DIRO
IFT 2425

EXAMEN INTRA

Max Mignotte

DIRO, Département d'Informatique et de Recherche Opérationnelle, local 2377
Http : //www.iro.umontreal.ca/∼mignotte/ift2425/
E-mail : mignotte@iro.umontreal.ca

Date : *18/02/2016*

I	Amplification d'Erreur (25 pts)
II	Erreur en Arithmétique Flottante (52 pts)
III	Méthode du Point Fixe/Newton/Bissection (45 pts)
Total	122 points.

Tous documents personnels, calculatrices et calculateurs autorisés

I. Amplification d'Erreur (25 pts)

L'équation des gaz parfaits permet d'écrire :

$$P \cdot V = n \cdot R \cdot T$$

où P est la pression (en pascals), V le volume (en m^3), n le nombre de moles, R est une constante $R \approx 8.31$ (en $J \cdot \mathrm{mol}^{-1} \cdot K^{-1}$) et T, la température (en Kelvin).
Supposons que V et R sont exactes et que la pression P et la température T sont connues à $\pm\epsilon_P$ et $\pm\epsilon_T$ près (avec ϵ_P petit devant P et ϵ_T petit devant T).

1. Calculer de deux façons différentes (en justifiant pourquoi ces deux méthodes s'appliquent ici) une borne supérieure de l'erreur relative et absolue que l'on fait sur une estimation de n.
   <20 pts>
   Dans quels types de conditions (température, pression ?), les erreurs numériques s'amplifieront ou s'atténueront ?
   <5 pts>

Réponse

$\boxed{\mathbf{1}}$

Ici , on note ϵ_P ou ϵ_T ce qu'on a l'habitude de noter en cours Δ_P ou Δ_T et $n = PV/RT$.
La première méthode sera celle utilisant les series de Taylor qui s'applique bien car ϵ_P petit devant P et ϵ_T petit devant T. Elle permet de trouver rapidement :

$$\begin{aligned} \epsilon_n &= \frac{V\,P}{R\,T^2}\,\epsilon_P + \frac{V}{R\,T}\,\epsilon_T \\ \frac{\epsilon_n}{n} &= \frac{\epsilon_P}{P} + \frac{\epsilon_T}{T} \end{aligned}$$

<10 pts>

La seconde est la méthode dite *de la fourchette* qui est valide ici, car $n = PV/RT$ est une fonction simple qui noues permet de trouver facilement $n_{min} = (P_{min}V)/(RT_{max})$ et $n_{max} = (P_{max}V)/(RT_{min})$ avec n une fonction continue et monotone (en fait croissante) quand $P \in [P_{min}, \ldots, P_{max}]$ et quand $T \in [T_{max}, \ldots, T_{min}]$. Et, puisque ϵ_n est la demi- largeur de l'intervalle $[n_{min}, \ldots, n_{max}]$;

$$\epsilon_n = \frac{1}{2}\left\{\frac{(P+\epsilon_P)\,V}{R\,(T-\epsilon_T)} - \frac{(P-\epsilon_P)\,V}{R\,(T+\epsilon_T)}\right\} = \frac{V}{2\,R}\left\{\frac{(P+\epsilon_P)}{(T-\epsilon_T)} - \frac{(P-\epsilon_P)}{(T+\epsilon_T)}\right\}$$

$$\text{et :} \qquad \epsilon_n/n = \frac{RT}{PV}\cdot\frac{1}{2}\left\{\frac{(P+\epsilon_P)\,V}{R\,(T-\epsilon_T)} - \frac{(P-\epsilon_P)\,V}{R\,(T+\epsilon_T)}\right\} = \frac{T}{2\,P}\left\{\frac{(P+\epsilon_P)}{(T-\epsilon_T)} - \frac{(P-\epsilon_P)}{(T+\epsilon_T)}\right\}$$

<10 pts>

Nota - : Si on developpe la deuxième expression et que l'on élimine tous les termes quadratiques en ϵ (c-a-d $\epsilon_T^2 \approx 0$ et $\epsilon_P\epsilon_T \approx 0$, modélisant ainsi le fait que ϵ_P et ϵ_T sont petits devant P et T, on retrouve l'expression approximée par la formule de Taylor, c-a-d ; $\epsilon_n/n = \epsilon_P/P + \epsilon_T/T$ (et même remarque pour la première expression associée à l'erreur absolue).

$\boxed{\mathbf{2}}$

Si on reprend l'expression la plus simple, cad : $\frac{\epsilon_n}{n} = \frac{\epsilon_P}{P} + \frac{\epsilon_T}{T}$, on voit bien les erreurs vont d'autant plus s'amplifier que la pression sera inférieur à 1. Cela se produira aussi pour la température mais une température inférieure à un Kelvin est assez rare (on a dans le vide une température de 2,73 degrés Kelvin en moyenne, cad la célèbre température résiduelle ou rayonnement fossile du Big Bang ! approximée à 3K).
<5 pts>

II. Erreur/Amplification d'Erreur en Arithmétique Flottant (57 pts)

1. Expliquer pourquoi le calcul numérique de ces trois expressions :

   (a) $\frac{1}{1+x} - \frac{1}{(1+x)^2}$

   (b) $1 - \sin(x)$

   (c) $\frac{(x-1)}{(\exp(x-1)-1)}$

   peuvent conduire à des problèmes d'erreurs numériques (identifier aussi pour quelle valeur de x). Identifier et citer (tous) le(s) problème(s) numérique(s) associé(s) et proposer une expression mathématiquement équivalente (ou dans le pire des cas, lorsqu'on ne peut absolument pas, une expression approximativement équivalente) qui permettrait d'éviter ces problèmes numériques et/ou augmenter la précision des calculs de ces trois expressions.
   <15 pts>

2. Donner un exemple, en arithmétique flottante, ou :

$$a \times (b \times c) \neq (a \times b) \times c$$

   <5 pts>

3. Montrer par une petite démonstration (la plus convaincante possible) que la somme de deux nombres approximés quasi égaux ne pose pas de problème (ou en tout cas, pas autant de problème que la soustraction de deux nombres approximés quasi égaux !). On prendra, pour simplifier la démonstration, que l'imprécision de ces deux nombres (en erreur absolue ; mais aussi en erreur relative puisqu'ils sont quasi égaux !) que l'on ajoute sont approximativement les mêmes.
   <8 pts>

4. Soit l'expression :

$$f(x) = \frac{\exp(x) - 1}{x} - 1$$

   (a) Identifier pour quelle valeur de x le calcul numérique de l'expression ci-dessus ($f(x)$) posera problème et dire quelle est la limite de cette expression quand x tend vers la valeur indiquée.
   <3 pts>

   (b) Énumérer ensuite **tous** les problèmes numériques associés à cette expression.
   <3 pts>

   (c) Proposer ensuite une expression mathématiquement équivalente (ou dans le pire des cas, lorsqu'on ne peut absolument pas, une expression approximativement mathématiquement équivalente la plus précise possible) qui permettrait d'éviter les problèmes numériques cités précédemment et augmenter la précision du calcul de cette expression.
   <4 pts>

   (d) Indiquer la plus grande valeur de x, donnée par votre calcultrice (sous la forme $x = 10^{-?}$) qui indiquera une valeur erronée de $f(x)$ avec plus de 100% d'erreur par rapport à la valeur exacte de cette expression. Si vous répondez à cette question, indiquer aussi la marque de votre calculatrice.
   <4 pts>

5. En traitement d'images ou en machine learning (algorithmes d'apprentissages), il arrive quelquefois que l'on doive estimer, dans un algorithme, une somme $S()$, de la forme suivante :

$$S(x_1, x_2, \ldots, x_n) = \ln\left[\sum_{k=1}^{n} \exp(x_k)\right]$$

avec des valeurs de x_k évoluant, pour donner un exemple, autour de valeurs négatives très petites, du style : $x_k \in [-100, \ldots, -250]$.
Expliquer pourquoi le calcul numérique de cette expression peut conduire à des problèmes d'erreurs numériques. Proposer une expression mathématiquement équivalente et/ou une stratégie qui permettrait de minimiser le ou les problème(s) numérique(s) identifiés précédemment et/ou d'augmenter la précision du calcul de cette expression.
<10 pts>

Réponse

$\boxed{1}$

[-a-] Pour la première expression, on aura un problème d'annulation de cse (chiffres significatifs exactes) due à la soustraction de deux nombres approximés quasi égaux. Cette annulation de cse se passe lorsque $x \to 0^+$. Dans ce cas, il vaut mieux ré-écrire cette expression en réduisant le tout au même dénominateur :

$$\frac{1}{1+x} - \frac{1}{(1+x)^2} = \frac{x}{(1+x)^2}$$

[-b-] Pour la deuxième expression, on aura aussi un problème d'annulation de cse qui se passe lorsque $x \to \pi/2$. On peut ré-écrire cette expression de telle façon que plus aucune soustraction de deux nombres incertains (i.e., imprécis) quasi-égaux apparaissent de la façon suivante :

$$1 - \sin(x) = \frac{(1-\sin(x))(1+\sin(x))}{1+\sin(x)} = \frac{1-\sin^2(x)}{1+\sin(x)} = \frac{\cos^2(x)}{1+\sin(x)}$$

Nota - : La solution qui consiste à faire un développement limité au voisinage de $\pi/2$, comme approximation de l'expression est possible mais reste moins intéressante que l'expression mathématiquement équivalente donnée ici.

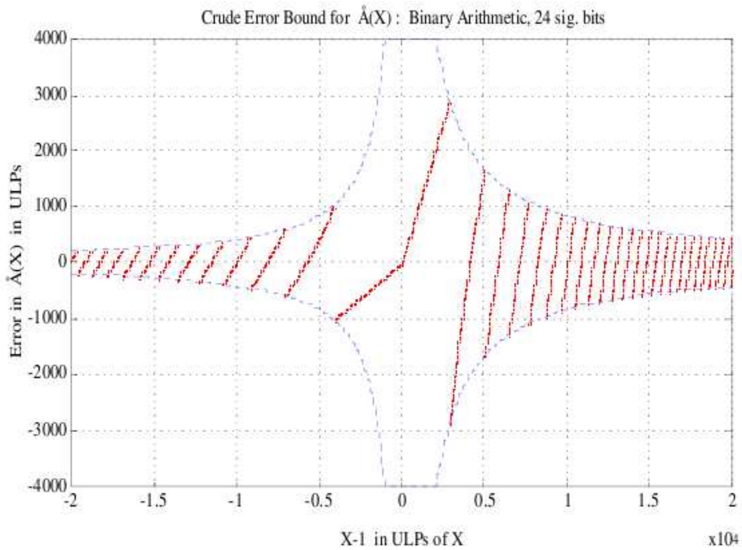
[-c-] Pour la troisième expression, il y a deux problèmes d'importance égale lorsque $x \to 1$. Le premier est aussi un problème d'annulation de cse au dénominateur. Le deuxième est une expression dangereuse indéterminée du type $0/0$ pouvant potentiellement engendrer en notation flottante un INF ou NaN. Le plus simple, pour cette expression est de considérer, pour des valeurs proche de 1, son developpement limité.

Dans un premier temps, on trouve immédiatement : $\exp(y) - 1 \approx y + (y^2/2) + (y^3/6) + O(y^4)$ et en faisant la division $y/(\exp(y)-1)$, on trouve : $1 - (y/2) + (y^2/12) + (y^3/24) + O(y^4)$ puis :

$$\approx 1 - [(x-1)/2] + [(x-1)^2/12] + [(x-1)^3/24] + O(x^4)$$

<15 pts>

Nota -1- : Le graphe montre l'erreur de cette expression ($A(x)$), pour sa forme originale, en ULP. ULP signifie *"Unit in the last place stored"*. Par exemple, pour un format flottant IEEE simple précision de 24 bits, un ULP $= 2^{-23}$. Pour un IEEE simple précision, l'erreur maximale se trouve donc, pour $x \to 1$ ou $(x-1) \to 0$ avec une erreur de ± 2900 UPL $\approx 2900 \times 2^{-23} \approx 3.46 \times 10^{-4}$.



Nota -2- : Aussi bizarre que cela puisse paraisse, le calcul de l'expression (c) est beaucoup plus stable numériquement si celui-ci est fait de la façon (équivalente mathématiquement), au voisinage 1 pour x, suivante : (Réference : W. Kahan (1998). "How JAVA's Floating-Point Hurts Everyone Everywhere.")

```
float CalcExprC( float x )
{
float y, z ;
y = x-1.0 ;
z = exp(y) ;
if (z != 1.0) z = y/(z - 1.0) ;
return (z) ;
}
```

2

Par exemple :

$$
\begin{aligned}
10^{50} \times \left(10^{50} \times 10^{-50}\right) &= 10^{50} \\
&\text{et} \\
\left(10^{50} \times 10^{50}\right) \times 10^{-50} &= \text{ERROR}
\end{aligned}
$$

<5 pts>

3

Puisque l'incertitude (ou erreur absolue) sur la somme de nombres incertains est la somme des incertitudes (ou erreur absolue) (cf. cours), donc pour deux nombres imprécis, U^* et V^*, quasi égaux, que l'on ajoute, on a donc la somme approximée S^* et l'incertitude suivante :

$$
\begin{aligned}
S^* &= (U^* + V^*) + \Delta S \\
S^* &= (U^* + V^*) + (\Delta U + \Delta V)
\end{aligned}
$$

Puisque $\Delta U \approx \Delta V$ et $U \approx V$ (deux nombres imprécis quasi égaux), cette somme aura une erreur relative comparable à l'erreur relative que l'on a initialement sur U ou sur V ($\approx \Delta U/U \approx \Delta V/V$). Le calcul de la somme de ces deux valeurs imprécises n'amplifiera donc par l'imprécision initiale, ni en erreur relative, ni en ereur absolu, que l'on a sur chacune des valeurs de départ.
<8 pts>

$\boxed{\textbf{4.(a)}}$

Il y a pour $x \to 0$, une succession de problèmes numériques. La limite quand $x \to 0$ est (donnée par la règle de l'hospital) $\lim_{x\to 0}(\exp(x)-1)/x = \lim_{x\to 0}\exp(x)/1 = 1$. Donc en théorie l'expression tend vers 0 quand $x \to 0$.

<3 pts>

$\boxed{\textbf{4.(b)}}$

Il y a pour $x \to 0$ 1-] la soustraction de deux nombres approximés quasi égaux pour $(\exp(x)-1)$, 2-] une forme indéterminée du type $0/0$ pour $(\exp(x)-1)/x$ et finalement 3-] une autre soustraction de deux nombres approximés quasi égaux entre -1 et $(\exp(x)-1)/x$.

<3 pts>

$\boxed{\textbf{4.(c)}}$

Dans ce cas, on a pas le choix de proposer pour obtenir une très bonne approximation de cette expression pour de faible valeur de x ($x \to 0$), un développement limité qui dans cet exemple se trouve très facilement :

$$f(x) = \sum_{n=1}^{\infty} \frac{x^n}{(n+1)!} \quad = \frac{x}{2!} + \frac{x^2}{3!} + \frac{x^3}{4!} + \frac{x^4}{5!} + \ldots$$

et on prendra le nombre de termes permettant d'arriver à la précision voulue.

<4 pts>

$\boxed{\textbf{4.(d)}}$

On a caclulé précedemment que, quand $x \to 0$; $f(x) = \frac{x}{2!} + \frac{x^2}{3!} + \ldots + \frac{x^n}{(n+1)!}$

Donc, on doit trouver une valeur de x, suffisement petite, qui donne une expression de $f(x)$ avec 100% d'erreur par rapport à $\approx x/2$.

Avec ma calculatrice CASIO *fx-7400G Plus* j'obtient :

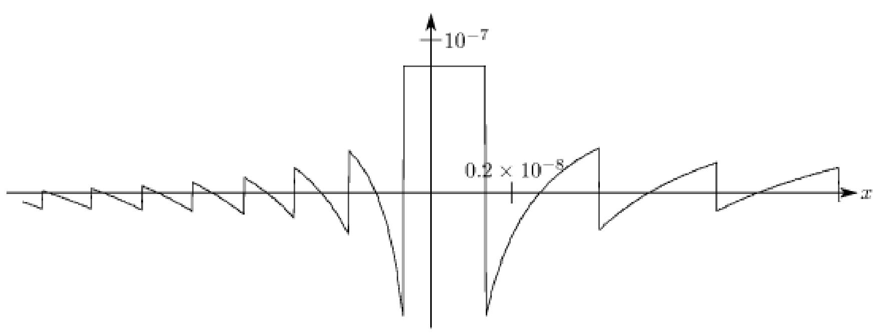
Pour $f(x = 10^{-6}) = 5 \times 10^{-7}$ $\triangleright$ Résultat avec 0% d'erreur

Pour $f(x = 10^{-7}) = 0$ $\triangleright$ Résultat avec 100% d'erreur

Pour moi, donc, tout va mal pour des valeurs plus petite ou égale à 10^{-7}, ce qui n'est pas vraiment si petit que ca.

<4 pts>

Nota :



Le graphe montre l'erreur de cette expression pour $x \in [-10^{-8} 10^{8}]$ pour un format flottant simple précision.

$\boxed{\textbf{5}}$

Sur ma calculatrice $\exp(-250) = 0$, ce qui est typiquement un problème d'UNDERFLOW.

Même si on peut imaginer une calculatrice en triple précision qui réussisse à assez bien estimer $\exp(-250)$, il reste encore la possibilité si n est très grand d'avoir une somme entre deux nombres de grandeur très différente (la somme partielle peut très vite être assez grande et on peut se retrouver avec le problème de l'addition de deux nombres d'ordre de grandeur différente).

Pour éviter le problème -1- (et seulement celui la, dans le cas d'une somme avec pas trop de termes), une solution consiste a prendre la maximum de l'ensemble $[-100, \dots, -250]$ et de soustraire ce maximum au deux bornes de cet ensemble pour obtenir $[0, \dots, -150]$ et de ré-écrire cette expression pour quelle soit mathématiquement équivalente. Ce cas ne pose plus de problème car $\exp(0)$ et $\exp(-150)$ ne sont plus en UNDERFLOW. L'expression mathématiquement équivalente est donc :

$$\begin{aligned}
S(x_1, x_2, \dots, x_n) &= \ln\left[\sum_{k=1}^{n} \exp(x_k - \max_{x_k} + \max_{x_k})\right] \\
&= \ln\left[\sum_{k=1}^{n} \exp(\max_{x_k}) \cdot \exp(x_k - \max_{x_k})\right] = \ln\left[\exp(\max_{x_k}) \sum_{k=1}^{n} \exp(x_k - \max_{x_k})\right] \\
&= \max_{x_k} + \ln\left[\sum_{k=1}^{n} \exp(x_k - \max_{x_k})\right]
\end{aligned}$$

Si le second problème subsiste, alors il serait bien de faire l'addition dans le sens des éléments croissant de la somme.

<10 pts>

III. Méthode du Point Fixe/Bissection/Newton (45 pts)

On se propose de trouver numériquement dans R une valeur approchée de la racine de la fonction

$$f(x) = \cos(x) - x \tag{1}$$

1. Méthode du Point Fixe
   (a) Montrer qu'il existe une racine unique r pour cette Eq. (1) dans l'intervalle $J = [0.0, 1.0]$. Montrer que l' intervalle J est un intervalle sur lequel la convergence vers une solution unique par la méthode du point fixe est assurée.
   <6 pts>
   (b) Utiliser le résultat de la question précédente pour calculer les 6 premières estimées $r_1, \dots, r_6$, en partant de $r_0 = 0.5$.
   <6 pts>
   (c) Pour obtenir à l'avance le nombre d'itérations nécessaires de la méthode du point fixe pour arriver à la racine souhaitée avec la précision voulue, on peut essayer de trouver (en utilisant le théorème des accroissements finis ou de la valeur moyenne) une majoration du type $|r_n - r| \leq K$, ou r_n désigne la valeur approchée, à la ni-ème itération, de cette racine. Trouver cette majoration et en déduire le nombre d'itérations nécessaire pour obtenir, par cette méthode, une valeur approchée à 10^{-2} près de cette racine.
   <5 pts>
   (d) Une autre possibilité consiste à arrêter la procédure itérative du point fixe lorsque r_{n+1} et r_n sont suffisamment proches, i.e., $|r_{n+1} - r_n|$ suffisamment petits. Dans ce cas, trouver une relation du type

$$|r_n - r| \leq K\,|r_{n+1} - r_n|\,,$$

avec K un majorant que l'on précisera et en déduire le nombre d'itérations nécessaire pour obtenir une valeur approchée de r à 10^{-2} près.
<8 pts>

2. Méthode de la Bissection
   (a) Si on avait utilisé la méthode de la bissection, avec combien d'itérations serait-on arrivé à une valeur approchée de la racine à 10^{-2} près ? (Nota : on vous demande d'obtenir une estimation de ce nombre en utilisant les propriété de la méthode de la bissection mais sans calculer les différentes estimation $r_0, r_1, \ldots,$ donné par cette méthode)
   <5 pts>
3. Méthode de Newton
   (a) Donner la relation itérative $r_{n+1} = g_{\text{newt.}}(r_n)$ intervenant dans la méthode itérative de Newton pour la résolution itérative de la racine r de l'équation $f(x) = 0$.
   <5 pts>
   (b) Montrer qu'avec $r_0 = 0.5$, la relation itérative de Newton converge.
   <5 pts>
   (c) Calculer les 4 premières estimées $r_1, \ldots, r_4$, en partant de $r_0 = 0.5$.
   <5 pts>

Réponse

1.(a)

L'étude des variations de la fonction f sur $J = [0.0, 1.0]$ montre que la fonction est continue et décroissante sur J, donc monotone car $f'(x) = -\sin(x) - 1 < 0$ sur J.
De plus, on a $f(0) = 1$ et $f(1.0) \approx -0.45$ donc $f(0)f(1) < 0$ et, puisque la fonction f est continue et décroissante sur J, il existe donc une racine r unique dans cet intervalle.
<3 pts>

De plus $f(x) = 0$ est équivalent à l'équation $x = \cos(x) = g(x)$ avec :

$$|g'(x)| = |-\sin(x)| < 1 \quad \forall x \in J = [0.0, 1.0]$$

La fonction $g(x)$ est donc contractante sur J et la convergence est assurée.
<3 pts>

1.(b)

En partant de $r_0 = 0.5$, on a, $r_n = g_1(r_{n-1})$ et,

$$\begin{aligned}
r_1 &= 0.8775825619 \\
r_2 &= 0.6390124942 \\
r_3 &= 0.8026851007 \\
r_4 &= 0.6947780268 \\
r_5 &= 0.7681958313 \\
r_6 &= 0.7191654459
\end{aligned}$$

qui va converger doucement vers la valeur $r = 0.7390851332$.
<6 pts>

1.(c)

En utilisant le théorème de la valeur moyenne, on obtient, puisque $g(r) = r$ et $r_n = g(r_{n-1})$ avec r_n la valeur approchée de la racine à la n-ième itération

$$\begin{aligned} r_n - r &= \frac{g(r_{n-1}) - g(r)}{(r_{n-1} - r)} \times (r_{n-1} - r) \\ &= g'(\zeta) \times (r_{n-1} - r) \qquad \text{avec} \quad \zeta \in J \end{aligned}$$

En utilisant l'inégalité $|g'(\xi)| < |\sin 1| \approx 0.8415$, on obtient les inégalités suivantes

$$|r_n - r| \leq \ 0.8415\, |r_{n-1} - r| \ \leq \ 0.8415^2\, |r_{n-2} - r| \ \leq \ldots \ \leq \ 0.8415^n \ \underbrace{(r_0 - r)}_{<1}$$

On obtiendra donc $|r_n - r| < 10^{-2}$ dès que $(0.8415)^n < 10^{-2}$, i.e., dès que $n = 27$, ce qui montre que cela converge lentement.

<5 pts>

Nota : Dans notre cas $r_{27} = 0.739090064$ et plus d'itérations ne permet pas de changer le chiffre des dix millièmes de cette estimation qui est aussi significatif au sens du larousse ce qui montre que la borne précédente est plutôt pessimiste ...

1.(d)

En utilisant le théorème des accroissement finis, on a :

$$\begin{aligned} r_{n+1} - r_n &= (r_{n+1} - r) - (r_n - r) \\ &= (g(r_n) - g(r)) - (r_n - r) \\ &= \frac{(g(r_n) - g(r)) - (r_n - r)}{(r_n - r)} \times (r_n - r) \\ &= (g'(\zeta) - 1) \times (r_n - r) \quad \text{avec } \zeta \text{ compris entre } r_n \text{ et } r \end{aligned}$$

De ce fait, on a : $\quad |r_n - r| = \dfrac{1}{|1 - g'(\zeta)|} \times |r_{n+1} - r_n| \quad$ avec $\quad \forall x \in J \quad$ et $\quad |g'(x)| < 0.8415$

d'où $\quad \dfrac{1}{|1 - g'(\zeta)|} \leq 6.31 \quad$ et on peut écrire $\quad |r_n - r| \leq 6.31 \cdot |r_{n+1} - r_n|$

Si on veut $|r_n - r|$ inférieur à 10^{-2}, il suffit de choisir n tel que $|r_{n+1} - r_n| < \frac{10^{-2}}{6.31} \approx 0.00158 \approx 1.58 \times 10^{-3}$.

<8 pts>

Nota : Dans notre, on arrive à $r_{14} \approx 0.0.7382$, $r_{15} \approx 0.7396$ et $r_{16} \approx 0.0.7387$ la différence $|r_{14} - r_{13}| \approx 2.08 \times 10^{-3}$ n'atteint pas le critère souhaité. Par contre la différence $|r_{15} - r_{14}| \approx 1.4 \times 10^{-3} < 1.58 \times 10^{-3}$ et la méthode de cette question nous indique que l'on est arrivé avec $n = 15$, i.e., avec r_{15} à la précision voulue. La différence entre $r = 0.7390851332$ et $r_{15} = 0.7396499628$ est $|r_{15} - r| \approx 5.65 \times 10^{-4}$ donc bien inférieure à 10^{-2} près. De plus on peut constater que cette borne est un peu moins pessimiste que la précédente.

2.

La méthode de la bissection permet de construire à partir de l'intervalle $[0.0, 1.0]$ contenant r, un nouvel intervalle de longueur moitié contenant r. En appliquant n fois consécutives la méthode, on obtient un intervalle de longueur $1/2^n$ contenant r. A l'itération n, on a un majorant de l'erreur absolu donné par $\Delta r = \frac{|1|}{2^n}$. Or on veut $\Delta r < 10^{-2}$, donc $2^n > 10^2 = 100$, ce qui, dans notre cas est vrai des que $n = 7$. Ce nombre d'itération est à comparer avec le $n = 27$ ou $n = 15$ obtenu par les deux critères différents utilisés pour la méthode itérative du point fixe précédente.

<5 pts>

3.(a)

La fonction $f(x)$ est dérivable sur $\mathbb{R}$ et la méthode itérative de Newton permet d'écrire :

$$r_{n+1} = r_n - \frac{f(r_n)}{f'(r_n)} = r_n - \frac{\cos(r_n) - r_n}{-\sin(r_n) - 1} = r_n + \frac{\cos(r_n) - r_n}{\sin(r_n) + 1}$$

<5 pts>

3.(b)

Dans ce cas, le plus simple est de montrer que la fonction $f(x)$, sur l'intervalLe $J = [0, 1]$ (intervalle dans lequel on prendra $r_0 = 0.5$ et dans lequel une racine unique s'y trouve) ne présente pas d'*extrema*, i.e., de valeurs pour laquelle, $f'(x)$ s'annule.

Dans notre cas, $f(x) = \cos(x) - x = 0$ et $f'(x) = -\sin(x) - 1 < 0$, toujours négative sur J donc aucun problème !

<5 pts>

3.(c)

En partant de $r_0 = 0.5$, on a, $r_n = g_{\text{newt.}}(r_{n-1})$ et :

$$\begin{aligned} r_1 &= 0.7552224171 \\ r_2 &= 0.7391416661 \\ r_3 &= 0.7390851339 \\ r_4 &= 0.7390851332 \end{aligned}$$

<5 pts>

Nota -1- : Sur ma calculatrice, l'estimation se stabilise très rapidement, à partir de $r_4 = 0.7390851332$! ! (à comparer avec la méthode du point fixe précédente qui se stabilisait à partir de $r_{>59}$). Cela nous indique aussi que la fonction dont on cherche la racine est très très linéaire au voisinage de la racine.
On peut voir l'effet d'une racine simple sur la convergence de Newton qui, dans ce cas, est quadratique.

Nota -2- : On aurait pu le prévoir avec la méthode de l'exercice 1.(d) en trouvant une borne supérieure à $|\frac{1}{|1-g'(\zeta)|}|$ avec $g'(\zeta) = 1 + \frac{-2\sin(\zeta)\cos(\zeta)-2}{(\sin(\zeta)+1)^2}$, donc avec $|\frac{1}{|1-g'(\zeta)|}| = |\frac{(\sin(\zeta)+1)^2}{-2\sin(\zeta)\cos(\zeta)-2}| <\approx 1.67$ sur J. Pour obtenir $|r_n - r|$ inférieur à 10^{-2}, il suffit de choisir n tel que $|r_{n+1} - r_n| < \frac{10^{-2}}{1.67} \approx 0.00598$ qui est obtenu dès r_3 (car $|r_3 - r_2| \approx 5.7 \times 10^{-5} << 10^{-2}$).
Ce qui montre, une fois de plus, que la convergence est très rapide avec cette méthode (convergence quadratique car on cherche une racine simple dans ce cas).