Mathématiques 1

MP

2021

4 heures Calculatrice autorisée

La loi du demi-cercle

Notations

— Pour tous entiers naturels p et q tels que $p \leqslant q$ on note $[\![p, q]\!]$ l'ensemble $\{i \in \mathbb{N} \mid p \leqslant i \leqslant q\}$.

— Pour tout entier naturel n tel que $n \geqslant 1$, on note $\mathcal{M}_n(\mathbb{R})$ l'ensemble des matrices à n lignes et n colonnes à coefficients réels, $\mathcal{S}_n(\mathbb{R})$ l'ensemble des matrices symétriques et à coefficients réels de taille n et $\mathcal{O}_n(\mathbb{R})$ l'ensemble des matrices réelles orthogonales de taille n.

— Pour toute matrice symétrique réelle $M \in \mathcal{S}_n(\mathbb{R})$ on note $\lambda_1(M) \geqslant \lambda_2(M) \geqslant \cdots \geqslant \lambda_n(M)$ ses valeurs propres rangées dans l'ordre décroissant.

— On note $M^\top$ la transposée d'une matrice M.

— Pour tout $(i, j) \in [\![1, n]\!]^2$, on note E_{ij} la matrice de la base canonique de $\mathcal{M}_n(\mathbb{R})$ dont tous les coefficients sont nuls sauf celui situé sur la i-ème ligne et la j-ème colonne qui vaut 1.

— Pour toute matrice $M \in \mathcal{M}_n(\mathbb{R})$, on note $\|M\|_F = \sqrt{\operatorname{tr}(MM^\top)}$ sa norme euclidienne canonique.

— Dans tout le problème, on note $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé. Toutes les variables aléatoires considérées sont définies sur Ω.

— Étant donnée une variable aléatoire discrète X à valeurs réelles admettant une espérance, on note $\mathbb{E}(X)$ son espérance. Si X admet une variance, on note $\mathbb{V}(X)$ sa variance.

Problématique

En essayant d'expliquer la répartition des niveaux d'énergie des noyaux des atomes lourds, Eugene Wigner a été amené, dans les années 1950, à étudier le spectre de matrices symétriques réelles aléatoires de grande taille. La figure 1 montre un histogramme qui représente la répartition des valeurs propres d'une matrice symétrique réelle de taille $n = 2500$ constituée de variables aléatoires indépendantes identiquement distribuées, d'espérance nulle et de variance égale à 1.

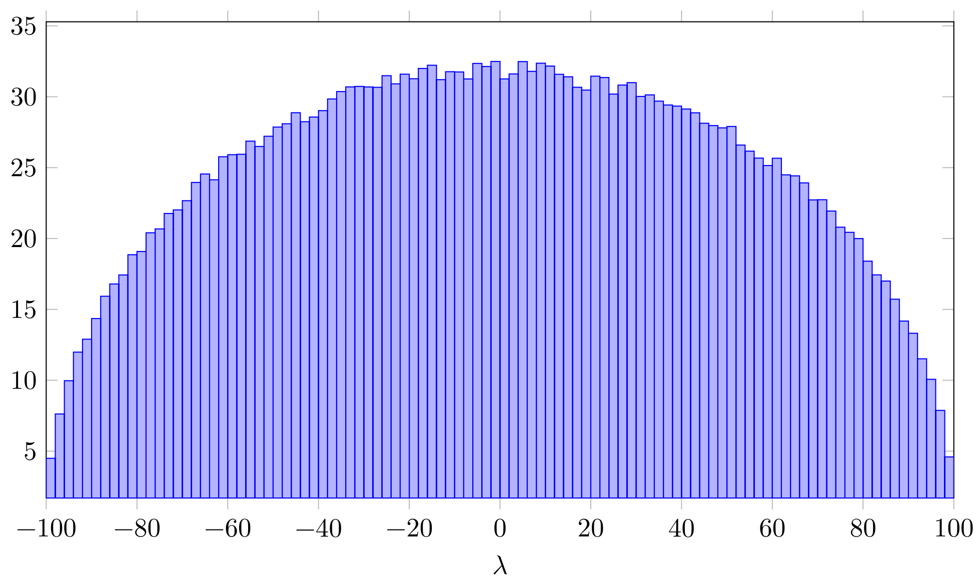

Figure 1

On voit que les valeurs propres se répartissent suivant un profil en demi-cercle de rayon $2\sqrt{n}$. En normalisant par un facteur $1/\sqrt{n}$ on obtient le profil d'un demi-cercle de rayon 2.

Plus précisément, on considère $(X_{ij})_{(i,j)\in(\mathbb{N}^\star)^2}$ une famille de variables aléatoires discrètes réelles telles que

— pour tout $(i,j)\in(\mathbb{N}^\star)^2$, $X_{ij}=X_{ji}$;

— les variables aléatoires X_{ij} sont de même loi, d'espérance nulle et de variance 1 ;

— pour $n\geqslant 1$, les variables aléatoires X_{ij}, pour $1\leqslant i\leqslant j\leqslant n$, sont mutuellement indépendantes.

Pour tout $n\in\mathbb{N}^\star$ et pour tout $\omega\in\Omega$, on note $M_n(\omega)$ la matrice $\big(X_{ij}(\omega)\big)_{1\leqslant i,j\leqslant n}\in\mathcal{M}_n(\mathbb{R})$.

Pour tout $\omega\in\Omega$, on note $\Lambda_{1,n}(\omega)\geqslant\cdots\geqslant\Lambda_{n,n}(\omega)$ les valeurs propres de $\dfrac{1}{\sqrt{n}}M_n(\omega)$.

On définit ainsi, pour $n\in\mathbb{N}^\star$ et $i\in[\![1,n]\!]$, des variables aléatoires réelles discrètes $\Lambda_{i,n}$ sur $(\Omega,\mathcal{A},\mathbb{P})$.

On se propose de montrer le résultat suivant :

Loi du demi-cercle

Pour toute fonction $f:\mathbb{R}\to\mathbb{R}$, continue et bornée,

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}f(\Lambda_{i,n})\right)\xrightarrow[]{n\to+\infty}\frac{1}{2\pi}\int_{-2}^{2}f(x)\sqrt{4-x^2}\,\mathrm{d}x.$$

Dans la partie I, on établit une inégalité portant sur les valeurs propres d'un couple de matrices symétriques. La partie II est consacrée à la résolution d'un problème de dénombrement qui est utilisée dans la partie III où la loi du demi-cercle est démontrée pour des variables aléatoires uniformément bornées. Dans la partie IV, on établit la loi du demi-cercle dans le cas général en utilisant les résultats des parties I et III.

I Inégalité de Hoffman-Wielandt

I.A – Soient A et B deux matrices de $\mathcal{S}_n(\mathbb{R})$.

Q 1. Montrer que, pour tout M dans $\mathcal{M}_n(\mathbb{R})$ et pour tous P et Q dans $\mathcal{O}_n(\mathbb{R})$, on a $\|PMQ\|_F=\|M\|_F$.

Q 2. On note $D_A=\operatorname{diag}(\lambda_1(A),\dots,\lambda_n(A))$ et $D_B=\operatorname{diag}(\lambda_1(B),\dots,\lambda_n(B))$. Montrer qu'il existe une matrice orthogonale $P=(p_{i,j})_{1\leqslant i,j\leqslant n}$ telle que $\|A-B\|_F^2=\|D_AP-PD_B\|_F^2$.

Q 3. Montrer que

$$\|A-B\|_F^2=\sum_{1\leqslant i,j\leqslant n}p_{i,j}^2\big(\lambda_i(A)-\lambda_j(B)\big)^2.$$

I.B – On note $\mathcal{B}_n(\mathbb{R})$ l'ensemble des matrices *bistochastiques* de $\mathcal{M}_n(\mathbb{R})$, c'est-à-dire l'ensemble des matrices $M=(m_{i,j})_{1\leqslant i,j\leqslant n}$ dont tous les coefficients sont positifs ou nuls et tels que $\sum\limits_{j=1}^{n}m_{i,j}=\sum\limits_{j=1}^{n}m_{j,i}=1$ pour tout $i\in[\![1,n]\!]$.

On note $f:\left|\begin{array}{ccc}\mathcal{M}_n(\mathbb{R}) & \to & \mathbb{R}\\ M & \mapsto & \sum\limits_{1\leqslant i,j\leqslant n}m_{i,j}\big(\lambda_i(A)-\lambda_j(B)\big)^2\end{array}\right..$

Q 4. Justifier que f admet un minimum sur $\mathcal{B}_n(\mathbb{R})$.

On se propose de montrer que ce minimum est atteint en la matrice identité.

Q 5. Soit $(i,j,k)\in[\![1,n]\!]^3$ tel que $j\geqslant i$ et $k\geqslant i$. Montrer que, pour $M\in\mathcal{M}_n(\mathbb{R})$ et pour $x\in\mathbb{R}^+$,

$$f(M+xE_{ii}+xE_{jk}-xE_{ik}-xE_{ji})-f(M)=2x(\lambda_i(A)-\lambda_j(A))(\lambda_k(B)-\lambda_i(B))\leqslant 0$$

Q 6. Soient $n\geqslant 2$ et $M=(m_{i,j})_{1\leqslant i,j\leqslant n}\in\mathcal{B}_n(\mathbb{R})$ une matrice différente de l'identité. On note i le plus petit entier appartenant à $[\![1,n]\!]$ tel que $m_{i,i}\neq 1$. Montrer qu'il existe une matrice $M'=(m'_{i,j})_{1\leqslant i,j\leqslant n}\in\mathcal{B}_n(\mathbb{R})$ telle que $f(M')\leqslant f(M)$ et $m'_{j,j}=1$ pour tout $j\in[\![1,i]\!]$.

Q 7. En déduire que

$$\min\{f(M) \mid M \in \mathcal{B}_n(\mathbb{R})\} = f(I_n).$$

I.C –

Q 8. En déduire que

$$\forall (A,B) \in \mathcal{S}_n(\mathbb{R})^2, \qquad \sum_{i=1}^{n} \big(\lambda_i(A) - \lambda_i(B)\big)^2 \leqslant \|A - B\|_F^2.$$

II Dénombrement des mots bien parenthésés

Dans cette partie, on s'intéresse à des chaines de caractères constituées uniquement des deux caractères parenthèse ouvrante et parenthèse fermante. On dit qu'un mot est *bien parenthésé* s'il commence par une parenthèse ouvrante et qu'à toute parenthèse ouvrante est associée une (unique) parenthèse fermante *qui lui est postérieure*. Par exemple le mot

( ) (( ))

est bien parenthésé. En revanche, le mot

( )) ( )

n'est pas bien parenthésé. Un mot bien parenthésé est ainsi forcément constitué d'un nombre pair de caractères, chaque parenthèse qui s'ouvre doit se refermer.

Pour tout entier $n \geqslant 1$, on note C_n le nombre de mots bien parenthésés de longueur $2n$. On pose par commodité $C_0 = 1$.

II.A –

Q 9. En énumérant les différents mots bien parenthésés de longueur 2, 4 et 6, montrer que $C_1 = 1$, $C_2 = 2$ et déterminer C_3.

Q 10. Montrer que, pour tout entier naturel n, $C_n \leqslant 2^{2n}$. Que peut-on en déduire pour le rayon de convergence de la série entière $\sum C_k x^k$?

Q 11. Montrer par un raisonnement combinatoire que, pour tout entier $k \geqslant 1$,

$$C_k = \sum_{i=0}^{k-1} C_i C_{k-i-1}.$$

On peut remarquer qu'un mot bien parenthésé est forcément de la forme $(m)m'$ avec m et m' deux mots bien parenthésés, éventuellement vides.

II.B – Pour tout $x \in \left]-\frac{1}{4}, \frac{1}{4}\right[$, on pose $F(x) = \displaystyle\sum_{k=0}^{+\infty} C_k x^k$.

Q 12. Montrer que, pour tout $x \in \left]-\frac{1}{4}, \frac{1}{4}\right[$, $F(x) = 1 + x\big(F(x)\big)^2$.

Q 13. Montrer que la fonction $f : \left| \begin{array}{ccc} \left]-\frac{1}{4}, \frac{1}{4}\right[& \to & \mathbb{R} \\ x & \mapsto & 2xF(x) - 1 \end{array} \right.$ ne s'annule pas.

Q 14. Déterminer, pour tout $x \in \left]-\frac{1}{4}, \frac{1}{4}\right[$, une expression de $F(x)$ en fonction de x.

Q 15. Déterminer le développement en série entière de la fonction $u \mapsto \sqrt{1-u}$. On écrira les coefficients sous la forme d'un quotient de factorielles et de puissances de 2.

Q 16. Montrer que, pour tout entier naturel n,

$$C_n = \frac{(2n)!}{(n+1)!\,n!}.$$

III Loi du demi-cercle, cas uniformément borné

On suppose, uniquement dans cette partie, que les variables aléatoires X_{ij} sont uniformément bornées :

$$\exists K \in \mathbb{R} \qquad \forall (i,j) \in (\mathbb{N}^\star)^2 \quad \left|X_{ij}\right| \leqslant K.$$

III.A – Pour tout entier naturel k on pose

$$m_k = \frac{1}{2\pi} \int_{-2}^{2} x^k \sqrt{4-x^2}\,\mathrm{d}x.$$

Q 17. Pour $k \in \mathbb{N}$, que vaut m_{2k+1} ?

Q 18. En utilisant le changement de variable $x = 2\sin t$, calculer m_0.

Q 19. À l'aide d'une intégration par parties, montrer que, pour tout entier naturel k,

$$m_{2k+2} = \frac{2(2k+1)}{k+2} m_{2k}.$$

Q 20. En déduire que

$$m_k = \begin{cases} C_{k/2} & \text{si } k \text{ est pair,} \\ 0 & \text{si } k \text{ est impair.} \end{cases}$$

III.B – Soit k un entier naturel. On se propose de montrer que

$$\lim_{n\to+\infty} \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} \Lambda_{i,n}^k\right) = m_k,$$

où $\Lambda_{i,n}^k = (\Lambda_{i,n})^k$ est la puissance k-ième de $\Lambda_{i,n}$.

Q 21. Justifier que la variable aléatoire $\displaystyle\sum_{i=1}^{n} \Lambda_{i,n}^k$ admet une espérance et que

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} \Lambda_{i,n}^k\right) = \frac{1}{n^{1+k/2}}\mathbb{E}\left(\mathrm{tr}(M_n^k)\right) = \frac{1}{n^{1+k/2}} \sum_{(i_1,\dots,i_k)\in[\![1,n]\!]^k} \mathbb{E}(X_{i_1i_2}X_{i_2i_3}\cdots X_{i_{k-1}i_k}X_{i_ki_1}).$$

On appelle *cycle* de longueur k à valeurs dans $[\![1,n]\!]$, tout $(k+1)$-uplet $\vec{\imath} = (i_1, i_2, \dots, i_k, i_1)$ d'éléments de $[\![1,n]\!]$. Les éléments $i_1, \dots, i_k$ sont appelés *sommets* du cycle $\vec{\imath}$. On dit aussi que le cycle *passe* par ces sommets. On note $|\vec{\imath}|$ le nombre de sommets distincts du cycle $\vec{\imath}$.

On appelle *arêtes* du cycle $(i_1, i_2, \dots, i_k, i_1)$ les couples non ordonnés (l'ordre des deux éléments de chaque couple n'est pas significatif) (i_1, i_2), (i_2, i_3), ..., (i_k, i_1).

Par exemple $\vec{\imath} = (1,3,5,3,2,2,1)$ est un cycle de longueur 6 dans $[\![1,5]\!]$. Les sommets de ce cycle sont les éléments 1, 2, 3 et 5, donc $|\vec{\imath}| = 4$. Les arêtes distinctes de ce cycle sont $(1,3)$, $(3,5)$, $(3,2)$, $(2,2)$ et $(2,1)$. Les arêtes $(3,5)$ et $(5,3)$ sont les mêmes.

Q 22. Montrer que le nombre de cycles de longueur k dans $[\![1,n]\!]$ passant par ℓ sommets distincts est inférieur ou égal à $n^\ell \ell^k$.

Q 23. En déduire que

$$\frac{1}{n^{1+k/2}} \sum_{\substack{\vec{\imath}\in[\![1,n]\!]^k \\ |\vec{\imath}|\leqslant(k+1)/2}} \left|\mathbb{E}(X_{i_1i_2}X_{i_2i_3}\cdots X_{i_{k-1}i_k}X_{i_ki_1})\right| \xrightarrow[]{n\to+\infty} 0.$$

On classe les cycles de longueur k en trois sous-ensembles :

— l'ensemble $\mathcal{A}_k$, constitué des cycles où au moins une arête n'apparait qu'une fois ;

— l'ensemble $\mathcal{B}_k$, constitué des cycles où toutes les arêtes apparaissent exactement deux fois ;

— l'ensemble $\mathcal{C}_k$, constitué des cycles où toutes les arêtes apparaissent au moins deux fois et il en existe au moins une qui apparait au moins trois fois.

Q 24. Montrer que, si le cycle $(i_1, i_2, ..., i_k, i_1)$ appartient à $\mathcal{A}_k$, alors

$$\mathbb{E}(X_{i_1 i_2} X_{i_2 i_3} \cdots X_{i_{k-1} i_k} X_{i_k i_1}) = 0.$$

Q 25. Montrer que, pour tout cycle $\vec{\imath}$ appartenant à $\mathcal{C}_k$, $|\vec{\imath}| \leqslant \dfrac{k+1}{2}$.

Q 26. Que peut-on dire de $\mathcal{B}_k$ si k est impair ? En déduire que $\lim\limits_{n \to +\infty} \mathbb{E}\left(\dfrac{1}{n} \sum\limits_{i=1}^{n} \Lambda_{i,n}^k\right) = 0$ dans ce cas.

On suppose dans la suite que k est pair et que $\vec{\imath} \in \mathcal{B}_k$ est un cycle passant par $\dfrac{k}{2} + 1$ sommets distincts. Autrement dit $|\vec{\imath}| = \dfrac{k}{2} + 1$.

On parcourt les arêtes de $\vec{\imath}$ dans l'ordre. À chaque arête de $\vec{\imath}$ on associe une parenthèse ouvrante si cette arête apparait pour la première fois et une parenthèse fermante si elle apparait pour la deuxième fois. Par exemple, au cycle $(1, 3, 2, 3, 1)$ correspond $(())$, au cycle $(1, 2, 1, 3, 1)$ correspond $()()$.

Q 27. Justifier que l'on obtient ainsi un mot bien parenthésé de longueur k.

Q 28. Dénombrer les cycles $\vec{\imath}$ qui correspondent à un mot bien parenthésé fixé.

Q 29. Déduire de ce qui précède que

$$\lim_{n \to +\infty} \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^{n} \Lambda_{i,n}^k\right) = C_{k/2}.$$

III.C –

Q 30. En déduire que, pour tout polynôme $P \in \mathbb{R}[X]$,

$$\lim_{n \to +\infty} \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^{n} P(\Lambda_{i,n})\right) = \frac{1}{2\pi} \int_{-2}^{2} P(x) \sqrt{4 - x^2} \, \mathrm{d}x.$$

III.D – Soit $A > 2$.

Q 31. Montrer que, pour tout $(p, q) \in \mathbb{N}^2$,

$$\mathbb{E}\left(\sum_{\substack{1 \leqslant i \leqslant n \\ |\Lambda_{i,n}| \geqslant A}} |\Lambda_{i,n}|^p\right) \leqslant \frac{1}{A^{p+2q}} \mathbb{E}\left(\sum_{i=1}^{n} |\Lambda_{i,n}|^{2(p+q)}\right).$$

Q 32. Montrer que

$$\lim_{n \to +\infty} \frac{1}{n} \mathbb{E}\left(\sum_{\substack{1 \leqslant i \leqslant n \\ |\Lambda_{i,n}| \geqslant A}} |\Lambda_{i,n}|^p\right) = 0.$$

Q 33. Soient f une fonction continue et bornée de $\mathbb{R}$ dans $\mathbb{R}$ et P un polynôme de degré p. Justifier qu'il existe une constante K telle que

$$\forall x \in \mathbb{R} \setminus]-A, A[, \qquad |f(x) - P(x)| \leqslant K |x|^p.$$

Q 34. En déduire que

$$\lim_{n \to +\infty} \frac{1}{n} \mathbb{E}\left(\sum_{\substack{1 \leqslant i \leqslant n \\ |\Lambda_{i,n}| \geqslant A}} |f - P| (\Lambda_{i,n})\right) = 0.$$

III.E –

Q 35. Soit f une fonction continue et bornée de $\mathbb{R}$ dans $\mathbb{R}$. Montrer que

$$\lim_{n\to+\infty} \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} f(\Lambda_{i,n})\right) = \frac{1}{2\pi}\int_{-2}^{2} f(x)\sqrt{4-x^2}\,\mathrm{d}x.$$

IV Loi du demi-cercle, cas général

On revient au cas général. On note $\mathbb{1}_A$ la variable aléatoire indicatrice d'un événement A.

Pour tout $(i,j)\in(\mathbb{N}^\star)^2$ et pour tout $C>0$, on pose

$$\sigma_{ij}(C) = \sqrt{\mathbb{V}(X_{ij}\mathbb{1}_{|X_{ij}|\leqslant C})}.$$

Si $\sigma_{ij}(C)\neq 0$, on pose

$$\widehat{X}_{ij}(C) = \frac{1}{\sigma_{ij}(C)}(X_{ij}\mathbb{1}_{|X_{ij}|\leqslant C} - \mathbb{E}(X_{ij}\mathbb{1}_{|X_{ij}|\leqslant C})).$$

IV.A –

Q 36. Soit X une variable aléatoire discrète d'espérance finie. Montrer que

$$\mathbb{E}(X\mathbb{1}_{|X|\leqslant C}) \quad \xrightarrow{C\to+\infty} \quad \mathbb{E}(X).$$

Q 37. En déduire que

$$\lim_{C\to+\infty}\sigma_{ij}(C) = 1.$$

Q 38. Justifier que, pour C assez grand, les variables $\widehat{X}_{ij}(C)$ sont bien définies et qu'elles sont alors bornées, centrées, de variance 1 et qu'elles sont mutuellement indépendantes pour $1\leqslant i\leqslant j$.

Q 39. Montrer que

$$X_{ij} - \widehat{X}_{ij}(C) = \left(1 - \frac{1}{\sigma_{ij}(C)}\right)X_{ij} + \frac{1}{\sigma_{ij}(C)}(X_{ij}\mathbb{1}_{|X_{ij}|>C} - \mathbb{E}(X_{ij}\mathbb{1}_{|X_{ij}|>C})).$$

Q 40. Montrer que

$$\lim_{C\to+\infty}\mathbb{E}\Big(\big(X_{ij} - \widehat{X}_{ij}(C)\big)^2\Big) = 0.$$

IV.B – Pour tout entier n tel que $n\geqslant 1$, on note $\widehat{M}_n(C) = (\widehat{X}_{ij}(C))_{1\leqslant i,j\leqslant n}$. Pour tout $\omega\in\Omega$, on note $\widehat{\Lambda}_{1,n}(\omega)\geqslant ...\geqslant\widehat{\Lambda}_{n,n}(\omega)$ les valeurs propres de $\dfrac{1}{\sqrt{n}}\widehat{M}_n(C)(\omega)$ rangées dans l'ordre décroissant.

On obtient ainsi des variables aléatoires réelles discrètes $\widehat{\Lambda}_{1,n},...,\widehat{\Lambda}_{n,n}$ sur $(\Omega,\mathcal{A},\mathbb{P})$.

Soit $f:\mathbb{R}\to\mathbb{R}$ une fonction K-lipschitzienne.

Q 41. Montrer que

$$\left|\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} f(\Lambda_{i,n})\right) - \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} f(\widehat{\Lambda}_{i,n})\right)\right| \leqslant \frac{K}{n}\mathbb{E}\left(\left\|M_n - \widehat{M}_n(C)\right\|_F\right).$$

Q 42. On suppose de plus f bornée. Montrer

$$\mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} f(\Lambda_{i,n})\right) \quad \xrightarrow{n\to+\infty} \quad \frac{1}{2\pi}\int_{-2}^{2} f(x)\sqrt{4-x^2}\,\mathrm{d}x.$$

IV.C –

Q 43. Montrer la loi du demi-cercle dans le cas général.

• • • FIN • • •