

928 : Problèmes NP-complets : exemples et réduction

L'objectif ne doit pas être de dresser un catalogue le plus exhaustif possible ; en revanche, pour chaque exemple, il est attendu que le candidat puisse au moins expliquer clairement le problème considéré, et indiquer de quel autre problème une réduction permet de prouver sa NP-complétude.

Les exemples de réduction seront autant que possible choisis dans des domaines variés : graphes, arithmétique, logique, etc. Un exemple de problème NP-complet dans sa généralité qui devient P si on contraint davantage les hypothèses pourra être présenté, ou encore un algorithme P approximant un problème NP-complet.

Si les dessins sont les bienvenus lors du développement, le jury attend une définition claire et concise de la fonction associant, à toute instance du premier problème, une instance du second ainsi que la preuve rigoureuse que cette fonction permet la réduction choisie.

Chapitre 5

Épreuves orales de modélisation

Lors de l'inscription au concours, quatre options sont proposées :

- A. Probabilités et Statistiques,
- B. Calcul scientifique,
- C. Algèbre et Calcul formel,
- D. Modélisation et Analyse de systèmes informatiques.

L'épreuve de modélisation comporte une période de préparation de quatre heures et une interrogation dont la durée sera d'une heure à compter de 2017.

Même s'il s'agit d'une épreuve plus appliquée ou moins académique que les deux autres épreuves orales, cela ne dispense en aucun cas les candidats de faire preuve de la rigueur mathématique requise : quand on utilise un théorème, il faut impérativement être capable d'en restituer un jeu d'hypothèses. Par jeu d'hypothèses correct, on entend que le théorème soit vrai et qu'il s'applique effectivement au contexte considéré. Le jury n'attend pas nécessairement un énoncé avec les hypothèses minimales.

Cette épreuve de modélisation doit permettre aux candidats de mettre en avant diverses qualités : les connaissances mathématiques (pour les options A, B, C) / informatiques et mathématiques (pour l'option D), la réflexion et la mise en perspective des connaissances, l'aptitude à les appliquer à des problèmes concrets de modélisation et à produire des illustrations informatiques pertinentes, les qualités pédagogiques de mise en forme d'un exposé construit et cohérent, la capacité à faire preuve d'initiatives pour s'exprimer et manifester des qualités pédagogiques et de synthèse. La capacité des candidats à répondre aux questions fait partie intégrante de l'évaluation de cette épreuve. Comme pour l'ensemble des oraux, le caractère vivant de l'exposé est un atout.

Le texte fourni est la base pour construire et exposer un traitement mathématique (et/ou informatique pour l'option D) d'un problème « concret » en s'appuyant sur des éléments, généralement partiels, disséminés dans le texte. La présentation doit s'appuyer sur un dosage cohérent et harmonieux entre introduction motivée de modèles, preuves mathématiques, illustrations informatiques, critiques éventuelles du texte, réponses aux questions, mise en lumière des connaissances. Les candidats doivent utiliser leurs connaissances mathématiques et informatiques pour justifier certains points mentionnés dans le texte, proposer un retour sur le modèle ainsi qu'une conclusion.

Texte L'épreuve de modélisation repose sur un texte d'environ 5 à 6 pages (hors l'exercice de programmation pour l'option D) que les candidats choisissent parmi les deux qui leur sont proposés.

En 2017, ces textes seront surmontés, pour les options A, B, C, du bandeau suivant :

Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. La présentation, bien que totalement libre, doit être organisée et le jury apprécie qu'un plan soit

annoncé en préliminaire. L'exposé doit être construit en évitant la paraphrase et en mettant en lumière les connaissances, à partir des éléments du texte. Il doit contenir des illustrations informatiques réalisées sur ordinateur, ou, à défaut, des propositions de telles illustrations. Des pistes de réflexion, indicatives et largement indépendantes les unes des autres, vous sont proposées en fin de texte.

Pour l'option D, le bandeau sera :

Il est rappelé que le jury n'exige pas une compréhension exhaustive du texte. La présentation, bien que totalement libre, doit être organisée et le jury apprécie qu'un plan soit annoncé en préliminaire. L'exposé doit être construit en évitant la paraphrase et en mettant en lumière les connaissances, à partir des éléments du texte. L'exercice de programmation doit être présenté lors de l'exposé. Des pistes de réflexion, indicatives et largement indépendantes les unes des autres, vous sont proposées en fin de texte.

Les textes se termineront par le texte suivant pour les options A, B, C :

Les pistes de réflexion suivantes ne sont qu'indicatives et il n'est pas obligatoire de les suivre. Vous pouvez choisir d'étudier, ou non, certains des points proposés, de façon plus ou moins approfondie, mais aussi toute autre question à votre initiative. Vos investigations comporteront une partie traitée sur ordinateur et, si possible, des représentations graphiques de vos résultats. À défaut, si vos illustrations informatiques n'ont pas abouti, il est conseillé d'expliquer ce que vous auriez souhaité mettre en œuvre.

Pour l'option D, la fin des textes comportera la mention suivante :

Les pistes de réflexion suivantes ne sont qu'indicatives et il n'est pas obligatoire de les suivre. Vous pouvez choisir d'étudier, ou non, certains des points proposés, de façon plus ou moins approfondie, mais aussi toute autre question à votre initiative. Vos investigations comporteront une partie traitée sur ordinateur, notamment à travers l'exercice de programmation.

Les textes sont souvent motivés par des problèmes concrets. Ils peuvent présenter des arguments rapides, voire heuristiques (signalés comme tels) mais ne contiennent pas d'assertion délibérément trompeuse et se concluent par une liste de suggestions. Même si la plupart des textes s'appuient sur des problématiques issues de sciences autres que les mathématiques, aucune connaissance dans ces domaines n'est exigée par le jury. Discuter la modélisation proposée par un texte consiste donc avant tout à dégager les comportements qualitatifs du modèle, la manière dont il dépend d'éventuels paramètres et s'il semble apte à rendre compte des phénomènes qu'il est censé représenter. Le jury s'attend à ce que les candidats ne se contentent pas d'un exposé qualitatif et démontrent certains résultats évoqués dans le texte. *A contrario*, les interprétations qualitatives du comportement des modèles sont parfois absentes des exposés. Pourtant, montrer que l'on comprend un modèle ne se réduit pas à prouver un théorème.

Chaque année, des textes sont rendus publics et sont disponibles sur le site de l'agrégation de mathématiques <http://agreg.org>. Ces textes sont représentatifs de l'épreuve et permettent aux candidats de se familiariser avec le format des textes, se faire une idée des attentes du jury, réfléchir à des illustrations numériques pertinentes dans le cadre du texte ou avoir des exemples d'exercices de programmation pour l'option D.

Préparation Durant les quatre heures de préparation, les candidats ont accès à la bibliothèque et disposent d'un ordinateur dont la configuration est décrite sur le site de l'agrégation de mathématiques, à l'adresse <http://agreg.org>. Les candidats retrouveront le même environnement sur l'ordinateur de la salle d'interrogation. Il n'est évidemment pas réaliste de découvrir les logiciels à disposition des candidats le jour de l'épreuve : la configuration informatique utilisée pour le concours et sa documentation

sont accessibles et téléchargeables sur le site <http://clefagreg.dnsalias.org/> et permettent de se familiariser avec l'environnement offert pour l'épreuve.

Il est recommandé aux candidats de consacrer une partie de leur temps de préparation à s'interroger sur le discours qu'ils tiendront, la stratégie d'exploitation du tableau et d'utilisation de l'outil informatique qui leur permettra de mettre au mieux en valeur leurs connaissances et leur compréhension du texte ou d'une partie de celui-ci. Proposer un exposé structuré et cohérent ne peut s'improviser au moment de l'oral et doit faire l'objet d'une réflexion préalable durant la préparation.

Oral Au début de l'interrogation, le jury commence par vérifier que les fichiers créés par les candidats lors de la préparation ont bien été transférés sur la machine locale (dont l'environnement est identifié à celui de la salle de préparation) et par rappeler les modalités de l'épreuve. Celle-ci est scindée en deux temps : un exposé de 40 minutes (35 minutes à partir de 2017), suivi de 25 minutes d'échanges avec le jury. Les candidats sont invités à commencer par donner la structure de leur présentation sous forme d'un plan et le déroulement de l'exposé doit être en cohérence avec cette structure. Grâce à ce plan, le jury pourra ainsi avoir une vision globale de l'exposé et aider, si besoin, les candidats à gérer leur temps. Comme les candidats se le voient rappeler en début d'épreuve, l'exposé doit être accessible à un public qui découvre les problématiques du texte et doit permettre d'en faire comprendre les enjeux à un public qui ne le connaîtrait pas. Le jury, tout en étant conscient des difficultés du concours, attend un minimum d'aisance au tableau, la manifestation d'une certaine volonté de capter l'attention de l'auditoire et un discours clair et précis.

Durant l'exposé, les candidats disposent de leurs notes, d'un tableau et d'un ordinateur. Ils peuvent alterner, quand bon leur semble, entre un exposé oral, quelques éléments rédigés au tableau de façon propre et lisible et la présentation de ce qui a été préparé à l'aide de l'outil informatique. Les candidats doivent gérer correctement le tableau et demander, si besoin, au jury les parties qu'ils peuvent effacer (le jury peut souhaiter conserver certains passages et y revenir lors des échanges avec les candidats). Le jury a sous les yeux un exemplaire des textes et les candidats peuvent y faire référence pour éviter de recopier les notations, les énoncés complets ou certaines formules. Même si les programmes ne fonctionnent pas comme ils l'auraient souhaité ou si les simulations numériques n'ont pas abouti, les candidats sont invités à expliquer ce qu'ils voulaient mettre en œuvre, illustrer ou programmer. Si les candidats n'ont pas accédé à l'ordinateur dix minutes avant la fin du temps qui leur est imparti, le jury les préviendra.

Il n'y a pas de « format type » pour cette épreuve. Des prestations très différentes, tant dans leur forme que dans leur contenu, sur un même texte, peuvent conduire également à des notes élevées. Comme dans tout oral, la construction de l'exposé doit être une préoccupation importante des candidats. Une réflexion s'impose afin de produire un tout cohérent et intelligible par un public qui, dans l'esprit de l'épreuve, découvre le texte à travers l'exposé des candidats. Une brève introduction à la problématique avant de s'engager dans une longue digression (succession de définitions, théorèmes, ...) sans lien avec le problème de départ ne peut donner satisfaction. Il en va de même d'un exposé se réduisant à la présentation de la problématique du texte et à des illustrations informatiques, ou à l'énumération linéaire des pistes de réflexion proposées par le texte, sans contribution des candidats. Un texte traité de façon partielle mais en profondeur peut donner une note élevée. À un survol superficiel de l'intégralité du texte sans apport mathématique / informatique ou critique scientifique, les candidats doivent préférer une discussion fouillée d'une portion du texte, bâtie sur des arguments mathématiques / informatiques solides, des simulations pertinentes accompagnées de commentaires de bon aloi. Néanmoins, dans l'ensemble, les candidats semblent avoir perçu la nécessité d'utiliser, au mieux, le temps qui leur est consacré. Le jury apprécie de voir de plus en plus de candidats qui se sont approprié le texte et en donnent une présentation pertinente et autre qu'une paraphrase linéaire.

Pour enrichir leur propos, les candidats sont invités à mobiliser leurs connaissances, sur des aspects variés du programme, en étayant les arguments seulement esquissés dans le texte par des énoncés précis.

Il est totalement illusoire de chercher à impressionner le jury par une logorrhée de mots savants : les textes proposés peuvent être discutés en exploitant un bagage technique qui n'utilise pas les éléments les plus sophistiqués du programme. En particulier, le jury ne manque pas de s'attarder sur toute notion amenée par les candidats durant leur présentation et il est toujours dommageable de s'aventurer sur des terrains méconnus. Bien plus qu'une démonstration de virtuosité technique, le jury attend que les candidats montrent leur maîtrise d'énoncés relativement simples « en situation » : c'est là que réside une des difficultés principales de l'épreuve. Nombre de candidats éprouvent des difficultés à formaliser précisément des notions de base du programme ou à mettre en œuvre certaines de leurs connaissances. *A contrario*, utiliser une portion excessive du temps de parole pour recycler un chapitre de cours ou un développement d'une leçon d'Analyse et Probabilités, d'Algèbre et Géométrie ou d'Informatique, en s'éloignant des enjeux du texte est considéré comme un hors sujet et est sévèrement sanctionné. La paraphrase pure et simple d'amples portions du texte ne constitue en aucun cas un exposé satisfaisant. Les textes fournissent souvent des esquisses de démonstrations qui sont précisément destinées à être complétées et commentées. Les candidats ne doivent pas se contenter de ces esquisses de démonstration. S'ils font mention de ces esquisses, le jury s'assurera que les candidats ont compris en profondeur. Le jury n'est pas dupe des candidats qui tentent de faire semblant de connaître une notion ou d'avoir compris un point du texte ou une démonstration. Il ne se laisse pas tromper non plus par les candidats qui font des indications du texte un argument d'autorité, tentative maladroite de masquer des insuffisances techniques. Un regard critique (« il faudrait prouver que... mais je n'ai pas réussi à le faire », « les hypothèses du théorème de XXX que je connais pour aborder des problèmes similaires ne sont pas satisfaites dans le cas présent » ...) est une attitude bien plus payante.

Échanges avec le jury Durant 25 minutes, le jury revient sur certains propos des candidats qui méritent précision. Il peut demander l'énoncé précis de théorèmes utilisés pour démontrer une assertion. Les échanges peuvent également porter sur la modélisation, les illustrations informatiques ou les exercices de programmation pour l'option D.

5.1 Recommandations du jury, communes aux options A, B, C

Le jury tient à souligner les attentes partagées entre les trois options A (Probabilités et Statistiques), B (Calcul scientifique) et C (Algèbre et Calcul formel).

Le jury regrette de ne pas voir davantage de dessins (soignés) ou schémas explicatifs qui peuvent rendre l'argumentation plus claire et convaincante. La capacité à revenir sur le problème de départ et à conclure quant à l'efficacité de l'approche mathématique proposée pour y répondre est une qualité très appréciée. La rigueur et la clarté de l'organisation, la gestion du temps, la pertinence des choix opérés parmi les différentes questions soulevées par le texte sont des éléments de l'évaluation. Les qualités de synthèse sont aussi appelées à s'exprimer : il vaut mieux indiquer les étapes cruciales d'un raisonnement que de se lancer dans un long calcul fastidieux qu'on aura du mal à mener à bien.

5.1.1 Illustration informatique

Le jury rappelle son fort attachement à cet aspect de l'épreuve, dont les ambitions sont clairement délimitées. Il ne s'agit en aucun cas, et pour aucune de ces trois options, d'un exercice de programmation. L'objectif est d'être capable d'utiliser l'outil informatique pour illustrer, de façon pertinente, le contenu du texte. La réalisation de cet objectif constitue une part incompressible de la note finale attribuée à l'épreuve. Une très bonne évaluation peut résulter d'une exploitation judicieuse de programmes simples, reposant largement sur les routines standards des logiciels fournis. La forme et la nature des illustrations proposées n'obéissent à aucun format préétabli. En revanche, elles doivent faire la preuve d'une véritable réflexion scientifique et être agrémentées de commentaires, sur les résultats et

les méthodes. Les illustrations informatiques devraient plus souvent permettre de faire un retour sur la problématique du texte et il convient de commenter les résultats obtenus dans le cadre du modèle proposé par le texte. Des illustrations peuvent mettre en valeur les limites du modèle étudié et le jury apprécie particulièrement que les candidats aient un regard critique sur le texte et ne prennent pas celui-ci pour vérité absolue. Le jury rappelle que même si les simulations ne sont pas abouties, il sait valoriser la démarche suivie lorsqu'elle est clairement argumentée et permettrait, avec des aménagements mineurs, de mettre en évidence des aspects intéressants du texte.

Si un bel effort est produit par la majorité des candidats, un nombre limité de candidats préfèrent ne pas toucher à l'outil informatique. Si les candidats pensent ainsi consacrer plus de temps à l'analyse du texte et à la préparation de leur exposé oral, cette stratégie n'est en aucun cas payante, bien au contraire.

Il est important de se rappeler qu'il s'agit d'une illustration. Présenter le code source de façon détaillée n'est pas indispensable, le jury sera plus sensible à une présentation de résultats d'exécution et/ou une représentation graphique clairement commentées oralement par le candidat. Dans le cas d'une représentation graphique, il est important de préciser, au moins oralement, ce qui est représenté en abscisses et ordonnées.

5.1.2 Liens avec les épreuves d'Analyse-Probabilités et Algèbre-Géométrie

Les candidats ne doivent pas hésiter à s'appuyer sur leurs connaissances spécifiques pour proposer des développements originaux dans les leçons d'Analyse et Probabilités ou d'Algèbre et Géométrie. Par exemple,

- A. les candidats de l'option A peuvent nourrir leurs leçons de thèmes et d'exemples probabilistes (la fonction caractéristique d'une variable aléatoire est une transformée de Fourier, le produit de convolution a un sens probabiliste, les espaces L^p d'une mesure de probabilité sont intéressants, l'étude de certaines séries aléatoires est possible, la fonction de répartition est une fonction monotone digne d'intérêt, le calcul intégral est en lien étroit avec les probabilités, ...).
- B. les candidats de l'option B peuvent centrer leur propos sur des problématiques motivées par les préoccupations du calcul scientifique (approximation de solutions d'équations linéaires ou non linéaires, d'équations différentielles, convergence d'algorithmes, analyse matricielle, ...).
- C. les candidats de l'option C peuvent s'appuyer sur leurs connaissances spécifiques en algèbre effective.

5.2 Option A : Probabilités et Statistiques

5.2.1 Commentaires généraux

Les deux aspects, probabiliste et statistique, forment un tout cohérent dans l'étude des phénomènes aléatoires et les textes proposés mêlent souvent ces deux points de vue. Cependant, il n'est pas impossible de se voir proposer le choix entre deux textes où l'aspect statistique est plus marqué : il est donc nécessaire que les candidats fassent un effort de formation statistique. La part importante de la statistique dans l'enseignement des mathématiques justifie cet investissement. Ainsi, le jury se satisfait de ne plus voir les candidats boudier les textes à coloration statistique, contrairement à ce qui a pu être constaté lors des sessions précédentes.

En outre, la difficulté des textes étant progressive, une honnête maîtrise des fondamentaux du programme permet, sans virtuosité technique, d'aborder l'épreuve favorablement. Il n'est pas rare qu'un texte démarre avec des résultats classiques (par exemple, la loi des grands nombres ou le théorème central limite). Être à l'aise avec ces notions permet de démarrer l'exposé en confiance.

Le jury met en garde contre les exposés pratiquant une paraphrase sans plus-value mathématique. Il est bien plus judicieux de traiter un passage éventuellement limité mais de manière détaillée et en fournissant tous les arguments pertinents.

5.2.2 Contenu théorique en probabilités-statistique

Au vu de son expérience, le jury estime judicieux, pour se préparer adéquatement à l'épreuve, de réfléchir aux techniques et notions suivantes.

- **Méthodes classiques en probabilités** (calcul de la loi de $f(X)$ à partir de la loi de X , utilisation de la fonction caractéristique pour la convergence en loi, probabilités conditionnelles et formules des probabilités totales ou de l'espérance totale, calcul de $\mathbf{E}(f(X, Y))$ en fonction de la loi du couple (X, Y)). À titre d'exemple, le calcul de $\mathbf{P}(X = Y)$ lorsque X, Y sont indépendantes et que X n'a pas d'atome pose souvent problème.
- **Différents modes de convergence** (presque-sûre, dans L^p , en probabilité, en loi). Leurs définitions, caractérisations et implications sont des questions fréquemment posées. Un peu de familiarité avec leur manipulation pourra être utile (par exemple, si $X_n \rightarrow X$ et $Y_n \rightarrow Y$, a-t-on $f(X_n) \rightarrow f(X)$ ou $(X_n, Y_n) \rightarrow (X, Y)$?).
- **La loi des grands nombres et le théorème central limite.** Ce sont des incontournables de l'épreuve et il faut en maîtriser les hypothèses et la conclusion. Il est utile de percevoir que le théorème central limite précise la vitesse de convergence dans la loi des grands nombres. D'un point de vue statistique, la loi des grands nombres met un estimateur en évidence et le théorème central limite permet de construire un intervalle de confiance (asymptotique). Rappelons que la loi faible des grands nombres a des hypothèses plus fortes et une conclusion plus faible que la loi forte. Elle n'a donc pas d'intérêt autre que pédagogique, en raison de la simplicité de sa preuve.
- **Intervalles de confiance.** Rappelons qu'un tel intervalle est aléatoire, fabriqué à partir des observations et que ses bornes ne doivent pas s'exprimer en fonction du paramètre à estimer. Notons aussi que le lemme de Slutsky s'avère utile pour remplacer des valeurs théoriques inconnues par des valeurs empiriques : il faut pouvoir expliquer ceci précisément.
- **Chaînes de Markov.** Énoncer les propriétés de Markov faible et forte pose souvent d'insurmontables difficultés. Il importe de reconnaître une construction récursive $X_{n+1} = f(X_n, \varepsilon_{n+1})$ d'une chaîne de Markov. La notion de mesure stationnaire doit pouvoir être interprétée de manière à la fois matricielle et probabiliste. La terminologie « état stable » parfois employée à la place de « mesure stationnaire » a donné lieu à des confusions (c'est une mesure et pas un état). Le sens des mots « irréductible », « récurrent », « aperiodique » mérite d'être précisément connu.
- **Espérance conditionnelle.** Outre la définition précise, il est indispensable d'avoir déjà fait quelques calculs concrets d'espérance conditionnelle (par exemple, $\mathbf{E}(f(X, Y)|Y)$ lorsque X, Y sont indépendantes ou lorsque le couple (X, Y) admet une densité, ou encore $\mathbf{E}(X|X + Y)$ avec X, Y indépendantes et de même loi).
- **Martingales.** Ce concept ne doit pas être confondu avec celui de chaîne de Markov. Il s'agit d'un outil parfois utile pour prouver des convergences presque-sûres ou L^2 .
- **Modèle linéaire.** Il s'agit essentiellement de comprendre que la minimisation d'une distance euclidienne est assurée par une projection orthogonale. Il est bon de savoir que la projection orthogonale sur F est simplement caractérisée par son action sur F et sur F^\perp .
- **Vecteurs gaussiens.** Par définition, une application affine conserve l'ensemble des vecteurs gaussiens, modifiant moyenne et covariance. Le théorème de Cochran, problématique pour de nombreux candidats, s'en déduit simplement car des projections orthogonales Π_V, Π_W sur des espaces orthogonaux vérifient $\Pi_V \Pi_V^T = \Pi_V$ et $\Pi_V \Pi_W = 0$.
- **Tests statistiques.** Le principe général est souvent mal connu. Rappelons qu'un test fixe deux hypothèses H_0, H_1 et un niveau α , qu'il majore la probabilité de rejeter H_0 sous H_0 par α et qu'il évalue aussi la puissance (probabilité d'accepter H_1 sous H_1). Les tests d'adéquation du

χ^2 et de Kolmogorov-Smirnov méritent d'être connus précisément.

À titre de satisfaction, le jury observe lors de cette session des progrès en ce qui concerne l'utilisation du lemme de Slutsky, la connaissance des lois usuelles et la construction d'intervalles de confiance.

5.2.3 Modélisation et mise en œuvre informatique

Le jury suggère aux candidats de réfléchir aux illustrations informatiques pertinentes en probabilités et statistiques. La présentation d'un test doit faire apparaître clairement la valeur de la statistique de test obtenue et le quantile de la loi pertinente. Si un intervalle de confiance I_n dépend de n observations, les candidats gagnent à présenter graphiquement l'évolution $n \mapsto I_n$. Illustrer le comportement asymptotique d'une suite de variables aléatoires demande d'avoir réfléchi aux différents modes de convergence. Une convergence presque-sûre se voit sur une trajectoire. Une convergence en loi demande de simuler un échantillon, de représenter l'histogramme ou la fonction de répartition empirique associés et de pouvoir expliquer pourquoi ces tracés empiriques sont proches de leurs pendants théoriques. De manière générale, lorsqu'un programme mêle une boucle et un aléa, il faut se demander si l'aléa doit être simulé avant la boucle ou bien à chaque étape de cette dernière.

S'il est parfaitement légitime d'utiliser les routines préprogrammées dans les logiciels disponibles, il pourra être pertinent d'avoir un peu réfléchi à leurs fonctionnements (calcul d'une fonction de répartition ou d'un quantile, simulation de lois usuelles, simulation de chaînes de Markov, calcul de mesures stationnaires).

Rappelons que de nombreux textes sont assortis d'un jeu de données numériques sur lequel les candidats sont invités à mettre en œuvre des illustrations. Le jury se réjouit de l'augmentation du nombre de candidats traitant effectivement ces données. Cela constitue une réelle plus-value pour l'exposé.

Enfin, s'il est bon que les choix de modélisation soient commentés, il ne s'agit pas de se livrer à une critique gratuite et systématique de toutes les hypothèses. Si une hypothèse semble restrictive, il sera judicieux d'expliquer en quoi elle simplifie les calculs. Si une généralisation est suggérée, il pourra être intéressant de signaler les complications techniques qu'elle entraînerait. Le jury valorise les efforts faits pour interpréter la signification pour le modèle des résultats mathématiques obtenus.

5.3 Option B : Calcul scientifique

5.3.1 Commentaires généraux

Si les modalités et les attentes de cette épreuve semblent connues par une majorité des candidats admissibles, beaucoup ne maîtrisent tout simplement pas les notions de base du programme général intervenant dans les textes. Afin d'aborder sereinement les textes proposés dans l'option B, le jury rappelle qu'un minimum d'aisance est requis avec les notions suivantes :

- Connaître le théorème de Cauchy-Lipschitz et être en mesure de l'appliquer pour des systèmes différentiels simples en citant correctement toutes les hypothèses nécessaires à son utilisation.
- Construire, analyser et mettre en œuvre la méthode d'Euler explicite.
- Connaître les principes des méthodes directes de résolution de systèmes linéaires (pivot de Gauss, LU), la notion de conditionnement, la recherche d'éléments propres de matrices, analyser et mettre en œuvre la méthode de la puissance.
- Analyser et mettre en œuvre la méthode de Newton (cas vectoriel).
- Construire la matrice correspondant à la discrétisation par différences finies de $(-\frac{d^2}{dx^2})$ et connaître ses propriétés.
- Être capable d'énoncer et appliquer le théorème des extrêma liés (typiquement pour des problèmes de minimisation de fonctionnelles convexes sur \mathbf{R}^N avec contraintes linéaires), analyser et mettre en œuvre un algorithme du gradient.

- Maîtriser les outils d'analyse de Fourier.
- Connaître les méthodes de quadratures classiques.

Le jury souligne que les textes exploitent ces notions dans leurs versions les plus élémentaires et ne requièrent aucun raffinement technique.

5.3.2 Recommandations spécifiques

Le jury émet les recommandations plus spécifiques suivantes :

- **Quadratures numériques.** L'erreur commise par la méthode des rectangles doit être connue et une preuve élémentaire est exigible des candidats.
- **Analyse des équations différentielles ordinaires et calcul différentiel.** La dérivation de fonctions de \mathbf{R}^m dans \mathbf{R}^n ne devrait pas poser de difficulté au niveau de l'agrégation. Les formules de développements de Taylor contiennent généralement un terme de reste, dont l'analyse est un point souvent crucial. Les candidats devraient faire preuve d'automatismes à la vue de la moindre équation différentielle ordinaire. Par exemple, un texte indiquant « la solution de l'équation différentielle [...] est définie pour tout temps et reste positive » doit amener à :
 1. Énoncer de façon précise le théorème de Cauchy-Lipschitz le plus adapté au problème (version linéaire si le problème est linéaire, version globale quand elle s'applique, ...).
 2. Expliquer comment il s'applique dans le contexte présent (détailler explicitement la fonction $(t, X) \in [0, T] \times \mathbf{R}^n \mapsto f(t, X) \in \mathbf{R}^n$ qui permet de mettre le problème sous la forme $X'(t) = f(t, X(t))$, distinguer la variable vectorielle X et la fonction $X : t \mapsto X(t)$ sont malheureusement des obstacles majeurs pour une grande proportion des candidats).
 3. En déduire la positivité de la solution et le caractère non borné du temps d'existence. Trop de candidats sont pris en défaut sur la notion de solution maximale.
- **Schémas numériques pour les équations différentielles.** Le jury considère la description des schémas d'Euler comme un élément central du programme de l'option. Les candidats doivent être capables de présenter clairement les principes guidant l'écriture de ces schémas et l'analyse de leurs propriétés de convergence, ainsi que les avantages et inconvénients des méthodes explicites et implicites. Trop rares sont les candidats capables de formaliser correctement une définition de la convergence d'un schéma numérique, qui est trop souvent confondue avec la consistance du schéma. La confusion récurrente entre l'approximation X_n et l'évaluation $X(t_n)$ de la solution exacte au temps t_n , l'incapacité à relier le nombre de points de discrétisation et le pas Δt , témoignent d'une compréhension déficiente du sujet. La mise en œuvre de ces méthodes peut être l'occasion de discuter des difficultés potentielles liées à la stabilité et aux contraintes portant sur le pas de temps.
- **Équations aux dérivées partielles.** Le jury précise que les textes qui évoquent des problèmes d'équations aux dérivées partielles peuvent être abordés avec des outils rudimentaires et ne nécessitent *a priori* aucune connaissance sur la théorie des distributions, bien que ces notions aient intégré le programme commun, ni ne réclament de dextérité particulière d'analyse fonctionnelle. Il est néanmoins important de savoir faire la différence entre un problème aux limites et un problème de Cauchy, et d'utiliser la discrétisation adaptée au type de problème considéré. Le jury a été quelque peu surpris que des candidats à cette épreuve découvrent la matrice associée à la discrétisation de $(-\frac{d^2}{dx^2})$ par différences finies le jour de l'oral.
- **Algèbre linéaire.** Des lacunes profondes et inquiétantes sont trop régulièrement relevées. Au grand étonnement du jury, de nombreux candidats ne font pas le lien entre matrices symétriques et formes quadratiques. Les raisonnements liés à la réduction des matrices sont trop souvent extrêmement laborieux et les méthodes pratiques de calcul (résolution de systèmes, déterminant, inverse d'une matrice, ...) méconnues. La notion de conditionnement est bien souvent trop imprécise et les liens entre rayon spectral et normes matricielles sont mal maîtrisés.

- **Optimisation.** Les théorèmes d'existence d'optima sont mal connus. On s'attend à ce que les candidats soient capable de préciser les hypothèses requises pour :
 1. obtenir l'existence d'un optimum,
 2. obtenir l'unicité,
 3. caractériser le minimum (sans contrainte ou avec contraintes d'égalité).
- **Analyse de Fourier.** Le jury s'étonne que de nombreux candidats aient du mal à préciser les hypothèses de régularité assurant la convergence de la série de Fourier. Le lien entre la régularité de la fonction et le comportement asymptotique de ses coefficients de Fourier doit être connu.

5.3.3 Modélisation et mise en œuvre informatique

Le jury de l'option B rappelle qu'une illustration réalisée avec les routines de base des logiciels fournis est tout à fait satisfaisante si elle est clairement présentée, motivée et discutée. Si **Scilab** ou **Python** sont certainement les logiciels les mieux adaptés, le jury relève qu'un certain nombre de candidats a pu fournir des résultats convaincants avec un logiciel comme **Octave**, **XCas** ou **Sage**. Le jury insiste sur le fait qu'il est important de commenter les résultats informatiques (courbes, solutions, erreurs, ...) obtenus et de les relier au problème de modélisation traité par le texte ou à des résultats mathématiques sous-jacents.

5.4 Option C : Algèbre et Calcul formel

5.4.1 Commentaires généraux

La ligne directrice de l'option C est dans un premier temps la recherche de l'*effectivité*, puis de l'*efficacité* (souvent en temps, parfois en espace) du calcul algébrique. Cela va des aspects les plus élémentaires (algèbre linéaire, groupes, calcul avec les entiers, les polynômes, les entiers modulo) aux aspects plus avancés (élimination, géométrie effective, codes correcteurs d'erreur). La quasi-totalité des textes proposés dans le cadre de l'option rentrent totalement dans une ou plusieurs de ces grandes problématiques.

Cette « grille de lecture » peut accompagner les candidats, les aider à construire leur exposé et, dans l'idéal, les amener à expliquer non seulement *ce que* le texte fait, mais aussi *pourquoi* il le fait. La capacité à percevoir ces problématiques fait la différence entre bonnes et excellentes prestations, et peut aussi expliquer les notes honorables de certains candidats dont le niveau mathématique était pourtant limité.

5.4.2 Calcul algébrique effectif

Lorsqu'un énoncé du texte affirme l'existence d'un objet (scalaire, vecteur, matrice, élément d'un groupe, entier, polynôme, ...), le jury apprécierait que les candidats mènent la réflexion suivante :

- peut-on calculer cet objet ?
- si oui, par quelle(s) méthode(s) ?
- ces méthodes sont-elles efficaces ? quel est leur coût ?

Les candidats ayant le réflexe de se saisir, spontanément, d'une question de complexité, sont encore très rares. Pourtant cette démarche est perçue très positivement par le jury. Pour prendre un exemple, quand le texte parle de cryptographie, comparer le coût du système proposé et le coût d'une attaque, même naïve, est une initiative intéressante et actuellement quasi-inexistante. Une telle étude est beaucoup plus à sa place qu'un exposé détaillé de RSA plaqué sur un texte qui n'en parle pas – la mention rapide de RSA dans un texte introduisant un système de chiffrement pour comparer des complexités restant bien sûr pertinente. Plus largement, une réflexion minimale sur les ordres de grandeur (est-ce

qu'un calcul faisable représente $10^1, 10^{10}, 10^{100}, 10^{1000}$ opérations élémentaires?) permettrait souvent de mieux situer les problèmes soulevés par un texte, ou de proposer des valeurs de paramètres réalistes quand ce sujet n'est pas évoqué par le texte.

5.4.3 Modélisation et mise en œuvre informatique

Le jury constate avec satisfaction une progression de la moyenne des candidats dans l'utilisation de l'outil informatique et regrette parfois que des candidats ayant préparé un travail d'illustration raisonnable, ne le mettent pas plus en valeur durant l'exposé.

Rappelons cependant que les candidats doivent être capables de justifier la pertinence de leur programme ou de leurs calculs dans le cadre du texte. Reprendre un morceau de code d'un livre est une démarche tout-à-fait acceptable à condition que les candidats comprennent exactement ce que fait le code et que cela fasse sens dans le cadre du texte.

Pour finir, le jury est parfois surpris de voir des candidats développer de longs et fastidieux calculs au tableau alors que l'utilisation de l'outil informatique leur aurait permis de gagner en temps et en clarté.

5.4.4 Aspects mathématiques

Sur les aspects mathématiques et la maîtrise des éléments du programme, les principales observations du jury sont les suivantes :

- Par rapport aux années précédentes, le jury se réjouit d'observer une amélioration en ce qui concerne l'algèbre linéaire effective. Les applications de la méthode du pivot de Gauss sont un peu mieux maîtrisées par les candidats. De plus, la complexité de cet algorithme est maintenant fréquemment connue des candidats même si l'origine de ce $O(n^3)$ donne parfois encore lieu à des explications pour le moins obscures.
- Le jury constate également une progression dans les connaissances théoriques des candidats sur les corps finis. Malgré cela, le calcul effectif dans ces corps n'est pas toujours maîtrisé.
- *A contrario*, la connaissance du résultant semble avoir régressé par rapport aux sessions précédentes. Si les candidats interrogés sur le sujet sont en général capables d'en donner une définition, ses propriétés élémentaires et surtout son utilisation pour éliminer des variables dans un système d'équations polynomiales semble très floue pour de nombreux candidats. Beaucoup de candidats le voient comme un critère d'existence d'un facteur commun et ne pensent plus (surtout dans le cas des polynômes à une variable) au PGCD qui est un objet bien plus simple à appréhender.
- Le jury observe que de nombreux candidats n'ont aucune connaissance sur les codes correcteurs d'erreurs. Un bon nombre d'entre eux font d'ailleurs la confusion entre codes correcteurs d'erreurs et cryptographie. Rappelons que les codes correcteurs sont une partie limitée du programme, et que très peu de connaissances sont exigibles (et exigées). Toutefois, il est nécessaire de s'y être confronté pour se familiariser avec les problématiques sous-jacentes, à savoir, typiquement qu'un bon code correcteur se décrit de façon compacte (et est donc en général linéaire), a une grande dimension et grande distance minimale (par rapport à sa longueur) et, aussi et surtout un algorithme de décodage efficace – rappelons que ce second point n'est pas vrai d'un code linéaire « quelconque ». Il faut avoir déjà un peu étudié le sujet pour comprendre les questions soulevées par presque tout texte sur les codes. Signalons enfin que la méconnaissance des corps finis est souvent rédhibitoire pour ce sujet.
- Si l'algorithme d'Euclide est bien connu des candidats, la plupart d'entre eux ne savent obtenir des relations de Bézout qu'en effectuant l'algorithme d'Euclide classique puis en procédant à une laborieuse « remontée ». Rappelons que l'algorithme d'Euclide **étendu** est explicitement au programme de l'option.

- Les attentes du jury en termes de complexité sont limitées mais il est attendu des candidats qu'ils sachent estimer le coût de certaines procédures classiques au programme : opérations sur les entiers et dans $\mathbf{Z}/n\mathbf{Z}$, sur les polynômes (multiplication, division euclidienne, évaluation, interpolation), pivot de Gauss, algorithme d'Euclide, ...

5.5 Option D : Modélisation et Analyse de Systèmes Informatiques

5.5.1 Commentaires généraux

Le jury apprécie le travail accompli pour la préparation de cette épreuve par les meilleurs candidats. Les candidats sont interrogés dans le même esprit que dans les autres options et les critères d'évaluation sont largement semblables, sauf en ce qui concerne l'exercice de programmation.

Dans la suite, nous procéderons en deux temps : nous évoquerons d'abord la vision que le jury a du déroulement de l'épreuve, avant de revenir sur les travers les plus fréquents. Chacune de ces deux parties évoquera les différents temps de l'épreuve : exposé, exercice de programmation, questions.

5.5.2 Attentes du jury

Les textes présentent généralement une problématique concrète, informatique ou de la vie de tous les jours, avant d'en proposer une formalisation plus ou moins complète et une analyse informatique plus ou moins détaillée. Ils sont souvent plutôt de nature descriptive et volontairement allusifs.

Exposé du candidat

Motivation. C'est aux candidats d'introduire le sujet du texte et de motiver la présentation qui va suivre – et cela n'arrive, hélas, que très rarement. Cette motivation est le plus souvent l'évocation de situations concrètes dans lesquelles on a besoin d'outils informatiques spécifiques. Ces situations peuvent être proposées par le texte lui-même, mais elles peuvent aussi être tirées de l'expérience personnelle des candidats. Toute contribution personnelle à ce niveau est toujours très appréciée !

Présentation du texte. Il est attendu des candidats une restitution argumentée d'une partie cohérente du texte, ainsi qu'un effort de formalisation sur les parties descriptives et allusives du texte. Il est bon d'essayer de donner une ou des preuves complètes d'énoncés du texte, ou de compléter les arguments parfois lapidaires fournis par ce dernier. Les énoncés considérés comme vraiment trop difficiles pour être prouvés dans le cadre d'une préparation en temps limité en partant des connaissances du programme sont systématiquement pointés comme devant être admis.

Il est attendu des candidats qu'ils soient fidèles à l'esprit du texte. Il ne s'agit pas de traiter l'intégralité des points du texte, mais que le traitement choisi soit cohérent : les candidats doivent par exemple pouvoir expliquer pourquoi ils ont choisi de développer certains points, et pas certains autres.

Exercice de programmation informatique

Nature de l'exercice. L'exercice de programmation proposé est en règle générale très simple et peut presque toujours être traité en une vingtaine de lignes. Il est décrit par une spécification claire (entrée / sortie / hypothèses éventuelles sur les données) qui doit impérativement être respectée¹.

La simplicité de l'exercice vient du fait que le jury souhaite avant tout tester une capacité (et non une virtuosité) à organiser un programme simple, clair, et pédagogique : le programme doit pouvoir être présenté et les choix (structures de données, style impératif vs fonctionnel, types), argumentés.

1. à une réserve près sur liste/tableau évoquée *infra*

Le jury n'accorde pas une importance excessive aux considérations d'élégance ou d'efficacité tant que ce qui est proposé reste dans les limites du raisonnable.

Exposé de l'exercice. L'exercice de programmation doit être présenté au jury, ce qui implique un commentaire pertinent en parallèle de la projection du code, qui aide à la compréhension toujours difficile dans un temps court d'un programme écrit par un autre.

La présentation au jury doit être faite, que le programme fonctionne — ce que l'on espère! — ou pas. C'est seulement dans un deuxième temps que les candidats lancent une exécution. Dans tous les cas, le jury évalue la qualité générale du code réalisé. Cette évaluation interactive permet aux candidats réactifs de repérer une erreur, voire de la corriger, de recompiler et de relancer l'exécution.

La dimension « présentation » de l'exercice de programmation est difficile et ne peut pas être découverte le jour de l'oral. S'interroger sur cette dimension lors de sa préparation – par exemple, en s'imaginant face à une classe – peut permettre de ne pas être pris au dépourvu.

Remarques sur le style de programmation. Si l'adoption d'un style de programmation plus élégant (par exemple, définition de types spécifiques ou de fonctions auxiliaires) peut augmenter un peu la vingtaine de lignes de code évoquée plus haut, il est à la fois peu raisonnable en matière de temps consacré à l'exercice, mais aussi de commodité pour le jury lors de l'évaluation, que l'ensemble du code ne puisse pas tenir sur un écran.

Les commentaires sont appréciés par le jury quand ils apportent un plus : spécification, pré ou post-conditions, invariants, complexité ; ils doivent rester en quantité raisonnable, ne pas empêcher la continuité et la lisibilité du code et ne pas être une simple paraphrase du code lui-même.

Enfin affirmons à nouveau que tous les exercices sont construits de manière à pouvoir être traités confortablement dans le cadre de tous les langages du programme. Par ailleurs, à compter de la session 2017, quand la spécification évoque un « tableau » (ou une « liste »), cela signifie toujours que les candidats peuvent choisir entre ces deux structures de données, mais doivent savoir argumenter leur choix – au minimum en expliquant que le langage ou le style de programmation choisi s'accommode mieux de l'un ou de l'autre.

On renvoie au rapport 2015 pour des commentaires plus détaillés, qui restent d'actualité, sur le style de programmation.

Tests. Les candidats doivent toujours proposer plusieurs jeux de test, dont si possible un qui ne soit pas totalement « jouet ». Les textes n'en proposant souvent au mieux qu'un, il faut donc prévoir un temps de réflexion sur ce point.

Questions du jury

Sur l'exercice de programmation. Le jury revient toujours, en général pour commencer la séance de questions, sur l'exercice de programmation, de manière plus ou moins approfondie. Il demande au moins aux candidats

- d'argumenter leurs choix s'ils ne l'ont pas fait au préalable,
- d'indiquer les bibliothèques et les fonctions avancées utilisées et d'expliquer leur comportement (voire de demander comment les candidats auraient pu faire sans) – et éventuellement, leur complexité en temps et en espace (cette dernière étant souvent mal comprise / connue). Cela est particulièrement vrai des nombreuses constructions avancées de Python.
- de préciser les hypothèses implicites faites sur les données.
- éventuellement, d'expliquer les messages d'erreurs observés à la compilation / l'interprétation / l'exécution.

Sur l'exposé des candidats. Ensuite le jury revient sur la partie du texte présentée par les candidats. L'interrogation s'adapte toujours au niveau des candidats. Les questions du jury portent au moins autant sur la démarche globale de modélisation et la compréhension des différentes approches du texte que sur l'étude technique des diverses propositions.

Il pose des questions destinées à affiner sa perception de la compréhension du texte par les candidats, à comprendre leur capacité à formaliser une question ou expliciter une preuve s'ils ne l'ont pas montrée d'eux-même, ou encore à tester leur regard critique sur le texte.

Enfin, il est fréquent que des questions d'informatique fondamentale soient posées en rapport avec le texte, pour percevoir la capacité à faire le lien entre connaissances théoriques et informatique souvent plus concrète ; les questions de calculabilité et complexité (tant analyse de complexité d'un algorithme que preuve de NP-complétude d'un problème, avec l'aide du jury), ou encore les questions d'algorithme, sont en particulier naturelles et fréquentes.

5.5.3 Quelques travers fréquents

Exposé

Le jury renvoie largement au commentaire commun aux quatre options sur cette partie : la paraphrase du texte, même agréablement conduite, est à proscrire, le jury cherchant à mesurer principalement l'apport des candidats par rapport au document que ces derniers ont reçu.

Le principal défaut observé est l'absence de formalisation ou de preuve dans le texte – qui conduit souvent à la paraphrase évoquée ci-dessus.

S'agissant des démonstrations, si une part de raisonnement informel, heuristique ou argumentatif, peut être admise, il est *indispensable* que les candidats précisent qu'ils vont utiliser des arguments de cette nature (par opposition à une preuve) et qu'ils aient, avant cela, montré leur capacité à effectuer des preuves complètes, rigoureuses, et soignées sur des énoncés non triviaux. Plus largement, annoncer comme « évidente », une assertion ne dispense pas de savoir donner un argument à la demande du jury ; des arguments élémentaires d'injectivité ou de cardinalité sont ainsi fréquemment mal identifiés par les candidats.

Exercice de programmation

Commençons par un commentaire positif : l'exercice de programmation est bien accepté, et bien traité par la quasi-totalité des candidats. Une écrasante majorité des programmes fonctionne et fait, de manière plus ou moins efficace ou élégante, ce qui est demandé, mais le jury regrette quelques lectures trop rapides de la spécification conduisant à des hors-sujet.

Observons néanmoins qu'il reste une marge de progression significative sur la présentation orale de l'exercice de programmation.

Programme. Le jury a souvent le sentiment que cet exercice est un peu surinvesti par les très bons candidats au regard des attentes. Certains candidats proposent souvent beaucoup plus que ce qui est demandé, au détriment des aspects formalisation évoqués précédemment. Il est généralement peu intéressant, par exemple, de proposer des optimisations simples (par exemple, remarquer qu'un graphe est non orienté et ne parcourir qu'une fois les arêtes au lieu de deux) qui peuvent être brièvement décrites à l'oral.

Pour le reste, les compléments proposés devraient s'efforcer de constituer une illustration d'un ou de plusieurs points du texte à l'instar de ce qui se pratique dans les autres options ; illustration d'un algorithme proposé par le texte, étude d'un exemple, expérimentations, étude statistique, mesure de temps ou de complexité en moyenne, ... De telles illustrations sont toujours très appréciées par le jury. Dans ce cadre, l'ensemble des outils informatiques présents sur l'ordinateur peut être utilisé. On peut même imaginer, pour aller loin dans cette direction, une présentation très expérimentale du texte mettant en lumière les problèmes qu'il évoque et exposant, sur un ou plusieurs exemples, les solutions proposées, leurs forces et leurs faiblesses. C'est un parti-pris différent de celui de la formalisation mais qui pourrait également donner d'excellents exposés.