Lecture 03 (hrs. 03,04) - September 25, 2025, 8:30 - 10:30 F3

(1.08) *Scilab code*.

```
function [z, v, info, k, mis] = bisezione(f, a, b, E, kmax)
//
// Use:
//     [ z,v,info,[k,[mis]] ] = bisezione(f,a,b,E,kmax)
//
//
// Approximate a zero of the function f: [a,b] -> R, which must
// be continuous, using the bisection method. The function f
// must assume non-zero values of opposite signs at a and b.
//
// The iteration stops when:
// (*) the function f evaluates to zero at the midpoint x_m
//     of the actual interval [a(k),b(k)];
// (*) the measure of the actual interval [a(k),b(k)] is lower than
//     E: in this case, in theory, z approximates a zero of
//     f with an absolute error no greater than E/2;
// (*) after kmax iterations.
//
// kmax: optional argument (predefined value: 50).
//
// z: final approximation (zero of f or midpoint
//    of the last generated interval);
// v: value of f at z;
// info = 0: a zero of has been found (f(z) = 0);
//      = 1: f(z) ~= 0 and the last interval considered has measure
//           less than E (mis < E);
//      = 2: f(z) ~= 0, mis >= E and the number of iterations has
//           reached the maximum allowed (k = kmax);
// k: number of iterations performed;
// mis: amplitude of the last determined interval.
//
//
// Initializations
//
if ~exists('kmax','l') then kmax = 50; end;
k_bis = 0; // counter of executed iterations
//
// Construction of the sequences
//
x_m = (a + b)/2;
f_m = f(x_m);
while (abs(b-a) >= E & f_m ~= 0 & k_bis < kmax),
  k_bis = k_bis+1;
  if sign(f_m) == sign(f(b)) then b = x_m; else a = x_m; end;
  x_m = (a + b)/2;
```

```
  f_m = f(x_m);
end;
//
// End of construction: assign output variables
//
z = x_m; v = f_m; k = k_bis; mis = abs(b-a);
if f_m == 0 then info = 0;
   else if abs(b-a) >= E then info = 2; else info = 1; end;
end;
//
endfunction
```

(1.09) Remark.

The *Scilab* loop

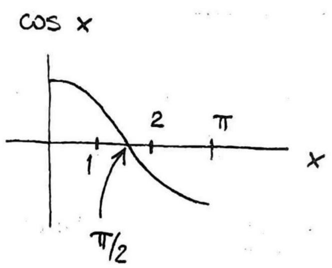
```
while condition,
  instructions;
end;
```

is equivalent to:

```
repeat:
  if condition is true then instructions;
       else leave the loop;
```

(1.10) Example.

Let f(x) = cos(x).



- The function is continuous on [a,b] = [1,2] and f(a) > 0, f(b) < 0
- Given E > 0 it is:

$$\text{measure } I(k) < E \quad \Leftrightarrow \quad \text{measure } I(0) / 2^k < E \quad \Leftrightarrow$$

$$\Leftrightarrow \quad 2^k > \text{measure } I(0) / E \quad \Leftrightarrow \quad k > \log_2(\text{ measure } I(0) / E)$$

  so: we expect to obtain an approximation of pi/2 with absolute error less than E in

  $$VA = \text{the least integer greater than or equal to } \log_2(\text{ mis } I(0) / E)$$

  iterations.
- We get (using the file EsempioBisezione.sce downloadable from the ‘altro materiale didattico’ section of the course web page):

E	info	mis	k	VA	kmax
10^{-5}	1	7.6 10^{-6}	17	17	50
10^{-10}	1	5.8 10^{-11}	34	34	50
10^{-15}	1	8.8 10^{-16}	50	50	50
10^{-16}	2	2.2 10^{-16}	60	54	60
10^{-16}	2	2.2 10^{-16}	100	54	100

From the last two rows of the table we can observe that when E = 10^{-16} the function *bisezione* stops because it has reached the maximum number of iterations allowed but, while in the first case (penultimate row) this is *consistent* with the theory, in the second case (last row) it is *not consistent* with the theory: the procedure *should have stopped after* 54 *iterations with* info = 1.

To understand why this happens, we need to study COMPUTER ARITHMETIC in more detail.

(1.11) Questions.

(A) *What numbers* can the calculator work with?
(B) *What can he do* with these numbers?

(1.12) Remark.

Let x be a *non-zero* real number, and β be an integer greater than or equal to two (*radix*). There is exactly *one* factorization of x in the form:

$$x = (-1)^s \beta^b g$$

where:

- s in {0,1}, *sign* of x
- b: integer number, radix β-*exponent* of x
- g: real number in $[1/\beta,1)$, radix β-*fraction* of x

(Proof:

- if x > 0 then s = 0, if x < 0 then s = 1;
- b is *the unique* integer number such that:

$$\beta^{b-1} < |x| \leq \beta^b$$

- $g = |x| / \beta^b$)

(1.13) Example.

(1) x = sqrt(5), β = 10 $\Rightarrow$ s = 0, b = 1, g = sqrt(5) / 10
(2) x = sqrt(5), β = 2 $\Rightarrow$ s = 0, b = 2, g = sqrt(5) / 4

(1.14) Remark.

The condition g real number in $[1/\beta,1)$ translates as follows: the positional radix β writing of g has the form:

$$0.c_1c_2c_3\ldots \text{ where } c_1 \neq 0$$

In particular: *if* $\beta = 2$ *it must be* $c_1 = 1$.

(1.15) Example.

(1) $x = 1/10$, $\beta = 10 \Rightarrow s = 0$, $b = 0$, $g = 1/10 = 0.1$

(2) $x = 1/10$, $\beta = 2 \Rightarrow s = 0$, $b = -3$, $g = 8/10 = 4/5 = 0.\overline{1100}$

(Step-by-step proof[1]:

(1)

- $4/5 = 0.c_1c_2c_3... \Rightarrow 8/5 = c_1.c_2c_3...$ and then:
- $[8/5] = [c_1.c_2c_3...]$ e $\{8/5\} = \{c_1.c_2c_3...\}$ i.e.:
- $c_1 = 1$ e $3/5 = 0.c_2c_3c_4...$

(2)

- $3/5 = 0.c_2c_3c_4... \Rightarrow 6/5 = c_2.c_3c_4...$ and then:
- $[6/5] = [c_2.c_3c_4...]$ e $\{6/5\} = \{c_2.c_3c_4...\}$ i.e.:
- $c_2 = 1$ e $1/5 = 0.c_3c_4c_5...$

(3)

- $1/5 = 0.c_3c_4c_5... \Rightarrow 2/5 = c_3.c_4c_5...$ and then:
- $[2/5] = [c_3.c_4c_5...]$ e $\{2/5\} = \{c_3.c_4c_5...\}$ i.e.:
- $c_3 = 0$ e $2/5 = 0.c_4c_5c_6...$

(4)

- $2/5 = 0.c_4c_5c_6... \Rightarrow 4/5 = c_4.c_5c_6...$ and then:
- $[4/5] = [c_4.c_5c_6...]$ e $\{4/5\} = \{c_4.c_5c_6...\}$ i.e.:
- $c_4 = 0$ e $4/5 = 0.c_5c_6c_7...$

We now observe that we have obtained a new writing for the initial number 4/5. We deduce that 4/5 has a periodic writing of period four.

End of the step-by-step proof.)

Note that in both examples it is $x = 1/10$, but in example (1) the fraction has a *finite-length* positional notation, while in example (2) it has an *infinite length*. The length of the positional notation depends on the radix.

---

1 If q is a real number, we indicate by [q] *the integer part* of q and by {q} *the fractional part* of q, that is {q} = q - [q].