Nicolas HUBERT – Antonin LAPICHE Groupe C

TP 1 : Précision des calculs dans une machine informatique

Question 2

Tmin et Tmax correspondent aux valeurs absolues minimale et maximale du terme lors des k itérations.

Le fichier termeg1 est la trace d'éxecution du programme. On peut y voir tous les résultats.

La ligne fprint permet d'écrire les résultats dans un fichier texte.

x	ex(cal)	ex(théo)	kMAX	erreur
-10	4,539993000E-05	4,539992976E-05	59	0%
-15	3,058852000E-07	3,059023205E-07	75	0%
-20	5,890016000E-10	2,061153622E-09	90	**71%**
-25	2,847098000E-07	1,388794390E-11	104	**2049950%**
-30	2,271280000E-05	9,357622970E-14	119	**24271975885%**

Pour un x inférieur à -20, on voit que les erreur sont trop importantes pour être négligées

Question 3

http://www-fourier.ujf-grenoble.fr/~parisse/mat249/mat249/node15.html

L'exponentielle a un rayon de convergence infini, ce qui signifie que si la série est réelle, il y a convergence sur l'intervalle ouvert] $-\infty;\infty$ [.La méthode utilisée provient des développements en série de Taylor qui permet d'avoir une série entière. Il faudra arrêter les itérations de manière à ce qu'on le stocke dans un double. On va donc s'arrêter lorsque le terme suivant sera plus petit que notre erreur relative. Plus x est grand, plus on devra itérer, ce qui est (trop) contraignant comme nous. On va donc utiliser une approximation de ln(2) et effectuer la division euclidienne de *x* par ln(2) avec un reste symétrique.

On calcule :

$$e^r = \sum_{i=0}^{\infty} \frac{r^i}{i!}$$

Et on trouve l'exponentielle de x : $e^x = 2^a e^r$

- Pk ca part en cachuete ?
- Comment on s'en protège dans le 2nd algo ?
- http://www.cdta.dz/sitedasm/telechargement/M1_2_Anane330.pdf

C. L'exponentielle
Le calcul de la fonction **exp(x)** d'un nombre en virgule flottante se fait par une première limitation du domaine de définition de la fonction par le fait que le résultat de calcul de cette fonction ne peut pas être supérieur au plus grand nombre représentable en double précision ni inférieur au plus petit nombre représentable en double précision de la norme IEEE 754. Les arguments qui donnent ces limites sont :
A = log (2^{-1023}) = -709.0895657
B = log (2^{1023}) = 709.0895657
L'exponentiel d'un nombre plus grand que B est un

dépassement de capacité alors que celui d'un nombre plus petit que A est arrondi vers zéro.
Pour calculer l'exponentielle d'un nombre x un flottant en double précision. On utilise l'identité suivante : $e^x = 2^{x/\ln 2}$.
La première étape est donc de calculer l'argument réduit : $x^* = x - k \cdot \ln 2$, où x^* [-ln2/2, ln2/2] et k un entier qui est égal à:
k= ϒx/ln2ƒ ou ϒ ƒ indique la partie entière la plus proche.
Le calcul de l'exponentiel devient alors :

$$e^x = e^{x^*+k \cdot \ln(2)} = e^{x^*} \cdot e^{k \cdot \ln(2)} = e^{x^*} \cdot 2^k.$$

Question 4

Après avoir exécuté le programme dans Matlab on observe que pour l'une des courbes, l'approximation est très rapide et exacte (algo optimisé) tandis que l'autre est plus longue et conduit à un résultat erroné.

Question 7

Après avoir exécuté le programme, pour résoudre le système par simple inversion on récupère la valeur de x. En comparant avec les résultats obtenus en Q5, il en ressort que les x sont identiques.