

(1) FUNCTION ZEROES AND COMPUTER ARITHMETIC

(1.01) The problem.

Given a continuous function  $f:[a,b] \rightarrow \mathbb{R}$  such that there exists  $t$  in  $\mathbb{R}$  s.t.  $f(t) = 0$ , determine  $t$ . The number  $t$  is called 'a zero of  $f$ '.

(1.02) Theorem (existence of a zero)

Let  $f:[a,b] \rightarrow \mathbb{R}$  be a continuous function s.t.  $f(a)f(b) < 0$ . Then: there exists  $t$  in  $(a,b)$  s.t.  $f(t) = 0$ .

(1.03) Remark.

The condition  $f(a)f(b) < 0$  is equivalent to the condition:

$f(a)$  is *not* = 0 &  $f(b)$  is *not* = 0 & sign  $f(a)$  *not equal* to sign  $f(b)$

(1.04) Bisection method.

Idea: use Theorem (1.02) to obtain a sequence of intervals  $I(k) = [a(k), b(k)]$  such that:

- for every  $k$ , there exists a zero of  $f$  in  $I(k)$
- $I(k+1)$  is a subset of  $I(k)$
- when  $k \rightarrow \infty$  it is measure  $I(k) \rightarrow 0$

(1.05) Description of the method.

$z = \text{Bisection}(f, a, b)$

input:  $f:(a,b) \rightarrow \mathbb{R}$  s.t.  $f(a)f(b) < 0$

- $a(0) = a$ ;  $b(0) = b$ ;  $I(0) = [a(0), b(0)]$ ;  $x(0) = (a(0) + b(0)) / 2$ ;
- for  $k = 1, 2, 3, \dots$  repeat:
  - if  $f(x(k-1)) = 0$  then STOP; else
    - if  $f(x(k-1))f(b(k-1)) < 0$  then  $a(k) = x(k-1)$ ;  $b(k) = b(k-1)$ ;
    - else  $a(k) = a(k-1)$ ;  $b(k) = x(k-1)$ ;
  - $I(k) = [a(k), b(k)]$ ;  $x(k) = (a(k) + b(k)) / 2$ ;

output: when a suitable *halt condition* is met:  $z = x(k)$ , midpoint of the last interval.

(1.06) Remark.

(A) measure  $I(k) = b(k) - a(k) = \text{measure } I(k-1) / 2^1 = \text{measure } I(k-2) / 2^2 = \dots = \text{measure } I(0) / 2^k$  and then:

when  $k \rightarrow \infty$  it is measure  $I(k) \rightarrow 0$

(B) if  $f$  is a continuous function then: for every  $k$ ,  $I(k)$  contains a zero of  $f$  and

when  $k \rightarrow \infty$  it is  $x(k) \rightarrow t$ , where  $f(t) = 0$

(Proof ...)

(1.07) Halt condition.

The bisection method is an *iterative method*, that is, a method that approximates the desired object by constructing a *sequence*. Since it is physically impossible to construct *all* the elements of the sequence, it is *necessary* to introduce a halt condition, that is, a condition that, when satisfied, stops the construction of the sequence.

An example of a halt condition is: given a *positive* real number  $\Delta$  ...

if measure  $I(k) < \Delta$  then STOP

Properties of the halt condition:

- (1) the *condition* measure  $I(k) < \Delta$  'is *computable*'
- (2) the condition is *certainly verified* after a *finite* number of iterations (see Remark (B) in (1.06)): the criterion 'is *effective*'
- (3) if  $f$  is a continuous function and  $k$  is such that measure  $I(k) < \Delta$  then:

- there exists  $t$  in  $I(k)$  s.t.  $t$  is a zero of  $f$
- $|x(k) - t| < \text{measure } I(k) / 2 < \Delta/2 < \Delta$

that is, the procedure returns a value  $x(k)$  that is an approximation of a zero of  $f$  with *absolute error*  $|x(k) - t|$  less than  $\Delta$ : 'the procedure returns an approximation as accurate as required by the user'.