f(1)$ - or equivalently, a 1-periodic continuous function on the whole real line.

Exercise 4 (3p)**.** Show that for any interval $I \subset [0,1)$ and $\varepsilon > 0$, there are 1-periodic continuous functions $f_\pm$ on the real line such that

$$f_- \leq \chi_I \leq f_+ \quad \text{and} \quad \int_0^1 \left(f^+(x) - f^-(x)\right) dx < \varepsilon.$$

Suppose now for a while that we have been able to prove that

$$\lim_n \frac{1}{n+1} \sum_{k=0}^{n} f(k\alpha \mod 1) = \int_0^1 f(x)\,dx \tag{2.3}$$

for every 1-periodic *continuous* function f. Assuming Exercise (4), we conclude that for any interval $I \subset [0,1)$ and every $\varepsilon > 0$,

$$\frac{1}{n+1} \sum_{k=0}^{n} f_-(k\alpha \mod 1) \leq \frac{1}{n+1} \sum_{k=0}^{n} \chi_I(k\alpha \mod 1) \leq \frac{1}{n+1} \sum_{k=0}^{n} f_+(k\alpha \mod 1)$$

for all n, and thus

$$A_- := \underline{\lim}_n \frac{1}{n+1} \sum_{k=0}^{n} \chi_I(k\alpha \mod 1) \ge \int_0^1 f_-(x)\,dx$$

and

$$A_+ := \overline{\lim}_n \frac{1}{n+1} \sum_{k=0}^{n} \chi_I(k\alpha \mod 1) \le \int_0^1 f_+(x)\,dx$$

It follows that $A_+ - A_- \le \varepsilon$. Since $\varepsilon > 0$ is arbitrary, we conclude that $A_- = A_+$, and thus the limit (2.2) must exist.

Exercise 5 (2p)**.** Show that $A_+ = |I|$.

We are now left with the task of proving (2.3). For this, we first need to agree on some notation. Let $C_1(\mathbb{R})$ denote the complex linear space of complex-valued and *1-periodic* continuous functions f on $\mathbb{R}$. Given $m \in \mathbb{Z}$, the function $e_m(x) := e^{2\pi i m x}$ clearly belongs to $C_1(\mathbb{R})$, as does every linear combination of such functions. We say that f is a *trigonometric polynomial* if there are integers $m_1, \dots, m_N$ and complex numbers $c_1, \dots, c_N$ such that

$$f(x) = \sum_{j=1}^{N} c_j e_{m_j}(x), \quad \text{for } x \in \mathbb{R}.$$

Clearly $f \in C_1(\mathbb{R})$. We denote by $\mathscr{P}_1 \subset C_1(\mathbb{R})$ the complex linear sub-space of all trigonometric polynomials.

Exercise 6 (3p)**.** Show (2.3) for every $f \in \mathscr{P}_1$. **Hint: Geometric summation.**

Let us now use the following fact from Fourier analysis as a black box. Let $\|\cdot\|_\infty$ denote the norm

$$\|f\|_\infty = \sup\{|f(x)| : x \in [0,1]\}, \quad \text{for } f \in C_1(\mathbb{R}).$$

Then, for every $f \in C_1(\mathbb{R})$ and $\varepsilon > 0$, there exist $p \in \mathscr{P}_1$ such that $\|f - p\|_\infty < \varepsilon$.

Exercise 7 (3p)**.** Use this result, in combination with the previous exercise, to prove that (2.3) holds for all $f \in C_1(\mathbb{R})$.

3. Normal numbers

If you pick your favorite irrational number, then the chances are high that if you start to write it out in decimal form (say in base 10), then the digits that you will see behave quite "randomly". A way to formalize this, is to introduce the notion of a *normal number*. We say that $x \in [0,1]$ is *normal* (in base 10) if for every $q = 0, 1, \dots, 9$,

$$\lim_{n\to\infty} \frac{d_n(q,x)}{n} = \frac{1}{10}, \tag{3.1}$$

where $d_n(q,x)$ equals 1 if the n'th digit in the 10-expansion of x equals q, and zero otherwise. In other words, x is normal if all numbers between 0 and 9 appear in the decimal expansion of x with the same frequency.

Exercise 8 (3p)**.** Consider the digits of $\pi/4$ in base 10 - does $\pi/4$ appear to be normal? [Caution: nobody knows for sure]. More precisely, compute (3.1) for $x = \pi/4$, $q = 1, 2, \dots, 9$ and $n = 10, 100, 1000$. Present your findings in a table with q as rows, n as columns, and in each cell, you write out $d_n(q,x)/n$.

The decimal expansion in base 10 of a number $x \in [0,1]$ can be written as

$$x = \sum_{k=1}^{\infty} x_k\, 10^{-k}, \quad \text{for } x_k = 0, 1, \dots, 9.$$

There are some discrepancies; for instance, $0.099999999\cdots$ and 0.1 are the same number.

Exercise 9 (2p). Show that $x_k = q$ if and only if $10^{k-1}x \mod 1 \in J_q$, where $J_q = [q/10, (q+1)/10)$.

Hence we can write

$$\frac{d_n(x,q)}{n} = \frac{1}{n}\sum_{k=0}^{n-1} \chi_{J_q}(10^k x), \quad \text{for all } n \geq 1,$$

so x is normal if and only if

$$\frac{1}{n}\sum_{k=0}^{n-1} \chi_{J_q}(10^k x) \to |J_q| = \frac{1}{10}, \quad \text{for every } q = 0, 1, \ldots, 9.$$

It is in general very difficult to prove normality for a fixed x. The aim of the rest of this assignment is to prove that "most" $x \in [0,1]$ are normal in the following sense: For every interval $J \subset [0,1]$,

$$\int_0^1 \left| \frac{1}{n}\sum_{k=0}^{n-1} \chi_J(10^k x) - |J| \right|^2 dx \to 0. \tag{3.2}$$

Exercise 10 (2p). Use Exercise 4 (as a black box if you haven't been able to solve it), to show that it suffices to establish (3.2) with a continuous 1-periodic function f instead of χ_J and $\int_0^1 f(t)dt$ instead of $|J|$, i.e.

$$\int_0^1 \left| \frac{1}{n}\sum_{k=0}^{n-1} f(10^k x) - \int_0^1 f(t)dt \right|^2 dx \to 0. \tag{3.3}$$

Exercise 11 (3p). Prove (3.3) for $f \in \mathscr{P}_1$.

Exercise 12 (2p). Use the black box from Fourier analysis in the last section to prove (3.3) for $f \in C_1(\mathbb{R})$.

References