

Résoudre, modéliser, concevoir avec les grands modèles de langage : situations et formes d'engagement dans le travail mathématique

Philippe R. Richard

Université de Montréal, Québec ; philippe.r.richard@umontreal.ca

Résumé. Cet article examine les transformations du travail mathématique induites par les grands modèles de langage, dans la tension entre plausibilité discursive et nécessité logique. À partir de situations de résolution, de modélisation et de conception, il met en évidence les effets différenciés des IA génératives et symboliques sur le raisonnement, la validation et la structuration des savoirs. En s'appuyant sur la théorie des espaces de travail mathématique, il introduit la notion de contrepoint technologique pour caractériser les interactions humain-IA comme une mise en tension entre régimes hétérogènes de production. L'analyse montre que les délégations souhaitables ne concernent pas le contrôle épistémique du travail mathématique, qui demeure une responsabilité humaine et une condition de l'idoneité du processus. L'intelligence artificielle apparaît ainsi moins comme un partenaire de coopération que comme un instrument critique de reconfiguration du travail mathématique.

Mots-clés : Intelligence artificielle ; travail mathématique ; modèles de langage ; engagement épistémique ; idoneité ; contrepoint technologique

Abstract. Solving, modelling, and designing with large language models: situations and forms of engagement in mathematical work. This article examines the transformations of mathematical work induced by large language models within the tension between discursive plausibility and logical necessity. Drawing on situations of problem-solving, modelling, and design, it highlights the differentiated effects of generative and symbolic AI on reasoning, validation, and the structuring of knowledge. Building on the Theory of Mathematical Working Spaces, it introduces the notion of technological counterpoint to characterize human-AI interactions as a tension between heterogeneous regimes of production. The analysis shows that desirable forms of delegation do not concern the epistemic control of mathematical work, which remains a human responsibility and a condition for the idoneity of the process. Artificial intelligence thus appears less as a cooperative partner than as a critical instrument for the reconfiguration of mathematical work.

Keywords: Artificial intelligence; mathematical work; large language models; epistemic engagement; idoneity; technological counterpoint

Introduction

Depuis toujours, la classe de mathématiques est un lieu où coexistent et s'articulent de multiples systèmes sémiotiques. La langue naturelle y joue un rôle central : elle permet de formuler, commenter, expliciter, mais aussi de contrôler et de mettre à l'épreuve les relations entre différentes formes de représentation — figures, graphiques, tableaux, diagrammes ou expressions symboliques. Comme l'a montré Duval (1995), aucun système de représentation

ne peut, à lui seul, rendre compte de l'objet mathématique ; la compréhension suppose une coordination active entre registres, largement médiée par les ressources du langage.

Cette pluralité sémiotique n'est pas nouvelle. Elle a cependant été profondément reconfigurée par l'introduction d'artéfacts numériques dans les pratiques scolaires — calculatrices, logiciels de calcul formel, environnements de géométrie dynamique ou outils de programmation. Ces dispositifs ont déplacé le travail mathématique vers des formes d'interaction fondées sur la manipulation, la visualisation et l'exploration : variation de paramètres, itération de processus, simulation de situations. Ils ont ainsi modifié les régimes d'engagement du sujet, en brouillant les frontières entre exploration, conjecture et validation.

L'arrivée récente des systèmes d'intelligence artificielle introduit une rupture plus radicale. Contrairement aux outils précédents, qui s'inscrivaient dans un dialogue avec les langages formels des mathématiques, les grands modèles de langage produisent des énoncés en langue naturelle selon des principes statistiques, sans garantie intrinsèque de validité logique. Il en résulte une tension spécifique : un discours peut apparaître cohérent et convaincant sans satisfaire aux exigences de nécessité propres au raisonnement mathématique. Autrement dit, en mathématiques, il ne suffit pas d'« avoir l'air vrai » ; encore faut-il pouvoir en répondre du point de vue de la preuve et du contrôle, ce que ces systèmes ne garantissent pas.

Cette tension ne se limite pas à une opposition entre vérité et erreur. Elle affecte le travail mathématique lui-même : la manière de formuler un problème, d'explorer des pistes, de vérifier une solution ou d'interpréter un résultat. Dans l'interaction avec des systèmes comme ChatGPT, Gemini ou Claude, le sujet est confronté à des productions qui peuvent à la fois soutenir, déstabiliser ou masquer les processus de raisonnement. Les fonctions du langage — expliciter, guider, structurer — se trouvent alors partiellement simulées, sans que les conditions de leur validité soient toujours assurées.

C'est dans cet espace de tension que se situe le présent travail. Plutôt que de considérer l'intelligence artificielle comme un simple outil d'assistance, nous cherchons à observer comment elle reconfigure, en situation, le travail mathématique lui-même. Nous nous appuyons pour cela sur une série de situations instrumentées, dans lesquelles l'interaction avec l'IA permet d'observer des phénomènes de déplacement, d'hésitation, de contrôle ou de perte de repères. Le cadre théorique mobilisé pour analyser ces phénomènes — notamment en termes de travail mathématique, de contrepoint technologique et d'idonéité — est développé dans (Richard, 2026).

1. Intelligence artificielle : paradigmes et tensions du travail mathématique

L'expression « intelligence artificielle » désigne aujourd'hui un ensemble hétérogène de techniques numériques, dont les effets sur le travail mathématique varient considérablement selon les formes, les usages et les systèmes mobilisés. Pourtant, dans le discours public et médiatique, l'IA est souvent réduite aux seuls grands modèles de langage (GML) ou aux systèmes génératifs de type ChatGPT, occultant ainsi la diversité historique et conceptuelle des approches existantes.

Cette simplification pose un problème concret dès que l'on considère des situations de travail mathématique. Elle tend à homogénéiser des outils qui n'engagent pas les mêmes formes

d'action, de formulation ou de validation. Dans certains cas, des systèmes génératifs produisent des réponses discursivement cohérentes sans garantir leur validité mathématique ; dans d'autres, des outils issus de l'IA symbolique permettent un contrôle plus explicite des opérations, mais au prix d'une interaction plus contrainte. Ces écarts ne relèvent pas seulement de différences techniques : ils modifient en profondeur les conditions d'engagement du sujet dans l'activité mathématique.

Pour appréhender cette diversité, il importe de distinguer plusieurs familles d'approches selon leur principe de fonctionnement, tout en identifiant les tensions structurantes qui les traversent. L'enjeu n'est pas seulement de décrire ces approches, mais de comprendre comment elles reconfigurent le travail mathématique lui-même — notamment dans les relations entre exploration, formulation et validation. Cette section propose d'abord une typologie technique synthétique, puis une typologie fonctionnelle orientée vers les usages. Elle introduit également deux notions centrales pour la suite : les régimes d'engagement et le contrepoint technologique.

1.1 Paradigmes de l'IA et formes de production mathématique

Les principaux paradigmes de l'intelligence artificielle peuvent être regroupés en plusieurs grandes familles, fondées sur des modèles computationnels distincts. Les approches symboliques, historiquement premières, reposent sur la manipulation explicite de symboles et de règles logiques à l'aide de moteurs d'inférence déterministes. Elles permettent de produire des raisonnements formels, vérifiables, proches de ceux attendus dans une démonstration mathématique. À l'inverse, les approches subsymboliques, issues de l'apprentissage statistique sur de grandes masses de données, identifient des régularités sans les expliciter : elles produisent des résultats performants, mais souvent opaques du point de vue des processus qui les sous-tendent.

Entre ces deux pôles, d'autres approches — probabilistes, évolutionnistes ou incarnées — proposent des formes intermédiaires de traitement de l'information, fondées respectivement sur la gestion de l'incertitude, l'exploration de solutions par variation, ou l'interaction avec un environnement. Enfin, les développements récents en intelligence artificielle tendent vers des formes hybrides, dites neurosymboliques, qui cherchent à articuler la puissance adaptative des modèles statistiques et la rigueur des systèmes symboliques, notamment en couplant des modèles de langage à des moteurs de calcul formel. Toutefois, pour le travail mathématique, cette diversité technique se laisse interpréter à travers une tension plus fondamentale entre deux régimes de production : d'un côté, des systèmes fondés sur la nécessité logique, capables de garantir la validité interne des énoncés ; de l'autre, des systèmes fondés sur la plausibilité discursive, capables de produire des énoncés cohérents en langue naturelle sans en assurer la validité. Cette tension, analysée de manière plus systématique dans (Richard, 2026), constitue un point d'appui central pour comprendre les effets différenciés des systèmes d'IA sur les processus de raisonnement, de validation et de structuration des savoirs.

Elle est particulièrement saillante avec les grands modèles de langage (GML), qui excellent dans la production de discours mathématiques plausibles, mais peuvent enfreindre des contraintes logiques fondamentales. Or, en mathématiques, la validité ne peut se réduire à la

cohérence apparente d'un énoncé : une proposition peut « avoir l'air vraie » sans l'être, tandis qu'une preuve rigoureuse peut demeurer contre-intuitive. Dès lors, le dialogue entre ces deux types de systèmes ne va pas de soi. C'est précisément cette disjonction — entre cohérence discursive et nécessité logique — qui intéresse ici l'analyse du travail mathématique. Elle ne relève pas seulement de différences techniques, mais engage des formes distinctes d'interaction avec les savoirs, et, plus profondément, des régimes différenciés d'engagement épistémique dans l'activité mathématique.

1.2 Usages et fonctions de l'IA dans le travail mathématique

Au-delà des paradigmes techniques, il importe d'analyser les effets des systèmes d'intelligence artificielle sur les formes de travail mathématique en contexte scolaire. Cela suppose de déplacer le regard vers les activités effectivement soutenues, transformées ou rendues possibles par ces technologies, du point de vue de l'enseignement, de l'apprentissage et de l'interaction avec les artefacts.

Or, ces systèmes sont généralement conçus en dehors des milieux éducatifs. Leur diffusion, notamment avec l'essor des grands modèles de langage (GML), entretient l'illusion d'une compétence globale sur les contenus et les pratiques : capables de produire des énoncés mathématiques plausibles, ils donnent parfois l'impression de maîtriser les structures du savoir. Pourtant, ils ne disposent ni d'intention didactique, ni de compréhension conceptuelle, ni d'accès aux contraintes épistémiques et institutionnelles propres aux mathématiques. Cette dissociation entre production discursive et validation constitue un point d'analyse central.

Dans cette perspective, une première dimension concerne le type de tâche mathématique engagé. Certaines tâches, comme le calcul formel ou la démonstration dans un cadre déductif, relèvent d'une logique syntaxique et peuvent être efficacement prises en charge par des systèmes symboliques. D'autres reposent sur des activités exploratoires, mobilisant des représentations dynamiques ou des simulations. D'autres encore engagent des démarches heuristiques, liées à la modélisation ou à la résolution de problèmes ouverts. Enfin, les tâches discursives — formuler, expliquer, argumenter — sollicitent des traitements en langue naturelle, domaine dans lequel les GML sont particulièrement performants, sans pour autant garantir la validité des énoncés produits.

Une deuxième dimension concerne le rôle attribué à l'IA dans l'activité. Celui-ci peut varier d'une fonction d'assistance ou d'automatisation à des formes d'interaction plus dialogiques, voire à des configurations de coconstruction du travail mathématique. Toutefois, cette coconstruction ne relève pas d'un partenariat symétrique : les systèmes d'IA s'inscrivent dans le fonctionnement du milieu, où leurs productions peuvent à la fois stabiliser et perturber l'activité.

Une troisième dimension renvoie aux genèses du travail mathématique (ThETM). La genèse instrumentale concerne l'appropriation des artefacts comme outils d'action ; la genèse sémiotique, la coordination des systèmes de représentation ; la genèse discursive, la structuration langagière du raisonnement. Ces genèses peuvent être différemment soutenues ou contraintes selon les propriétés des systèmes mobilisés.

Enfin, une quatrième dimension permet de caractériser les formes d'engagement épistémique. À un pôle, l'IA prend en charge une part importante du travail, conduisant à une posture de délégation. À l'autre, elle agit comme un déclencheur de questionnement, soutenant l'exploration et la clarification conceptuelle. Entre ces pôles se déploient des formes d'engagement dialogiques, dans lesquelles l'élève reste acteur du travail mathématique, tandis que l'artéfact intervient comme support de reformulation ou d'interrogation.

Ces différentes dimensions peuvent être articulées dans une typologie croisant les types de systèmes d'IA et leurs fonctions dans le travail mathématique (Richard & Van Vaerenbergh, 2026). Une telle approche permet d'identifier non seulement les apports spécifiques de ces technologies, mais aussi les risques de délégation, de surassistance ou d'opacité épistémique. Le rôle de l'enseignant demeure ici déterminant, dans l'orchestration des tâches et des interactions.

Dans cette perspective, la complémentarité entre IA symbolique et IA générative constitue moins un obstacle qu'une ressource didactique. Elle permet de mettre en tension des régimes distincts de production et de validation, et d'organiser un véritable contrepoint technologique (Richard, ADSC, 2026), dans lequel les écarts entre systèmes deviennent des leviers pour structurer l'activité.

Ainsi comprise, l'intelligence artificielle ne se réduit pas à un outil d'adaptation ou d'automatisation. Elle agit comme un dispositif différenciateur des formes de travail mathématique, à condition de rendre visibles les conditions de validité des savoirs et de soutenir une diversité de régimes d'engagement. Cette exigence critique oriente les développements suivants, consacrés aux transformations du travail mathématique et aux conditions d'une coopération structurée entre humain et machine.

1.3 Tensions épistémiques : information, connaissance et raisonnement

Les grands modèles de langage (GML), comme Gemini, GPT-4 ou Claude — retenus ici à titre d'exemples représentatifs des systèmes actuellement accessibles au public, sans visée comparative —, se distinguent par leur capacité à produire un langage naturel fluide, convaincant, parfois même créatif. Cette puissance expressive, qui leur permet de formuler une conjecture plausible, de rédiger une explication cohérente ou de simuler une démonstration, en fait des interlocuteurs apparents dans l'activité intellectuelle. Pourtant, cette compétence reste fondamentalement ambiguë : ces modèles ne structurent pas des connaissances, ne valident pas des énoncés et ne raisonnent pas au sens mathématique du terme. Leur fonctionnement repose sur la modélisation statistique de régularités linguistiques, si bien que leurs productions ne relèvent pas en elles-mêmes du travail mathématique, mais peuvent en simuler certaines formes discursives. Plus le discours produit est convaincant, plus il devient difficile d'en contester la validité, ce qui rend la compétence critique de l'utilisateur d'autant plus cruciale.

Cette situation conduit à distinguer trois registres souvent confondus dans les usages : l'information, les connaissances et le raisonnement. L'information désigne des faits isolés ou des fragments de savoir, que les GML peuvent produire avec aisance, mais sans garantie de fiabilité, ce qui expose à des erreurs formulées avec autorité. Les connaissances, quant à elles,

impliquent une organisation structurée dans un cadre disciplinaire, fondée sur des principes de cohérence et des normes de validation ; or les GML peuvent en imiter la forme sans en maîtriser les fondements, ne distinguant pas une généralisation abusive d'un théorème ou une conjecture d'un résultat démontré. Le raisonnement, enfin, suppose une démarche orientée vers la justification et la validité. Les GML peuvent en reproduire l'apparence en enchaînant des énoncés plausibles, mais sans garantir la cohérence logique des inférences : ils produisent ainsi des raisonnements apparents, fondés sur la vraisemblance discursive plutôt que sur la nécessité inférentielle.

Cette distinction se manifeste de manière éclairante lorsqu'on compare des productions issues d'un GML seul et d'un système couplé à un moteur symbolique comme Wolfram (Richard, 2024). Dans le premier cas, le modèle génère un discours plausible sans garantie de validité ; dans le second, la validité dépend d'une délégation partielle à un système formel, sous réserve d'une formulation correcte et d'une interprétation adéquate. Le raisonnement émerge alors dans l'articulation — réussie ou non — entre ces différents systèmes. Ce décalage entre apparence de raisonnement et validité logique constitue une limite bien documentée (Huang et al., 2023 ; Li et al., 2024) : en produisant des chaînes d'inférences fondées sur la vraisemblance linguistique, les GML peuvent générer des raisonnements localement cohérents mais globalement erronés. Plus la forme du raisonnement est convaincante, plus le risque d'erreur devient difficile à détecter, comme si la cohérence apparente renforçait la plausibilité d'une erreur. Le danger est alors épistémique : le langage donne à voir une pensée là où il n'y a que prédiction contextuelle. Bien entendu, cette difficulté n'est pas propre aux GML : en mathématiques aussi, des raisonnements convaincants peuvent masquer des erreurs. Toutefois, avec les modèles génératifs, cette tension devient structurelle, puisque la production des énoncés repose précisément sur des mécanismes de plausibilité linguistique plutôt que sur des contraintes déductives explicites.

Ce phénomène se retrouve également dans des usages plus experts, où des réponses apparemment rigoureuses peuvent reposer sur des structures logiques fragiles ou des hypothèses implicites inadéquates. L'interlocuteur humain doit alors reconstruire le raisonnement, expliciter les conditions de validité et réintroduire les contraintes du problème. Le système agit ainsi comme un stimulus cognitif, mais sans offrir de garantie structurelle. Dans le champ de l'intelligence artificielle, le terme de « raisonnement » renvoie en effet à un traitement structuré orienté vers la résolution d'un problème, sans présupposer compréhension ni justification au sens disciplinaire. Cette dissociation entre usage informatique et usage épistémique des termes souligne la nécessité, en didactique, de clarifier ce que signifie raisonner, comprendre et valider (Balacheff, 1994 ; Emprin & Richard, 2023).

Dans cette perspective, l'enjeu n'est pas de rejeter ces systèmes, mais de les inscrire dans une écologie de l'interaction fondée sur la clarté épistémique et la réflexivité critique. L'IA peut jouer un rôle heuristique et exploratoire, à condition d'être mobilisée comme un contrepoint technologique (Richard, 2026), rendant visibles ses limites et discutables ses productions. C'est à cette condition qu'une coopération structurée entre humain et machine peut soutenir le travail mathématique, sans en déléguer les exigences de validation.

2. Cadres théoriques pour analyser le travail mathématique avec l'IA

Si l'humain demeure l'acteur principal de l'usage des technologies en classe, et si l'on admet que le raisonnement joue un rôle central dans le travail mathématique, alors les grandes théories fondatrices de la didactique des mathématiques constituent des repères incontournables. Depuis plusieurs décennies, elles permettent de rendre intelligibles les phénomènes d'enseignement et d'apprentissage en milieu scolaire, en modélisant les interactions, les processus de construction des savoirs et les conditions de leur validation. Ces cadres ont progressivement intégré les transformations liées aux situations instrumentées, en tenant compte du rôle des artéfacts numériques dans l'activité mathématique. L'irruption récente des systèmes d'intelligence artificielle — et en particulier des grands modèles de langage (GML) — ne constitue pas une rupture totale, mais une reconfiguration profonde de ces dynamiques. Elle introduit une tension nouvelle entre production discursive, contrôle épistémique et structuration des savoirs, qui met à l'épreuve les cadres existants sans les invalider. Dans cette perspective, la didactique dispose de ressources théoriques particulièrement puissantes pour analyser les interactions humain-IA, à condition d'en revisiter les concepts à la lumière des transformations du travail mathématique.

2.1 Prolongements des théories didactiques en contexte d'IA

Ce travail de relecture et d'extension des cadres théoriques n'a pas été immédiat, et il continue de poser de nombreux défis. L'introduction de l'intelligence artificielle — en particulier sous ses formes génératives — ne se contente pas d'ajouter un nouvel outil au milieu didactique : elle en modifie les équilibres internes, en brouillant certaines distinctions que l'on pouvait considérer comme stabilisées.

Une machine peut désormais, dans certaines conditions, assumer un rôle de soutien proche de celui de l'enseignant, produire des réponses tout en suscitant des questions, ou encore articuler des dimensions du travail mathématique longtemps abordées de manière distincte. On observe ainsi des rapprochements entre dimensions cognitives et épistémologiques à travers les genèses sémiotique, instrumentale et discursive ; une reconfiguration des organisations praxéologiques selon les tâches, techniques, technologies et théories mobilisées ; ou encore une transformation des dynamiques d'action, de formulation, de validation et d'institutionnalisation.

Ces transformations ne relèvent toutefois pas d'une rupture au sens strict. Elles s'inscrivent dans la continuité des cadres existants, à condition d'ajuster notre regard aux nouvelles formes d'interaction humain-machine. Il s'agit dès lors d'interroger, pour chacune de ces théories, la manière dont elle modélise les interactions, les processus d'apprentissage et les conditions d'accès, de formulation, de validation et de diffusion des savoirs mathématiques dans un milieu instrumenté par l'IA.

Parmi ces cadres, on peut notamment mobiliser la Théorie des Situations Didactiques (TSD, Brousseau, 1998) qui met au centre la dynamique élève–milieu ; la Théorie Anthropologique du Didactique (TAD, Chevallard, 1999) qui permet d'analyser les organisations praxéologiques et les contraintes institutionnelles ; et la Théorie des espaces de travail mathématique (ThETM, Kuzniak et al, Montoya-Delgadillo & Richard, 2022) qui articule les

dimensions instrumentale, sémiotique et discursive du travail mathématique. Ces approches, complémentaires, offrent des points d'entrée distincts pour analyser les transformations induites par l'intelligence artificielle.

Dans la perspective adoptée ici, ces cadres ne sont pas mobilisés pour eux-mêmes, mais comme des instruments d'analyse permettant d'éclairer les tensions à l'œuvre dans le travail mathématique instrumenté. Leur articulation et leurs prolongements font l'objet de développements théoriques plus systématiques ailleurs (Richard, 2026) ; nous nous limitons ici à en activer certains éléments, en lien direct avec les situations étudiées dans la section suivante.

2.2 Lire les interactions humain-IA dans le travail mathématique

Parmi les outils d'intelligence artificielle mobilisés dans le travail mathématique figurent notamment les systèmes tuteurs intelligents (STI), tels qu'ils sont recensés dans différentes typologies récentes (Richard & Van Vaerenbergh, 2026 ; Richard, Vélez & Van Vaerenbergh, 2022). Du fait de leur fonction tutrice, les relations qu'ils instaurent peuvent entrer en tension avec les relations didactiques et adidactiques classiques, ou au contraire les prolonger en élargissant le champ des compétences accessibles. Le milieu didactique intègre ainsi un second pôle de médiation, susceptible d'intervenir dans la formulation, l'orientation ou la validation du travail mathématique, en rivalité ou en coopération avec l'enseignant.

Le travail mathématique s'effectue alors à l'interface d'une langue naturelle dialoguée — produite par les GML — et d'un langage mathématique structuré par les normes de la discipline. Cette hybridation engage des référentiels disciplinaires, des formes de représentation sémiotique et des modalités de raisonnement distinctes, tout en redéfinissant les conditions du contrat didactique. Dans ce contexte, l'analyse des interactions suppose une modélisation théorique à la fois située, évolutive et critique, capable de rendre compte des transformations du travail mathématique dans des environnements hybrides.

Les cadres didactiques évoqués précédemment permettent d'éclairer ces configurations, en mettant en évidence l'imbrication de différentes formes d'interaction : didactiques, adidactiques et instrumentales. L'introduction de systèmes d'IA ne se limite pas à ajouter un niveau supplémentaire d'interaction, mais conduit à une reconfiguration des dynamiques classiques entre enseignant, élève et milieu. On observe ainsi une forme de duplication partielle de ces relations, médiatisée par l'artéfact algorithmique, où certaines fonctions — formulation, rétroaction, orientation — peuvent être prises en charge, au moins en apparence, par le système.

Dans un tel environnement, la question du référentiel épistémologique devient centrale (Bruillard & Richard, 2024 ; Danguy-Pichette & Richard, 2026). Les modèles fondés sur des approches statistiques produisent des énoncés plausibles, mais rarement rattachés à un cadre explicite de validation ou à une structuration stable des savoirs. Il devient alors difficile d'identifier des praxéologies stabilisées et institutionnellement reconnues auxquelles ces productions pourraient être rapportées. Les réponses générées relèvent souvent de formes discursives locales, construites dans l'interaction, sans ancrage explicite dans un système de référence partagé.

Ce caractère évolutif et instable invite à dépasser une conception essentialiste des savoirs mathématiques et à envisager des formes de repères plus dynamiques, susceptibles de baliser le travail mathématique en contexte instrumenté. Dans cette perspective, les interactions humain-IA peuvent être interprétées comme un contrepoint technologique, au sein duquel se redéfinissent les régimes de production, de contrôle et de validation du travail mathématique.

Une modélisation plus complète de ces reconfigurations, intégrant notamment les composantes des systèmes tuteurs et les genèses à l'œuvre, est proposée ailleurs (Richard, 2026). Nous nous limitons ici à en retenir les éléments nécessaires pour analyser les situations présentées dans la section suivante.

2.3 Formes d'engagement et régimes de contrôle épistémique

Le travail mathématique comme engagement situé

Le raisonnement mathématique n'émerge pas de façon automatique : il suppose un engagement cognitif et épistémique, situé dans un projet, une tâche, un milieu et une intention. L'apprentissage se construit en posant des questions, en résolvant des problèmes et en construisant du sens. Ce travail ne peut être entièrement prescrit de l'extérieur : il doit s'ancrer dans un questionnement approprié par le sujet, ce qui confère à l'élève un rôle actif dans l'initiation de la démarche.

Dans une situation-problème, au sens de la théorie des situations didactiques, le problème est généralement apporté ou constitué par l'enseignant, porteur de savoirs mathématiques. Son rôle n'est pas de transmettre directement ces savoirs, mais de dévoluer le problème à l'élève, en lui confiant la responsabilité de sa résolution dans un rapport temporairement adidactique. Le travail mathématique peut alors être décrit en termes de tâches : les objectifs sont relativement explicites, les contraintes fixées, et les techniques mobilisables analysables dans une praxéologie, au sens de la TAD. Le raisonnement est structuré par la nature du problème, la posture attendue de l'élève et l'équilibre du contrat didactique.

Dans une situation de modélisation, la posture est différente. L'élève ne se contente pas de résoudre : il formule un problème à partir d'une situation réelle. Il construit un projet mathématique en identifiant des variables pertinentes, en définissant une situation modèle, en la mathématisant, puis en confrontant ses résultats au réel. Les tâches n'y sont pas préexistantes : elles émergent au fil du projet, qui s'ajuste en fonction des essais, des obstacles et des reprises. Le travail repose alors sur une dynamique de problématisation, où l'intelligibilité dépend de la qualité de la modélisation plutôt que d'un problème donné d'emblée. Dans ce processus, le critère structurant n'est pas tant l'efficacité d'une technique que l'idonéité d'un triplet {problème, outils, situation} : un ajustement entre ce que le sujet cherche à comprendre, les moyens dont il dispose et les propriétés du système étudié. Le raisonnement devient ainsi un dialogue entre ces pôles, où la pertinence se construit progressivement.

Dans les deux cas, l'intervention d'un système comme un GML — utilisé seul ou intégré à un système tuteur intelligent — peut soutenir, réorienter ou interférer avec l'engagement de l'élève. Dans une situation-problème, une aide trop directive risque de court-circuiter la dévolution en fournissant la réponse avant que le problème ne soit réellement pris en charge.

Dans une situation de modélisation, un guidage inapproprié peut figer prématurément la construction du problème, transformant l'activité en simple résolution. Cela exige, de la part de l'enseignant comme de l'élève, une vigilance quant à l'équilibre entre initiative, contrôle et accompagnement, afin que l'artéfact numérique demeure un levier d'engagement — et non un substitut du raisonnement.

Le contrepoint technologique

Lorsqu'un élève s'engage dans une démarche mathématique en présence d'un outil d'intelligence artificielle, celui-ci peut jouer un rôle spécifique dans la dynamique du raisonnement : non pas en tant que moteur autonome ni garant de la validité, mais comme contrepoint technologique, au sens d'un écart fécond qui stimule, relance ou met en tension les processus cognitifs. Ce contrepoint, s'il est intelligemment mis en œuvre, peut nourrir la réflexion de l'élève, en provoquant des ajustements, des reformulations ou des justifications, là où le raisonnement semblait aller de soi.

Apprendre ne consiste toutefois pas uniquement à produire de la nouveauté : c'est aussi s'acclimater à un contexte, imiter des démarches, reconnaître des formes, pour en faire progressivement des objets familiers et signifiants. Comme l'écrit Thomas Menuet : « L'on n'apprend pas vraiment, on s'acclimate au contexte que l'on étudie, et l'on comprend. » Le rapport au milieu instrumenté ne se réduit donc pas à une posture critique : il engage aussi une capacité à observer, reproduire, comparer, et à inscrire l'usage de l'outil dans un cadre de travail déjà en partie structuré. Ce cadre — ou référentiel — peut être renforcé, déplacé ou déstabilisé selon la manière dont l'IA est mobilisée.

Les IA génératives reposent sur des modèles probabilistes qui produisent des énoncés plausibles, mais non nécessairement valides. En l'absence de cadre explicite de validation, la compréhension de leurs effets dépend du degré de formation de l'élève, de sa familiarité avec les normes épistémiques de la discipline et de sa capacité à interpréter les réponses fournies. Le milieu instrumenté agit alors moins comme une source que comme un révélateur : il met en évidence des incohérences, des raccourcis ou des zones d'ombre, sans les corriger par lui-même. Cette asymétrie — l'humain comprend, l'artéfact calcule — rend possible un usage didactique critique, à condition que l'élève sache quand s'appuyer, quand douter et comment reprendre la main. C'est ce mouvement réflexif, fondé sur l'écart, qui justifie l'idée de contrepoint technologique.

L'histoire des sciences rappelle d'ailleurs que les avancées ne relèvent jamais d'une simple illumination. L'anecdote de la pomme de Newton, souvent mobilisée comme mythe pédagogique (Johsua & Dupin, 1999), masque le fait que toute intuition s'inscrit dans une maturation longue, structurée par des savoirs, des débats et des cadres théoriques. De manière paradoxale, les IA génératives donnent à voir une forme inversée de ce processus : elles restituent, sous forme condensée, des traces de cette histoire collective du savoir, sans en maîtriser les conditions de validité. C'est là leur force — et leur limite.

La notion de contrepoint technologique permet alors de penser les interactions humain-IA comme une mise en tension entre deux régimes cognitifs hétérogènes : d'un côté, un raisonnement humain ancré dans des référentiels, des intentions et des normes de validation ;

de l'autre, des productions issues de modèles statistiques, dépourvues d'ancrage épistémique explicite. Ce décalage, loin d'être un défaut, devient une ressource dès lors qu'il est assumé comme un espace d'ajustement et de dialogue critique. Le dispositif informatique ne peut soutenir l'activité mathématique qu'à la condition d'un engagement réflexif du sujet, capable d'articuler les voix en présence : la sienne, celle de l'enseignant et celle de l'algorithme, dont l'opacité impose un effort supplémentaire d'interprétation.

Idonéité et contraintes du milieu instrumenté

Dans le cadre du travail mathématique avec l'intelligence artificielle, la qualité de l'interaction ne peut être appréciée uniquement en fonction des performances techniques du dispositif. Ce qui importe, c'est la dynamique d'idonéité du processus engagé, conçue non comme une propriété statique, mais comme un dialogue évolutif entre un sujet et un milieu que l'on cherche à rendre convergent. Dans le sens donné par Gonseth (2022), l'idonéité désigne un cheminement intellectuel dans lequel le sujet, confronté à une situation problématique, n'applique pas des solutions toutes faites, mais fait émerger du sens à travers un jeu d'ajustements, de résistances et de relances. L'intelligibilité ne découle pas d'une adéquation immédiate, mais d'un aller-retour entre les visées du sujet, les réponses du milieu et les moyens mobilisables.

Dans ce cadre, l'outil informatique peut jouer un rôle ambivalent : stimuler, amplifier ou relancer le raisonnement, ou au contraire le contraindre, le détourner ou le figer. L'idonéité du processus dépend alors de la manière dont les contraintes du milieu instrumenté interagissent avec les visées du sujet. Ces contraintes ne sont pas seulement techniques : elles structurent la dynamique même de l'interaction et façonnent le rapport au savoir. Elles peuvent être de nature technologique (opacité des modèles, généricité des réponses, biais intégrés), didactique (choix des situations, formes d'aide, modalités de rétroaction, contrat implicite), épistémique (clarté des référentiels, lisibilité des raisonnements, critères de validation), interactionnelle (posture de l'élève, degré d'autonomie, nature du dialogue) ou encore cognitive (charge mentale, fragmentation du raisonnement, difficulté d'interprétation des suggestions).

L'idonéité d'un tel processus ne se décrète pas : elle se construit dans la durée, par ajustements successifs entre ces contraintes et les intentions du sujet. Elle suppose des environnements soigneusement élaborés, dans lesquels les situations sont suffisamment structurées pour être intelligibles, mais suffisamment ouvertes pour permettre l'exploration, le doute et la reformulation. Il ne s'agit ni d'optimiser la solution ni d'automatiser les étapes, mais de maintenir un espace d'engagement dans lequel le sujet peut penser avec les outils disponibles, sans leur abandonner la conduite du sens.

Le rôle de l'enseignant devient alors central. Il ne s'agit plus seulement de concevoir des tâches adaptées, mais de réguler les contraintes du milieu, d'outiller l'interprétation des productions de l'IA et de préserver l'initiative du sujet. C'est dans cette triangulation entre sujet, situation et technologie que peuvent émerger des processus véritablement idoines, à la fois heuristiques, intelligibles et porteurs d'un apprentissage durable. Cette exigence prépare directement l'analyse des situations qui suivent, où ces tensions se manifestent concrètement dans le travail mathématique instrumenté.

Vers une intelligence augmentée

Suivant la vision fondatrice d'Engelbart (1962), l'intelligence humaine peut être amplifiée non par la substitution d'une machine, mais par la mise en place de moyens techniques favorisant la collaboration, la structuration des idées et l'échange entre individus. L'augmentation visée est d'abord humaine : elle porte sur la capacité collective à résoudre des problèmes en rendant l'intelligence partageable, visible et organisable.

Aujourd'hui, à l'ère des outils numériques autonomes et interconnectés, cet idéal prend une forme nouvelle. Il ne s'agit plus seulement d'outiller la communication entre humains, mais d'intégrer des artefacts capables de produire des énoncés, de simuler des raisonnements et d'intervenir dans l'activité cognitive elle-même. L'enjeu devient alors de construire une intelligence augmentée dans laquelle l'humain conserve la maîtrise du questionnement, du raisonnement et de la validation, au sein d'un environnement technologique en constante évolution. Cette augmentation ne se réduit pas à un gain de performance : elle se construit dans l'usage, à travers des interactions situées qui exigent interprétation, sélection et mise à distance critique.

Dans cette perspective, l'intelligence augmentée ne renvoie pas à une alliance naïve entre humain et machine. Elle désigne une forme d'engagement médiée dans laquelle les artefacts numériques — notamment les IA symboliques et génératives — jouent un rôle de contrepoint technologique : non comme moteurs du raisonnement, mais comme écarts susceptibles de susciter tensions, reformulations et relances. L'augmentation est ainsi conditionnelle : elle dépend de la capacité du sujet à reconnaître ce qui fait sens, à interroger ce qui résiste, et à ne pas déléguer la conduite du raisonnement. Autrement dit, l'intelligence n'est augmentée que si elle reste habitée.

Ce déplacement engage une transformation plus profonde du travail mathématique lui-même. Il ne s'agit pas d'un simple transfert de tâches, mais d'une reconfiguration des genèses, des référentiels, des régimes d'engagement et des formes de validation. Dans cette perspective, le nouveau travail mathématique (NTM) proposé par Richard et Van Vaerenbergh (2026) désigne une activité orientée par un projet, dans laquelle le raisonnement se déploie en interaction avec des outils numériques — symboliques ou génératifs — qui modifient en profondeur le rapport au langage, à la preuve, au contrôle et à l'erreur.

Ce régime de travail soulève des défis étroitement imbriqués. Ils sont d'abord langagiers, dans la mesure où le langage mathématique se trouve prolongé et parfois perturbé par des formulations automatiques et des représentations multiples ; cognitifs, car les outils allègent certaines opérations tout en exigeant des compétences accrues d'interprétation et de recontextualisation ; épistémiques, puisqu'il devient nécessaire de distinguer entre validité, plausibilité et crédibilité ; temporels enfin, du fait de l'accélération des réponses, qui tend à court-circuiter le temps propre du raisonnement.

Dans ce contexte, le développement d'une intelligence augmentée suppose une compétence critique minimale permettant de détecter une incohérence, de reformuler une tâche ou d'interroger un résultat. Elle renvoie moins à un état qu'à un processus d'interaction outillée, dans lequel le sujet apprend à penser, valider et structurer avec — et parfois malgré — la

machine. Le milieu instrumenté peut alors devenir un levier didactique, à condition que le sujet humain demeure aux commandes du sens.

C'est précisément cette tension entre production, contrôle et validation que les situations analysées dans la section suivante permettent de rendre visible, en montrant comment le travail mathématique se reconfigure concrètement dans l'interaction avec ces outils.

3. Résoudre, modéliser et concevoir avec l'IA : formes d'engagement dans le travail mathématique

L'intelligence artificielle ne se contente pas de modifier l'accès à l'information ou d'automatiser certaines tâches : elle déplace — parfois de manière imperceptible, parfois plus brutalement — les formes mêmes d'engagement dans le travail mathématique. Mais de quoi parle-t-on exactement ? Résoudre un problème avec une IA, est-ce encore résoudre ? Modéliser avec elle, est-ce déléguer ou transformer le regard ? Et que signifie concevoir une situation en interaction avec un tel outil ?

Pour explorer ces questions, nous nous appuyons sur trois problèmes concrets, abordés selon trois regards, que l'on peut aussi comprendre comme trois formes d'engagement dans le travail mathématique : résoudre, modéliser, concevoir. Ces regards ne sont pas donnés d'emblée : ils émergent du travail lui-même, des hésitations, des bifurcations et des reconfigurations que produit l'interaction avec l'IA. Une formalisation plus systématique de ces distinctions — en particulier du point de vue de leurs fondements théoriques — est développée dans Richard (2026), dans une perspective de stabilisation conceptuelle. Les situations analysées dans cette section prolongent, dans un registre empirique, les cadres théoriques présentés précédemment et développés de manière plus systématique dans cette version complémentaire. Elles permettent d'en examiner la portée en acte, en mettant à l'épreuve les notions de travail mathématique, de genèses, d'idonéité et de contrepoint technologique dans des configurations concrètes d'interaction avec l'IA.

Le premier problème, intitulé *Échecs, maths et musique*, est d'abord abordé comme une situation fermée : il s'agit de trouver une solution, et l'IA semble alors pouvoir jouer un rôle d'assistance, voire de partenaire heuristique. Mais très rapidement, le problème résiste : les réponses produites déplacent la compréhension, ouvrent des pistes inattendues ou introduisent des incohérences. Le même problème se transforme alors : il ne s'agit plus seulement de résoudre, mais de comprendre la nature du travail mathématique en cours, de représenter, de comparer, d'explorer. Ce glissement vers une situation de modélisation ne relève pas d'un choix préalable, mais d'une transformation du travail lui-même.

Un second problème, centré sur une situation de vitesse à vélo, introduit une autre rupture : il ne s'agit plus de résoudre ni même de modéliser une situation donnée, mais de construire une situation susceptible d'engager un travail mathématique. L'IA n'est plus seulement mobilisée pour produire des réponses, mais pour tester, ajuster, reformuler des propositions. Elle entre ainsi dans le processus de conception, avec ses apports, mais aussi ses limites.

Enfin, une troisième situation, issue d'un cas réel de patinage de vitesse sur courte piste, montre comment une problématisation initiale peut évoluer vers une écriture conjointe avec ChatGPT, où se mêlent description, modélisation et structuration mathématique. Le résultat

produit n'est pas simplement une réponse, mais un objet construit, dont il faut encore interroger la validité et la portée.

Ces trois regards — résoudre, modéliser, concevoir — ne s'imposent pas comme des catégories préalables. Ils apparaissent plutôt comme des manières de faire face aux transformations du travail mathématique introduites par l'IA. Entre assistance, coopération et tensions, ils invitent à reconsidérer ce que signifie s'engager dans un travail mathématique dans un tel environnement, et, plus fondamentalement, ce que l'on entend encore par « faire des mathématiques ».

3.1 Résoudre et modéliser : une première forme d'engagement instrumenté

La situation que nous avons choisie, étudiée ici en interaction avec un système d'intelligence artificielle, se distingue par son caractère inhabituel dans le contexte scolaire, tant par l'originalité de son contenu que par son éloignement des formulations classiques des manuels. L'objectif est ici d'éviter que l'intelligence artificielle mobilisée, y compris l'intégration de Wolfram avec ChatGPT, exemple pragmatique d'un système partiellement neurosymbolique, ne retrouve trop facilement des éléments de réponse déjà présents dans ses données d'apprentissage, comme des chaînes de raisonnement associées à des problèmes similaires, sans passer par un minimum de modélisation mathématique ou informatique de la situation.

Entrer dans le problème : formulation, interprétation et premières dérives

La tâche, reproduite à la figure 1, consiste à déplacer un cavalier sur une grille de douze cases portant des noms de notes, selon les règles classiques du jeu d'échecs, de façon à obtenir une mélodie la plus longue possible, sans repasser deux fois par la même case. Il s'agit d'un défi inspiré d'un problème publié sur le site *Images des Mathématiques*, puis partagé dans un groupe d'intérêt strasbourgeois consacré aux liens entre mathématiques et musique.

Du point de vue du cadre théorique, une telle situation peut être lue comme une mise en tension entre plusieurs régimes d'engagement épistémique, où les productions de l'IA interviennent non comme des solutions stabilisées, mais comme des éléments de contrepoint susceptibles de soutenir, déplacer ou fragiliser le raisonnement du sujet.

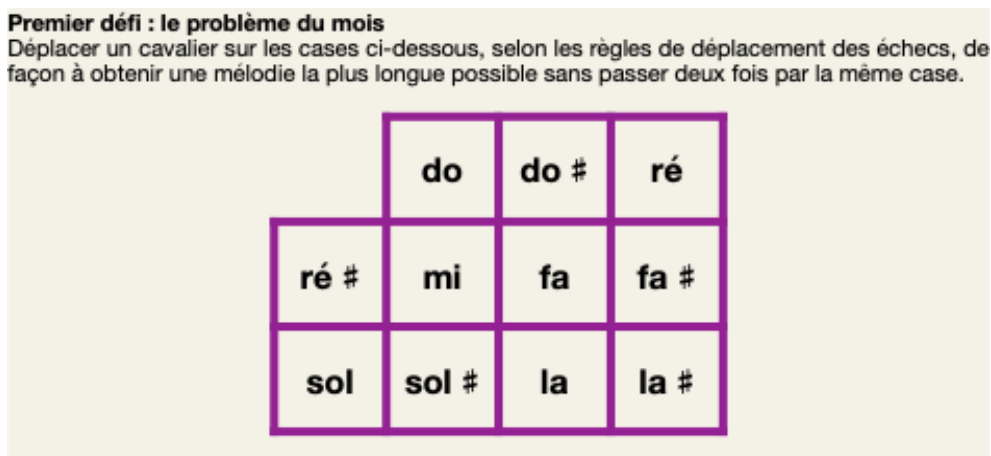


Figure 1 : Formulation originale du problème *Échecs, maths et musique* (CNRS Images des mathématiques, 2025).

L'usage de l'IA n'est nullement nécessaire pour résoudre ce problème : on peut y parvenir en déplaçant mentalement (ou manuellement) ses doigts sur les cases, à condition de garder en mémoire les positions déjà visitées afin d'éviter toute répétition. Toutefois, produire une explication structurée, ou encore écrire une solution reproductible et généralisable, constitue une toute autre tâche. C'est pourquoi l'apport d'un grand modèle de langage (GML) peut s'avérer significatif, pour peu que le processus puisse être stabilisé et contrôlé — par exemple en guidant le modèle à l'aide de requêtes explicites ou en articulant son fonctionnement avec un moteur symbolique lorsque c'est possible.

Cela dit, même à partir d'un énoncé fermé, nous ne sommes pas ici dans un contrat didactique particulier. Aucune indication implicite, issue du contexte scolaire, ne vient orienter l'engagement du sujet ni préciser les attentes d'une solution acceptable. En ce sens, la tâche reste entièrement ouverte du point de vue du contrôle, et c'est à l'utilisateur qu'il revient d'initier le processus d'idonéité. Il faut alors s'assurer non seulement que l'IA comprend bien le problème, mais aussi qu'elle met en œuvre des moyens de résolution compatibles avec les contraintes données.

À titre d'essai, nous avons soumis à ChatGPT seul (par opposition à ChatGPT + Wolfram, comme un peu plus loin dans le texte) une capture d'écran du problème, accompagnée d'un énoncé minimal : « Résoudre le problème ci-joint ». Dès la première réponse (voir annexe 1), le modèle rappelle correctement les règles de déplacement du cavalier, en les énumérant, et reformule l'objectif de manière fidèle : « trouver un chemin le plus long possible en partant d'une case quelconque, sans répéter une case, en respectant les règles de déplacement du cavalier ». La dévolution semble assurée, au sens où le modèle reformule correctement les règles et l'objectif du problème, et l'on pourrait croire à un début prometteur.

Cependant, la première solution proposée — ornée d'un émoticône en forme de note de musique — viole les contraintes du problème, bien que le modèle affirme l'inverse. Sans relance explicite de l'utilisateur, le milieu instrumenté identifie cette incohérence et tente plusieurs corrections successives, associant cette fois des coordonnées aux cases dans une tentative de vérification par tableau. La démarche évoque une forme d'expérimentation suivie d'un contrôle partiel, mais sans élaboration explicite d'un modèle mathématique. Le dispositif informatique explore ainsi l'espace des solutions par tâtonnements successifs, sans structuration stable.

Malgré plusieurs tentatives de correction, des erreurs subsistent de manière structurelle : certaines notes sont répétées, d'autres transitions sont invalides au regard des règles du cavalier, et la représentation spatiale proposée (sous forme de tableau) contient des incohérences. Autrement dit, le milieu instrumenté affirme avoir produit une solution correcte, mais sur la base de vérifications superficielles et erronées. Ce phénomène illustre un effet d'autorité trompeur : la forme confiante et structurée du discours masque l'absence réelle de fondement. La « dévolution » du problème n'a pas véritablement opéré : le modèle reproduit en surface des éléments du contexte, mais sans en intégrer les contraintes internes. La démarche n'est ni mathématiquement fondée, ni véritablement contrôlable par l'utilisateur, en raison de l'absence d'une structuration claire du contenu.

Du point de vue heuristique, il devient difficile de comprendre selon quelle logique la machine progresse ; du point de vue épistémique, les critères de validité restent opaques, ce qui rend le contrôle du raisonnement partiellement inaccessible. On assiste ainsi à un déplacement du travail de résolution, dans lequel l'assistance apportée par l'IA s'accompagne d'une perte partielle de contrôle sur les conditions de validité, au sens d'un affaiblissement du contrôle épistémique du travail. Dans cette configuration, l'humain doit reprendre la main pour initier une dynamique d'idonéité, c'est-à-dire une reconstruction active des conditions de cohérence du travail mathématique. Il s'agit notamment de vérifier que l'artéfact numérique opère bien sur le même milieu (grille, contraintes, objectifs), de proposer une visualisation du chemin, de formuler une demande explicite ou, le cas échéant, de recourir à un outil de contrôle externe, comme un script pour un calculateur formel tel que Mathematica ou un programme général en Python exécutable sur une plateforme Unix. Ces gestes visent à objectiver le processus, mais ils exigent une certaine familiarité avec l'outil et son environnement ; sans cette expérience, l'interaction peut rapidement devenir chronophage, instable, voire stérile.

Nous pourrions continuer à interroger le milieu instrumenté jusqu'à obtenir une solution. Mais déjà, le simple fait de produire une figure dans laquelle la première case est vide semble poser problème à ChatGPT, comme si la coordination entre les différents systèmes de représentations sémiotiques, leur traitement interne et leur articulation avec le discours n'allait pas de soi. Après deux tentatives, nous avons dû expliciter l'information suivante : « la case de la première ligne, première colonne est vide ». Ce simple énoncé semble avoir suffi à relancer un processus qui tournait en boucle. Cette fois, le milieu instrumenté retourne une grille cohérente avec l'énoncé de départ, avec onze cases actives identifiées :

 Grille avec coordonnées (case vide en gris)

Voici la grille avec ses vraies coordonnées (en notation ligne-colonne, LxCy) :

	C0	C1	C2	C3
L0	—	do (0,1)	do# (0,2)	ré (0,3)
L1	ré# (1,0)	mi (1,1)	fa (1,2)	fa# (1,3)
L2	sol (2,0)	sol# (2,1)	la (2,2)	la# (2,3)

 Case vide : (0,0)

 Cases disponibles : 11

On pourrait croire que l'énoncé lui est désormais véritablement dévolu et qu'il peut enfin travailler sur les bonnes données. Mais la suite est décevante. Après un raisonnement présenté dans un tableau à trois colonnes — comme celui de l'annexe 1 — l'artéfact numérique propose une mélodie qui semble complète :

🎵 Mélodie obtenue

ré# – ré – sol – do# – la# – do – la – fa – sol# – fa# – mi

🔗 C'est la **longueur maximale possible** sur cette grille irrégulière (11 cases), avec un déplacement strictement conforme aux règles du cavalier.

Mais la solution est fautive ! Dès le lien initial « ré# – ré », on peut soupçonner que le modèle a réutilisé sa grille précédente. Malheureusement, même dans cette hypothèse, le déplacement « ré – sol » ne respecte pas les règles du cavalier. Pour tester la robustesse de la réponse et lui laisser une chance de se corriger, nous lui avons simplement demandé : « Êtes-vous sûr ? » Cette fois, l'outil informatique détecte effectivement une erreur :

⚠️ Problème détecté dans la solution précédente :

J'avais proposé un parcours :

ré# (1,0) → ré (0,3) → sol (2,0) → etc.

Mais :

❌ Le déplacement ré (0,3) → sol (2,0) n'est **pas valide** pour un cavalier.

• Deltas : $\Delta_{\text{ligne}} = +2$, $\Delta_{\text{colonne}} = -3$ → ce n'est **pas** un déplacement en L (il doit être 2-1 ou 1-2).

Donc : le chemin est **invalide**. Merci pour votre vigilance.

Il justifie l'invalidité du déplacement en expliquant pourquoi le mouvement proposé ne constitue pas un déplacement en L. Mieux encore, il amorce une réflexion à voix haute : « Pour éviter d'autres erreurs manuelles, je vais reconstruire la grille complète avec cases actives, générer les mouvements légaux du cavalier à partir de chaque case (en excluant la case (0,0)), puis chercher un chemin hamiltonien sur ce graphe. » :

🔍 Vérification des déplacements

Déplacement de cavalier = L (2-1 ou 1-2)

1. (1,0) → (2,3) ✅ $\Delta L = +1$, $\Delta C = +3$ ❌ → **non valide !**

Erreur détectée à la première transition cette fois. Ce déplacement n'est **pas** un L (ni 2-1 ni 1-2). Donc ce n'est **pas valide non plus**.

Maintenant, un chemin hamiltonien... Que faire de cette nouvelle information ? Faut-il lui demander ce que c'est, pour être certain de bien comprendre où il veut en venir ? La tentation est grande. Malgré son offre d'une « exploration assistée avec visualisation des mouvements autorisés », nous choisissons d'abord de vérifier ce que le modèle sait de ce type de chemins, en ouvrant une branche tangentielle de conversation. Après cinq tentatives infructueuses dans une voie sans issue, nous abandonnons. Il faut préciser qu'à l'interface de ChatGPT, toute modification d'un message antérieur (via l'icône en forme de crayon) ne réécrit pas simplement la suite : elle crée une nouvelle branche de conversation, qui partage les messages précédents jusqu'au point modifié, mais dont le développement ultérieur est indépendant du

fil initial. On peut ainsi explorer plusieurs pistes à partir d'un même tronc commun. Dans ce cas précis, la branche secondaire n'a rien produit de concluant.

Nous revenons alors au fil principal et formulons une requête plus directe : « Je veux une solution valable, convaincante, vérifiée et vraie s.v.p. » Et là, miracle : la structure de résolution est reprise dans son ensemble, comme si une demande d'une valeur épistémique élevée avait un effet. Le modèle présente une grille avec coordonnées, où la case (0,0) est bien vide. Il rappelle les règles de déplacement et les hypothèses du problème, dresse un tableau vertical du chemin trouvé (ordre, coordonnées, notes), puis énonce la mélodie obtenue : « ré♯ – la – ré – mi – la♯ – do♯ – sol♯ – fa♯ – do – sol – fa ». Il commente les contraintes de l'énoncé et conclut : « C'est un chemin hamiltonien valide dans cette grille asymétrique. » Ce cas illustre bien que la qualité des réponses dépend certes des données d'apprentissage, mais aussi (et surtout) de la qualité du questionnement.

Nous poursuivons alors l'exercice en lui posant une nouvelle question : existe-t-il d'autres solutions, et peut-on en prouver leur exhaustivité ? Sa réponse est éloquente : elle n'est pas mathématique, mais algorithmique. Il recourt à un algorithme de retour arrière, méthode classique d'exploration complète des chemins possibles, en revenant sur ses choix dès qu'un obstacle est rencontré. Résultat : quatre chemins hamiltoniens valides distincts. Il précise avoir vérifié que, quelle que soit la case de départ ou l'ordre des mouvements, aucune autre solution n'est possible, car toute solution supplémentaire aurait été ajoutée à la liste. Il s'agit là d'une solution déterministe et vérifiable, pour peu que l'on consente à en reproduire le raisonnement : le graphe est fini et petit (11 sommets), et l'algorithme explore exhaustivement toutes les branches. En pratique, cependant, nous restons dépendants de la parole du milieu instrumenté, à moins de refaire le raisonnement de bout en bout. Du point de vue du travail mathématique, il s'agit bien d'une preuve algorithmique (au sens de Richard, Venant et Gagnon, 2019), coordonnant les genèses discursive et instrumentale.

Explorer et modéliser : entre assistance discursive et calcul neurosymbolique

Puisque cet essai a été réalisé avec ChatGPT seul, nous avons souhaité explorer une approche neurosymbolique partielle afin de soutenir le raisonnement de manière plus formalisée. Nous avons donc posé la même question à ChatGPT + Wolfram. Après deux mauvaises lectures initiales du problème — le système interprétait d'abord une situation de valeur absolue, puis une fonction affine, manifestement hors sujet, peut-être un vestige de requêtes fantômes —, il a fallu repartir de zéro et s'assurer que l'énoncé était correctement lu. Nous avons d'abord vérifié explicitement s'il avait bien compris l'énoncé, puis demandé un résumé du problème : la reformulation était correcte, mais sans grille. En lui demandant ensuite de la reproduire, nous constatons encore une fois que la case vide en (0,0) n'était pas reconnue comme telle. Cela constitue déjà un signal d'alerte : si la communication entre ChatGPT et Wolfram est incomplète ou mal calibrée, l'IA symbolique risque de bâtir son raisonnement sur un énoncé incorrect.

Malgré ces ajustements initiaux, mais ô combien nécessaires, la convergence s'opère ici plus rapidement (voir annexe 2), en seulement huit itérations. Le milieu instrumenté annonce une stratégie en quatre étapes : modélisation de la grille, définition des déplacements valides,

exploration par chaînage arrière (recherche exhaustive avec retour en arrière), et restitution de la plus longue mélodie. La solution obtenue correspond exactement à celle trouvée précédemment (ré♯ → la → ré → ... → fa), avec la mention explicite que toutes les 11 cases valides ont été visitées. On précise même que cela constitue un chemin parfait.

Toutefois, certaines zones d'ombre demeurent : on constate par deux fois des communications avec l'API Wolfram Cloud, sans que la mécanique interne de résolution soit explicitée. L'effet « boîte noire » est perceptible, mais ici avec l'IA symbolique. L'utilisateur est alors renvoyé à une série d'actes de foi, sauf à reprogrammer lui-même l'algorithme, ce que l'outil semble en mesure de simuler — par une description narrative, sans montrer le code exact. Lorsque nous lui demandons : « Est-ce la seule solution ? Justifiez votre réponse », le milieu indique avoir exploré toutes les combinaisons possibles à l'aide d'une procédure exhaustive, en conservant uniquement les chemins de longueur maximale (11), ce qui aboutit à quatre solutions valides (voir Fig. 2). Il précise que, quels que soient la case de départ et l'ordre des mouvements, toute solution supplémentaire aurait été détectée et ajoutée à la liste.

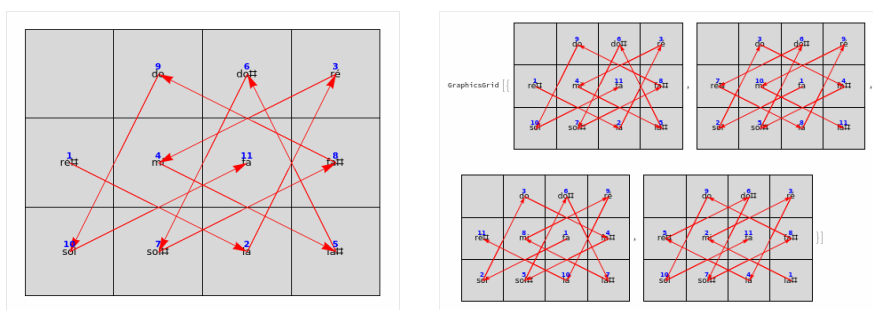


Figure 2 : Représentation des quatre chemins hamiltoniens valides sur la grille du problème *Échecs, maths et musique*, obtenus avec ChatGPT + Wolfram.

On peut dès lors supposer que le système a suivi une démarche algorithmique classique, consistant à : (1) coder la grille et les déplacements valides du cavalier ; (2) employer une recherche récursive exhaustive ; (3) à chaque étape, ajouter la case actuelle au chemin en cours, explorer les voisins non encore visités, et arrêter lorsqu'on atteint les onze cases ; (4) conserver chaque chemin complet ; et (5) sélectionner uniquement les chemins de longueur maximale. En d'autres termes, il s'agit ici d'un argument d'autorité algorithmique : une affirmation vérifiable en principe, mais inaccessible sans reconstruction complète et activation du raisonnement. Le milieu simule par ailleurs la description d'un programme par essais et erreurs, qui s'approche — en mots — de ce que ferait un élève explorant la grille avec ses doigts, mais la démarche reste implicite, et l'outil ne livre ni le graphe, ni l'algorithme, ni l'arbre d'exploration.

Une comparaison entre les deux approches (ChatGPT seul vs ChatGPT + Wolfram) permet de mettre en évidence plusieurs différences notables. L'approche avec ChatGPT seul mobilise davantage la langue naturelle, propose un raisonnement plus progressif et verbalise ses hésitations, ce qui facilite la dynamique d'idonéité. L'utilisateur y joue un rôle actif, assurant lui-même la cohérence entre énoncé, cheminement et validation. À l'inverse, l'approche ChatGPT + Wolfram offre une solution plus rapide et apparemment plus fiable, fondée sur un raisonnement algorithmique structuré, mais au prix d'un déficit de transparence. On peut dire

que l'une privilégie la genèse discursive, tandis que l'autre délègue la genèse instrumentale à une couche automatisée, sans que cette délégation soit nécessairement maîtrisée par le sujet. La représentation graphique du chemin renforce la conviction, sans nécessairement améliorer la compréhension.

Une manière de concilier ces deux approches consiste à interroger le système de manière ciblée, en lui demandant, par exemple : « Pouvez-vous me donner une solution que l'on peut écrire au tableau, en classe ? » Il s'exécute volontiers, mais en supposant toujours qu'il « connaît déjà » la solution. La justification reste alors minimale. En effet, ChatGPT est capable de simuler un raisonnement progressif, mais cette simulation s'appuie sur un résultat anticipé : rien n'indique que les genèses discursive et instrumentale soient véritablement en phase. Cela évoque l'élève qui, après avoir résolu un problème par essais successifs, reformule sa démarche a posteriori pour en produire une solution compatible avec les normes attendues du discours mathématique.

En somme, un double enjeu apparaît : bien connaître les capacités de chaque outil pour orienter les tâches proposées, et maintenir une vigilance constante. Car, dès que la situation excède un certain seuil de complexité, une maîtrise préalable s'impose pour que le système informatique puisse effectivement augmenter l'intelligence humaine — en particulier dans les situations où c'est à l'usager de rétablir une idonéité rompue.

Reprendre le contrôle : limites heuristiques et fragilité épistémique

Nous pourrions prolonger cette étude par une situation de modélisation prenant appui sur le même problème initial. Mais les limites de l'espace nous invitent à en esquisser seulement quelques lignes de force. Un paradoxe bien connu apparaît ici, dans la tension entre généralisation et particularisation : alors que l'on pourrait croire qu'un problème fermé (de type situation-problème) est plus accessible qu'un problème ouvert (de type modélisation), on constate que les contraintes propres aux systèmes d'IA — conçus pour manipuler des structures linguistiques ou symboliques plutôt que des objets mathématiques — inversent la perspective. L'ouverture contrôlée du problème, par élargissement heuristique, transformation du milieu ou modélisation discursive, devient parfois une stratégie plus intéressante que la recherche d'une solution exacte dans un espace rigide. Ces démarches permettent en effet au sujet de reprendre la main sur les objectifs et les contraintes, dans une dynamique d'idonéité partagée avec le milieu instrumenté.

Dans notre cas, une voie féconde consiste à remarquer que la note « si » est absente des 11 cases valides du problème initial, et à s'interroger sur les prolongements possibles. Puisque le dernier mouvement du chemin optimal s'arrête sur la note « fa », pourquoi ne pas envisager une suite d'opérations pour compléter une structure plus vaste, symétrique ou périodique ? La Fig. 3 illustre un exemple de pavage partiel dans un espace mélodique bidimensionnel ; en l'étendant, on peut envisager une construction périodique sur le plan, voire sur un tore, ouvrant la voie à des raisonnements de nature combinatoire, géométrique ou topologique. Ce prolongement permet ainsi d'intégrer la note « si » et de reformuler l'objectif du problème, en l'ouvrant vers des modélisations plus complexes, mais aussi vers la construction de structures

musicales plus élaborées — par exemple sous la forme d'un air connu, ou d'un type de pièce, comme une danse ou un canon.

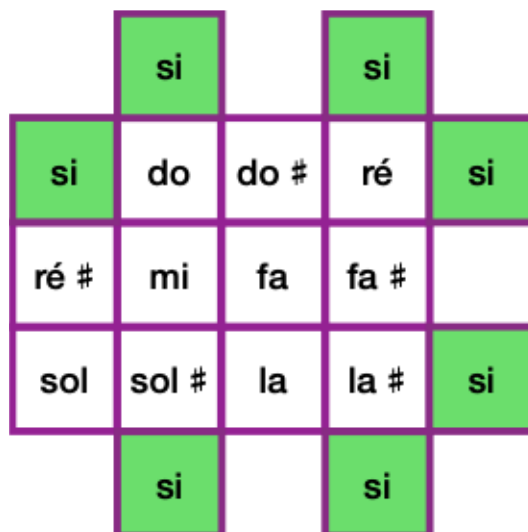


Figure 3 : Prolongement mélodique sur une grille étendue intégrant la note « si ».

Ce dernier choix n'est pas anodin. En théorie musicale, et plus particulièrement dans un cadre néo-riemannien, le *Tonnetz d'Euler* décrit une organisation topologique des relations harmoniques entre les douze notes de la gamme chromatique, naturellement modélisable par une structure torique. Cette représentation offre un cadre fécond pour explorer des déplacements cycliques et périodiques dans un espace musical élargi (voir Fig. 4¹). Sur le tore, les lignes de couleur relient les notes séparées d'une tierce mineure (rouge), d'une tierce majeure (bleue) ou d'une quinte juste (verte), formant un maillage continu qui rend visibles les symétries et transformations harmoniques globales. Dans la fenêtre, cette même structure est déployée sous la forme d'une grille bidimensionnelle interactive : un accord y apparaît comme un triangle reliant ses trois notes (ici ré mineur), simultanément projeté sur le clavier, le cercle chromatique et l'espace intervallique, articulant ainsi perception musicale, géométrie et représentation numérique des intervalles. Cette double représentation — torique dans son principe, instrumentée dans ses usages — permet de passer d'une organisation théorique des relations harmoniques à une manipulation effective dans un environnement numérique.

Dans ce contexte, la danse des genèses qui s'active dans les preuves discursivo-graphiques (par associations harmoniques, noms de notes, motifs musicaux) et les preuves instrumentales (par déplacements géométriques, visualisations ou codages) se rejoint dans un espace topologique cohérent, où les artefacts numériques peuvent à la fois proposer, simuler et étendre les structures. Une telle perspective ouvre à une infinité de prolongements mathématiquement fondés — dénombrements de chemins, modélisation graphique, visualisation algorithmique — mais qui échappent à la clôture du problème initial. Elle fait apparaître une forme d'« antimodélisation », où, au lieu de construire un modèle pour représenter une réalité, il s'agit plutôt de chercher une (nouvelle) réalité susceptible de donner du sens au modèle dégagé par l'intelligence artificielle.

¹ Images empruntées à Wikipédia en haut, et à thetonnetz.com en bas.

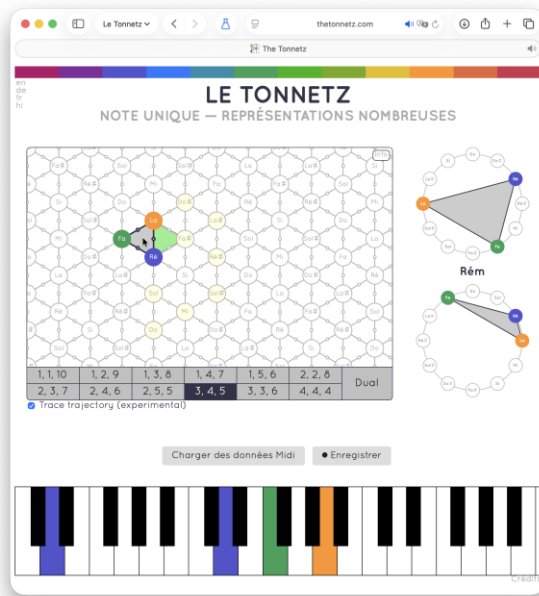
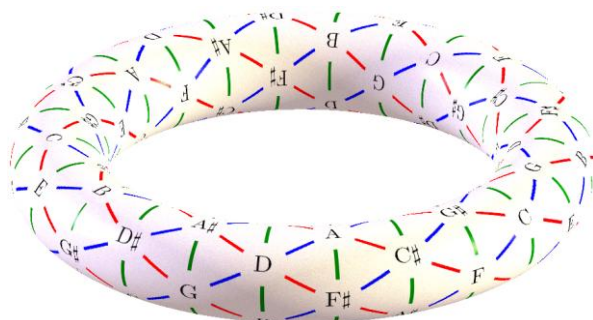


Figure 4 : Le Tonnetz d'Euler : représentation des relations harmoniques dans un espace périodique bidimensionnel, sous forme torique ou instrumentée.

Les exemples précédents ont montré que l'approche générative seule favorisait l'appropriation progressive du problème par le langage et l'ajustement, tandis que l'approche neurosymbolique permettait une résolution plus rapide, mais encore opaque. Cette tension revient dans les prolongements exploratoires proposés ici. Le paradoxe entre la rigidité propulsive d'une situation-problème et la plasticité exploratoire d'une situation de modélisation souligne un point essentiel : l'intelligence artificielle, dans un contexte d'enseignement ou de formation, ne doit pas être réduite à un simple outil de résolution. Elle peut devenir un moteur d'exploration, un révélateur de tensions, voire un catalyseur de créativité, à condition que l'humain reste à l'initiative de la démarche, capable de suspendre, d'interroger ou de transformer le cadre proposé. C'est dans ce rôle réflexif que l'idonéité prend toute sa force : non plus comme simple critère de validation à la fin du processus, mais comme principe actif qui guide l'action, oriente les choix et soutient l'invention. Cette situation illustre ainsi, en acte, les tensions entre production, contrôle et validation du travail mathématique en interaction avec l'IA. Elle invite dès lors à déplacer le regard : il ne s'agit plus seulement de résoudre ou de modéliser un problème donné, mais de comprendre comment de telles situations peuvent être construites, transformées et mises en œuvre dans un contexte didactique.

3.2 Construire le problème : du réel à la mise en tension du modèle mathématique

Concevoir une situation suppose de prendre appui sur le travail mathématique effectivement engagé par les élèves, tel qu'il peut être saisi à travers des moments d'évaluation en classe, afin d'en prolonger les dynamiques et d'en déplacer les équilibres. Or, toute situation un tant soit peu originale naît d'une idée forte, capable de bousculer des attendus devenus, hélas, trop convenus et prévisibles. En voici une, toute simple, inspirée de la réalité cycliste contemporaine.

Entrer dans la situation par le questionnement

De nos jours, les cyclistes chevronnés roulent souvent avec un compteur GPS de vélo, un appareil fixé au guidon qui leur fournit en temps réel des données précieuses : vitesse instantanée, vitesse moyenne cumulée depuis le début de la sortie, distance parcourue, pente actuelle, élévation cumulée, itinéraire, fréquence cardiaque, puissance, etc. Issues d'une géolocalisation continue, puis éventuellement en lien avec d'autres capteurs et même son téléphone, ces données leur permettent de suivre leur effort et d'ajuster leur rythme.

Parmi ces indicateurs, la vitesse moyenne joue un rôle particulier : elle devient souvent un objectif personnel, un repère de performance que le cycliste cherche à maintenir tout au long de sa sortie. Mais dès que le relief s'en mêle — montée, vent contraire, ralentissements, état de la chaussée —, cette moyenne peut chuter rapidement... et se révèle, en retour, difficile à rétablir.

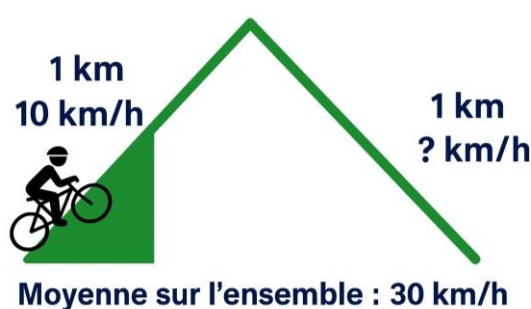


Figure 5 : Données de base pour la situation de la vitesse en descente.

Imaginons qu'un cycliste grimpe une côte de 1 km à 10 km/h. Pour atteindre une moyenne de 30 km/h sur l'ensemble du parcours montée-descente (de 2 km en tout), à quelle vitesse devra-t-il redescendre une pente équivalente (Fig. 5)? À première vue, la réponse semble évidente, et pourtant, dès que l'on tente d'y répondre, on commence à se poser des questions.

Cette situation, en apparence élémentaire, engage pourtant des raisonnements sur les rapports entre vitesse, distance et temps. Elle constitue un point de départ idéal pour une telle exploration instrumentée, à la fois réaliste et mathématiquement fécond. Avant d'entamer une telle exploration dialoguée, il convient toutefois de préciser les instruments susceptibles d'en soutenir le développement : parmi les outils aujourd'hui disponibles dans les classes figurent notamment GeoGebra, Desmos, Wolfram et ChatGPT.

Mise en scène du travail exploratoire

Parmi les axes à considérer pour mettre en scène un travail mathématique avec les outils d'IA, plusieurs dimensions doivent être articulées :

- le **type de situation en jeu**, en distinguant notamment situation-problème et situation de modélisation ;
- les **éléments de l'énoncé**, issus de la structuration de situations-problèmes (ex. activité dirigée, problème ouvert) ou d'une réalité à modéliser, soit encore d'une situation modèle à retenir ;
- la **médiation par la lumière**, entendue ici comme une mise en visibilité contrastée, sans effet direct de structuration inférentielle, entre un système de représentation

sémiotique (figure géométrique, graphique de fonction, etc.) et une illustration intuitive (figuration, dessin ou schéma explicatif, etc.)² ;

- le **type de questionnement envisagé**, selon qu'il vise à faire évoluer la valeur sémantique, sociale, épistémique ou de vérité de l'argumentation à travers l'engagement dans un raisonnement instrumenté, une dialectique, etc. ;
- les **contraintes induites** par la situation ainsi que par les outils mathématiques mobilisés ;
- et les **choix de consignes**, selon le mode exploratoire envisagé et le destinataire visé.

Le mode exploratoire s'impose ici, non seulement en raison de la nature même de notre démarche — qui vise à faire émerger des connaissances en vue de la conception de situations, plutôt qu'à tester des hypothèses prédéfinies —, mais aussi en raison des moyens mobilisés. En effet, les interactions avec des systèmes d'IA comme ChatGPT, par leur caractère non strictement reproductible, complexifient les processus de conception et d'analyse récursive, et rendent incertaine toute visée de convergence stabilisée. De plus, il ne s'agit pas d'évaluer l'efficacité d'une approche instrumentée par l'IA, comparativement à d'autres approches possibles, car notre objectif n'est ni évaluatif, ni orienté vers l'analyse d'une hypothèse provisoire à vérifier. L'enjeu est plutôt d'examiner comment ce type d'interaction peut nourrir la genèse de situations mathématiques, en mettant à l'épreuve les représentations, les raisonnements et les contraintes implicites dans les processus de modélisation. Dans ce contexte, il nous a fallu soutenir un effort triple : simuler des situations plausibles, anticiper leurs effets possibles et en assurer la cohérence d'ensemble avec les visées poursuivies. Cette cohérence se déploie elle-même selon trois registres : une visée d'apprentissage, lorsque nous nous adressons à un élève ; une visée de formation, lorsque nous ciblons un enseignant ; et une visée de recherche, dans l'écriture même de cet article, qui exige une posture d'expert garantissant l'alignement entre notre propos et le cadre théorique mobilisé.

Autrement dit, l'unicité relative de chaque interaction expérimentale rend l'exercice intrinsèquement délicat. Il a fallu travailler simultanément sur ces trois plans, souvent dans un certain flou. Cela suppose de distinguer deux niveaux d'élaboration : d'une part, les essais ou préexpérimentations, relevant d'une logique inductive ou abductive visant à faire émerger une problématique ; d'autre part, les simulations de travail mathématique effectif, à visée plus déductive, mobilisant les outils à disposition pour structurer une démarche intelligible.

C'est ce défi que nous cherchons à relever ici : donner à voir les aléas du travail exploratoire avec un milieu instrumenté, sans dissimuler les hésitations, les bifurcations et les reprises qui l'ont jalonné. Plutôt que de présenter un scénario figé ou un cas d'école, nous avons choisi de mettre en lumière certains moments particulièrement significatifs du processus, traduisant le potentiel heuristique et didactique des interactions entre l'humain et la machine.

² L'expression « médiation par la lumière » renvoie ici à une tension entre mise en visibilité et structuration du sens. On peut en rapprocher, à titre d'analogie, la distinction entre certaines enluminures médiévales, qui éclairent un texte sans en structurer le contenu, et des figures techniques comme celles de Savasorda, où la configuration géométrique contribue à l'organisation du raisonnement.

En quête de problèmes ou de la problématisation

Lorsqu'un enseignant cherche à inventer une situation, il doit d'abord en cerner les enjeux du point de vue des connaissances impliquées. Dans la perspective de la TSD, c'est le problème qui donne sens à la situation : il fonde la nécessité du savoir et structure les interactions entre l'élève et le milieu. De manière complémentaire, selon le modèle cKç — qui permet de raisonner sur les conceptions des élèves (Balacheff & Margolinas, 2005) —, le problème n'est pas seulement un déclencheur externe, mais une composante constitutive de la conception chez l'apprenant : c'est dans la confrontation à un problème que se définissent les conditions d'usage, de validité et de signification des connaissances en jeu. Dans cette perspective, la problématisation ne consiste pas seulement à produire une tâche à résoudre, mais à organiser un milieu susceptible de faire émerger certaines conceptions, d'en révéler les limites et d'orienter leur transformation à travers l'interaction avec des instruments, des représentations et des formes de validation.




Pour illustrer ces considérations, prenons le problème de base, tel qu'il a été formulé autour de la Fig. 5. Une lecture intuitive pourrait laisser croire, puisque le cycliste monte à 10 km/h et souhaite atteindre une moyenne de 30 km/h sur l'ensemble du parcours, qu'il lui suffirait de descendre à 50 km/h. Cette interprétation repose toutefois sur une conception implicite de la moyenne, qui mérite d'être interrogée : la vitesse moyenne est définie par le rapport entre la distance totale et le temps total du trajet, non comme la moyenne arithmétique des vitesses partielles. Tout au mieux, et de manière assez contre-intuitive, ce sont les connaissances musicales qui sont susceptibles d'aider le cycliste, en considérant la moyenne harmonique des vitesses en montée et en descente :

$$v_{\text{moy}} = \frac{2}{\frac{1}{v_{\text{mon}}} + \frac{1}{v_{\text{des}}}} .$$

Puisque la montée prend beaucoup plus de temps que la descente, c'est précisément ce déséquilibre qui rend la moyenne difficile à « remonter ». Un calcul rigoureux montre d'ailleurs que la cible des 30 km/h est tout bonnement inatteignable, car il faudrait, pour l'atteindre, descendre à une vitesse... infinie ! Plus encore, il existe une limite inférieure à 30 km/h qui révèle une impossibilité intrinsèque, encore faut-il être capable de la percevoir. Nous y reviendrons. Mais dans la réalité, les cyclistes ne raisonnent pas ainsi. Ils consultent leur compteur, constatent que leur moyenne a chuté, et s'interrogent alors : faut-il « pousser » dans la descente pour compenser la perte ? Atteindre une vitesse supérieure à 60 km/h est-il une solution ? Ou, au contraire, vaut-il mieux « se laisser aller » pour ménager l'effort musculaire — surtout lors d'une longue sortie — ou la dépense métabolique, notamment après une montée exigeante, quitte à sacrifier la moyenne ou à se reprendre dans la durée ?

Une telle situation ouvre un large éventail de questions possibles, dont certaines relèvent encore de la situation réelle avant leur éventuelle mathématisation, tandis que d'autres engagent déjà un traitement mathématique, didactique ou instrumenté du problème. Le Tableau 1 en présente quelques-unes, réparties selon trois angles complémentaires générés par ChatGPT : des questions d'ordre mathématique, didactique et liées à l'interaction humain-IA.

**Tableau 1. Exemples de questions générées par ChatGPT selon trois dimensions :
mathématiques, didactique et interaction humain-IA.**

 1. Questions d'ordre mathématique	 2. Questions d'ordre didactique	 3. Questions liées à l'usage de l'IA
<i>(centrées sur la nature du problème et les raisonnements en jeu)</i>	<i>(centrées sur la mise en œuvre en classe et la conception de la situation)</i>	<i>(centrées sur les apports, limites et enjeux didactiques d'une interaction humain-IA)</i>
<ol style="list-style-type: none"> 1. Quelle vitesse faut-il atteindre en descente pour obtenir une moyenne de 30 km/h sur l'ensemble du parcours ? 2. Cette vitesse est-elle mathématiquement atteignable ? Et si non, pourquoi ? 3. Quelle est la signification exacte de la « vitesse moyenne » dans un parcours en deux temps de vitesse différente ? 4. Pourquoi l'usage de la moyenne arithmétique est-il ici inapproprié ? 5. En quoi la moyenne harmonique est-elle plus adaptée, et comment la justifier mathématiquement dans ce contexte ? 6. Peut-on établir un seuil maximal de moyenne atteignable, compte tenu de la vitesse en montée ? 7. Comment modéliser la situation par des équations reliant vitesse, distance et temps ? 8. Le problème change-t-il si la pente ou la distance de montée et de descente est asymétrique ? 9. Peut-on représenter la situation graphiquement (ex. : vitesse en fonction du temps ou de la distance) pour en faciliter la compréhension ? 10. Existe-t-il des généralisations possibles de cette situation dans d'autres contextes physiques ou mathématiques ? 	<ol style="list-style-type: none"> 1. Le contexte proposé (cyclisme, performance, récupération) est-il pertinent et mobilisateur pour les élèves visés ? 2. Peut-on structurer cette situation comme une activité dirigée, une situation-problème ou un projet de modélisation ? 3. Quels obstacles conceptuels ou erreurs fréquentes peut-on anticiper (ex. : confusion entre vitesse et moyenne, entre temps et distance) ? 4. Quels appuis ou ressources (graphique, tableau, simulation, calculatrice, IA) seraient pertinents pour soutenir la résolution ? 5. Comment articuler cette situation à d'autres contenus du programme (fonction, équation, proportion, etc.) ? 6. La situation permet-elle un travail différencié selon les profils ou les niveaux des élèves ? 7. Comment favoriser la mise en commun des stratégies dans une optique d'apprentissage collectif ? 8. Peut-on introduire des variations paramétriques (vitesse cible, distance, pente) pour enrichir l'analyse ou favoriser la généralisation ? 9. Cette situation a-t-elle une valeur sociale, épistémique ou formative qui mérite d'être explicitée auprès des élèves ? 10. Peut-elle s'intégrer à une approche interdisciplinaire (liens avec l'éducation physique, la physique, la biologie ou même la musique) ? 	<ol style="list-style-type: none"> 1. L'IA (ex. ChatGPT) est-elle capable de repérer spontanément la contradiction dans l'usage d'une moyenne arithmétique dans ce contexte ? 2. Sait-elle mobiliser la moyenne harmonique et en justifier l'usage dans une situation concrète ? 3. Peut-elle expliquer clairement la différence entre vitesse instantanée, vitesse moyenne, et moyenne pondérée ? 4. Est-elle capable de proposer des représentations variées du problème (formule, graphique, tableau, image) ? 5. L'IA peut-elle simuler un raisonnement erroné, dans une visée d'enseignement par l'erreur ? 6. Quelle est la robustesse de ses réponses si l'on modifie légèrement les données (ex. : vitesses, distances, pente, objectifs) ? 7. Peut-elle accompagner une démarche d'investigation en reformulant le problème, en générant des variantes ou en posant des contre-questions ? 8. Dans quelle mesure l'IA transforme-t-elle le rapport de l'élève (ou de l'enseignant) au problème mathématique ? 9. L'IA pourrait-elle être mobilisée non pour résoudre le problème, mais pour proposer des analogies (ex. musique, économie, écologie) favorisant l'appropriation conceptuelle ? 10. Quels indicateurs permettraient d'évaluer la qualité des interactions humain-IA : pertinence des relances, cohérence des justifications, clarté des représentations, etc. ?

Ces questions proviennent d'un échange exploratoire avec ChatGPT, amorcé à partir du problème du cycliste et progressivement réorienté selon différentes intentions mathématiques, didactiques et instrumentales. Elles sont souvent pertinentes, mais mériteraient d'être (ré)organisées en fonction d'une intention ou d'un projet didactique précis. Ce n'est toutefois pas notre objectif ici, puisque nous ne cherchons pas à proposer des activités toutes faites. On remarque néanmoins que cet outil offre plusieurs pistes prometteuses, ne serait-ce que pour tenter de faire parler le contexte. Par exemple : s'il existe un seuil de vitesse moyenne hors d'atteinte avec une montée à 10 km/h, quelle est en la valeur limite ? Peut-on la visualiser ? En revanche, d'autres questions pourtant bien naturelles restent en plan, par exemple : si l'on monte à 10 km/h et que l'on descend à 50 km/h, qu'indiquerait le GPS au bout de la

descente ? C'est comme si ChatGPT agissait en bon assistant, parfois même plutôt doué, mais sans jamais pouvoir dépasser l'humain qui reste maître du questionnement, acceptant volontiers qu'il ne puisse pas penser à tout.

Pour approfondir certaines de ces pistes, nous avons formulé trois requêtes ciblées. Celles-ci visent à documenter la capacité du système à effectuer des calculs, à produire des représentations graphiques et à raisonner sur des cas extrêmes, dans une logique exploratoire visant à enrichir la problématisation :

R1. Répondre à cette question :

« Si l'on monte à 10 km/h et que l'on descend à 50 km/h, qu'indiquerait le GPS au bout de la descente ? »

R2. Fournir une représentation graphique de la vitesse moyenne que l'on atteint lorsque l'on grimpe une côte de 1 km à 10 km/h en fonction de la vitesse en descente sur une côte équivalente de 1 km.

R3. Dans le contexte du problème, si nous étions capables de rouler à la vitesse de la lumière en descente, pourrions-nous atteindre 30 km/h malgré tout ? Pourquoi ?

Pour **R1**, après avoir dévoilé un calcul étape par étape, le milieu indique une moyenne affichée de 16,67 km/h, ce qui correspond à une bonne valeur approchée de $50/3$ km/h. Ce résultat, bien que mathématiquement juste, peut surprendre tout cycliste sur le plan intuitif : nombreux sont ceux qui s'attendraient à une moyenne plus élevée, possiblement proche de 30 km/h. La déception n'est donc pas d'ordre mathématique, mais bien physique : elle résulte du temps disproportionné passé à grimper. En réalité, la seule façon d'augmenter la vitesse moyenne dans un tel scénario consiste d'abord à monter plus vite — ce qui suppose évidemment un effort accru —, ou à prolonger la phase de roulage à vitesse élevée après les 2 km du parcours.

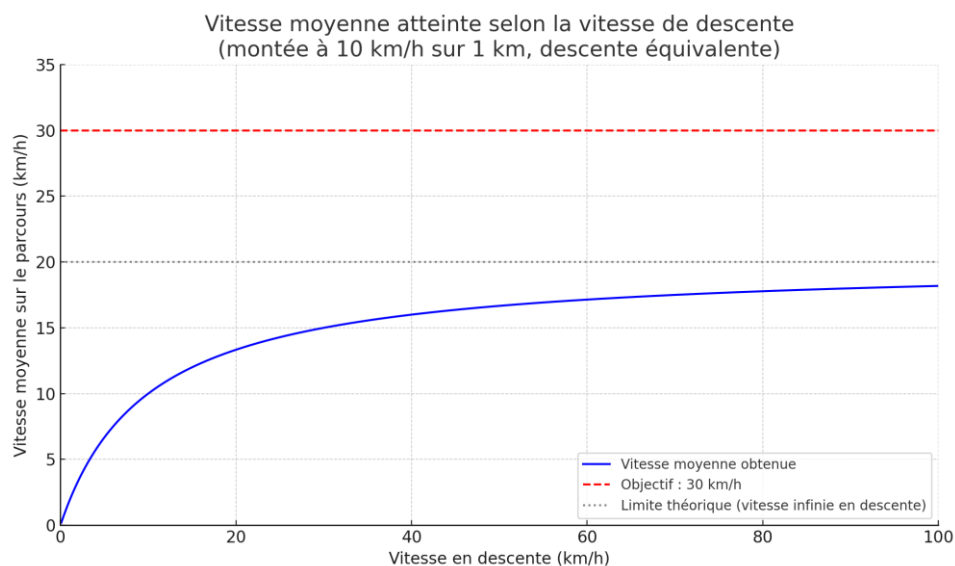


Figure 6. Représentation graphique produite par ChatGPT (Python) de la vitesse moyenne en fonction de la vitesse de descente.

Pour **R2**, il retourne une première représentation graphique qui n'était pas mal, sauf qu'il avait initialement choisi une échelle de 0 à 1 000 km/h pour l'axe horizontal — plutôt absurde dans un contexte cycliste. Après notre demande de correction, il a rapidement produit le graphique de la Fig. 6. Celui-ci met en évidence la relation entre la vitesse de descente et la vitesse moyenne obtenue sur le parcours, sous la forme d'une courbe asymptotique qui tend vers une limite de 20 km/h lorsque la vitesse en descente devient très grande. Autrement dit, même dans le cas d'une descente quasi instantanée, la moyenne reste contrainte par le temps de montée, qui demeure incompressible. Le simple fait d'augmenter fortement la vitesse en descente n'améliore donc que marginalement la moyenne globale, ce qui met en lumière une illusion d'optimisation fréquemment observée. L'écart entre l'objectif de 30 km/h et cette borne théorique constitue ainsi un indicateur puissant du piège conceptuel induit par une intuition arithmétique immédiate, avant toute acculturation à ce type de situation.

Pour **R3**, il répond par un « non » bien senti, en expliquant que, même si l'on descendait à la vitesse de la lumière (environ 1 080 000 000 km/h), le temps de descente tendrait vers zéro, mais le temps de montée resterait à 0,1 heure. Il ajoute :

La vitesse moyenne serait alors :

$$v_{\text{moy}} = \frac{2}{0,1 + 0} = 20 \text{ km/h},$$

ce qui confirme que 20 km/h est une borne asymptotique qu'on ne peut dépasser dans cette configuration.

Ce « 0 » peut étonner, on dirait même qu'il encapsule un 0^+ , ce qui paraît se confirmer lorsqu'on insiste en demandant « pourquoi 20 km/h ? », et qu'il nous répond « quand la vitesse de descente $v \rightarrow \infty$, alors $\frac{1}{v} \rightarrow 0$ ». Finalement, il conclut que la vitesse moyenne maximale théorique est 20 km/h, quel que soit le gain en descente, qu'il est donc mathématiquement impossible d'atteindre une moyenne de 30 km/h, même à la vitesse de la lumière, et que c'est ce qui rend le problème si riche tant du point de vue conceptuel et didactique.

Lorsqu'on consulte le code Python qu'il a employé pour générer le graphique, on constate qu'il mobilise deux bibliothèques fondamentales : *numpy*, pour les calculs numériques (comme la génération d'une gamme de vitesses en descente), et *matplotlib.pyplot*, pour la représentation graphique elle-même. On remarque également que la plage des vitesses de descente commence à 0,1 km/h plutôt qu'à 0, ce qui évite une division par zéro dans le traitement des cas extrêmes.

Cet ensemble d'interactions montre comment l'artéfact numérique permet non seulement de simuler des comportements souhaités ou attendus, mais aussi de provoquer des tensions heuristiques utiles à la construction de nouveaux problèmes. On voit ainsi se dessiner, en filigrane, un déplacement progressif des intuitions initiales vers des formes de modélisation plus rigoureuses, mobilisant à la fois des représentations graphiques, des raisonnements sur les limites et des interprétations fines du comportement informatique. Ce déplacement s'opère dans un environnement où le milieu instrumenté permet de générer rapidement des réponses, mais aussi — et peut-être surtout — d'inviter à les questionner, les ajuster, voire les détourner

à des fins didactiques. C'est à cette articulation entre formulation, figuration et représentation que nous nous intéressons à présent, en examinant comment ChatGPT contribue à structurer un double registre d'images et de signes dans le travail mathématique, non sans difficulté.

Distinction entre illustration et représentation sémiotique avec l'IA

En début de texte, nous avons évoqué brièvement les distinctions entre IA symbolique et IA générative, sans développer le rôle des approches dites « floues », comme les réseaux neuronaux, la logique floue ou les algorithmes évolutionnistes. Celles-ci permettent pourtant de traiter des situations imprécises ou non déterministes. Elles proposent une autre manière d'aborder la modélisation, que ce soit pour rendre compte d'une réalité à travers la mathématisation d'une situation modèle, ou pour représenter des connaissances à l'aide de systèmes de signes formels. L'objectif n'est pas nécessairement de reproduire la structure logique du raisonnement humain, mais plutôt d'en simuler certains comportements, ou de générer des réponses plausibles dans des environnements multiregistrés, en jouant sur une forme d'élasticité sémantique propre à l'activité humaine.

Cela ouvre la voie à une analyse plus fine du contrepoint technologique, des tensions, mais aussi des synergies possibles entre des régimes d'intelligibilité hétérogènes : l'un symbolique, basé sur des enchaînements formels contrôlés ; l'autre génératif, fondé sur des approximations statistiques pilotées par les données. C'est précisément dans la construction d'images et de signes que cette tension devient très visible, comme le montrent les GML par rapport aux traitements symboliques. Les productions d'un outil informatique, qu'il s'agisse d'un graphique, d'un schéma ou d'une figure, oscillent entre l'illustration intuitive, plus proche d'un effet visuel, et la représentation sémiotique, c'est-à-dire une configuration de signes mobilisables dans un raisonnement mathématique. Cette distinction est cruciale dans l'enseignement des mathématiques, où une image n'est pas nécessairement un support d'énonciation ou de preuve, mais peut au contraire détourner l'attention des propriétés pertinentes, faute de coordination avec le discours et le contexte de la situation.

La Figure 6, produite par ChatGPT avec le soutien de Python et de ses bibliothèques de calcul et de visualisation, incarne bien cette tension. D'un côté, elle représente graphiquement un lien entre deux grandeurs (vitesse de descente et vitesse moyenne), mobilisant des conventions symboliques bien établies (axes, unités, asymptotes, etc.) : il s'agit là d'une représentation sémiotique rigoureusement interprétable. De l'autre, sa courbure asymptotique agit aussi comme une illustration intuitive du phénomène : plus la descente est rapide, plus la moyenne semble plafonner — intuition que même un non-mathématicien peut saisir. Cependant, cette illustration a été produite à l'aide d'un langage informatique, sur la base d'information textuelle (dynamique de traduction), pas mathématique. Lorsque la représentation est produite dans l'environnement Wolfram, son statut devient celui d'un système de signes reconnus, à la manière des représentations figurales, graphiques, en arbre ou en diagramme (conversion entre registres). C'est cette bifacialité sémiotique que nous souhaitons souligner : une même représentation peut être conçue ou interprétée comme une illustration, un système de signes, ou les deux, selon l'intention de lecture. Ce point de bascule est didactiquement fécond, à condition de savoir où il se situe précisément dans la tâche, ou plus globalement dans le contrat didactique. Il montre aussi que l'outil informatique,

dans un environnement de travail mathématique, ne se contente pas de représenter un raisonnement déjà fait : il participe à sa formulation, en coproduisant des objets graphiques ou textuels susceptibles d’orienter la pensée... ou de la confondre.

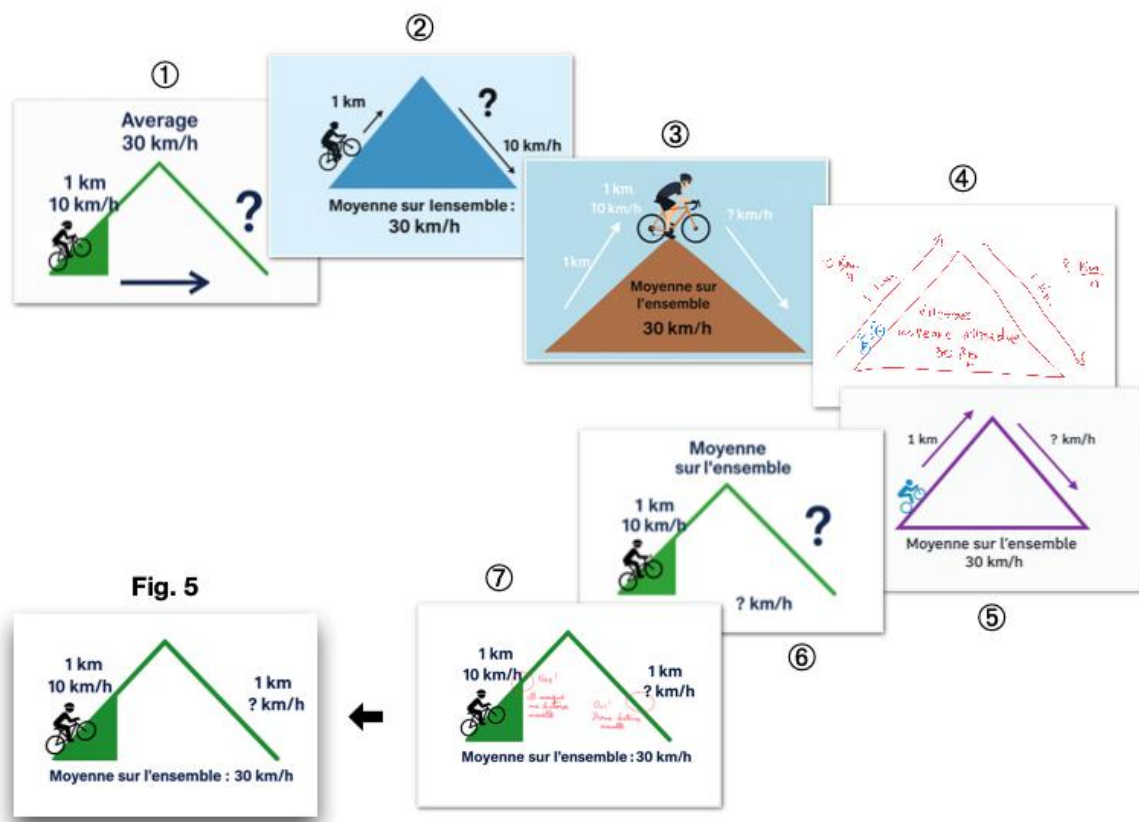


Figure 7. Tentatives de production d’une figure à l’aide de ChatGPT seul.

La Figure 7 rend compte des difficultés rencontrées lorsqu’on cherche à produire une représentation graphique à l’aide de ChatGPT seul. Malgré des requêtes répétées en langage naturel, et en dépit d’éléments visuellement pertinents (profils, flèches, annotations), les différentes tentatives ne parviennent pas à articuler de manière cohérente l’espace, les données numériques et la logique mathématique sous-jacente. Cette situation a nécessité de multiples itérations, l’ouverture de conversations parallèles et le recours à des moyens externes — notamment des annotations manuelles ou des outils graphiques — pour obtenir un visuel exploitable. Elle met en évidence une limite importante des GML dans la coordination des registres figuratif et mathématique, même lorsque la situation est bien « dévolue » et conceptuellement accessible.

C’est ici que la perspective de Sangalli (2001) devient éclairante. L’auteur souligne que l’introduction du « flou » dans les raisonnements mathématiques ne constitue pas une perte de rigueur, mais une adaptation à la complexité du réel. Cette idée fait écho à la tension que nous observons ici entre cohérence formelle et plausibilité discursive dans les systèmes à cheval entre différentes approches d’IA. Les représentations issues de l’IA générative — parfois instables, mais expressives — viennent perturber le cadre sémiotique rigide des approches symboliques, sans pour autant le disqualifier. C’est cette cohabitation, parfois conflictuelle, qui mérite attention dans une perspective didactique, et qui est très exigeante pour l’usager.

Enfin, il serait erroné de considérer les mathématiques simplement comme une discipline dans un contexte où, selon le niveau d'analyse, les connaissances mobilisées ne relèvent pas exclusivement de l'informatique. Les mathématiques ne se contentent pas de parler des objets modélisés par l'informatique : elles en constituent les fondements mêmes. C'est par les structures mathématiques que les modèles sont construits, ajustés, validés — et donc pensés. Lorsqu'un élève ou un chercheur interagit avec une IA pour explorer un problème, ce sont les mathématiques qui rendent cette interaction possible, intelligible et féconde. Il devient donc crucial de comprendre les rôles respectifs que jouent l'illustration et la représentation dans ces échanges, et ce que cela engage sur le plan de l'apprentissage et du raisonnement.

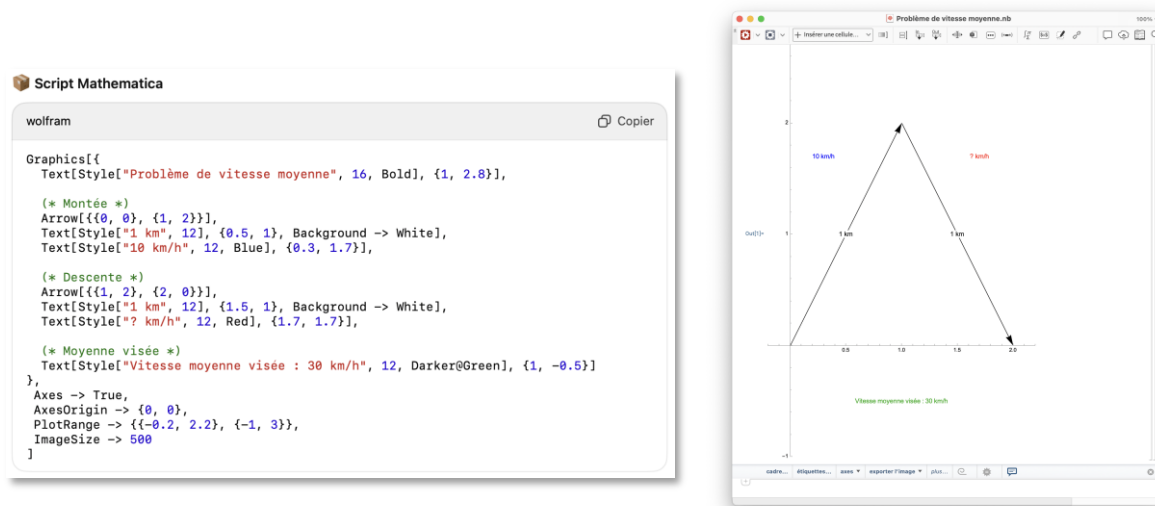


Figure 8. Représentation produite dans l'environnement ChatGPT + Wolfram.

Nous avons d'ailleurs éprouvé cette tension de manière très concrète lors de la production des figures 7 et 8. Pour générer une figure synthétique comme la figure 5, il a fallu près de vingt itérations avec ChatGPT seul, sans jamais obtenir une représentation mathématiquement exploitable. Les requêtes textuelles se sont multipliées, des branches parallèles de conversation ont été ouvertes, des moyens hors système mobilisés (comme l'écriture sur tablette numérique), et l'alignement spatial entre données et éléments figuratifs a exigé un ajustement patient — qui frisait l'insolence récursive. Pourtant, l'énoncé était clair, structuré, parfaitement interprétable par un humain, et apparemment aussi à la lecture des éléments discursifs du milieu. Mais le passage à la représentation visuelle résistait, en raison d'une coordination défailante entre les modules graphiques et textuels.

Le basculement vers l'environnement ChatGPT + Wolfram (Fig. 8) modifie profondément le statut des productions obtenues. En générant un script en Wolfram Language, l'IA ne produit plus seulement une image, mais une structure formelle interprétable, dans laquelle chaque élément graphique correspond à un objet mathématique défini selon des conventions explicites. La distinction entre code et figure devient alors opératoire : le premier constitue une écriture formelle manipulable, tandis que la seconde en est une instantiation visuelle contrôlée. Ce passage vers un système sémiotique programmable renforce la stabilité des représentations, réduit le nombre d'itérations nécessaires et permet une conformité accrue aux normes du travail mathématique. Il marque ainsi un déplacement du registre illustratif vers un registre de représentation au sens fort.

3.3 Écrire avec l'IA : une coélaboration sous contrôle humain

Cet exemple, plus autonome dans sa forme, a pourtant été coécrit avec ChatGPT à partir d'un questionnement personnel et d'un processus d'idonéité où c'est l'humain qui décide de la forme finale. Il illustre non seulement un prolongement naturel du travail mathématique, mais aussi une démarche instrumentée d'écriture, dans laquelle l'IA assiste l'élaboration progressive d'un texte explicatif à visée didactique :

En **patinage de vitesse sur courte piste**, on trouve des épreuves sur des distances comme 333 m, 777 m ou encore 1000 m. Pourquoi ces valeurs apparemment irrégulières ? Voyons ce qu'il en ressort du point de vue mathématique.

1. Où placer la caméra de photo-finish pour une épreuve de 1000 m ?

Chaque tour de piste mesure environ 111,12 m, selon le règlement de l'ISU (International Skating Union). Une course de 1000 m correspond alors à exactement 9 tours : $9 \times 111,12 = 1000,08$ m. On en conclut que l'arrivée coïncide, à très peu de chose près, avec la ligne de départ. La caméra de photo-finish est donc placée au point de départ, qui fait également office de point d'arrivée.

2. Est-ce qu'on termine vraiment au point de départ ? Ou à un epsilon près ?

D'un point de vue mathématique, $1000,08 \text{ m} \neq 1000 \text{ m}$. Il y a une différence de 8 centimètres, soit un epsilon de 0,08. Faut-il déplacer la ligne d'arrivée pour respecter cette précision ? La réponse est non : en pratique, cette variation est négligeable. Elle se situe bien en deçà des tolérances permises en compétition et n'altère en rien la validité de la mesure ni l'équité entre les concurrents.

3. Que dit le règlement officiel de l'ISU à ce sujet ?

L'ISU fixe les distances en nombre de tours, et non en mètres exacts. Ce sont donc les tours qui sont normés, non les distances métriques. Par exemple :

- La course de 1000 m correspond à 9 tours, soit environ 1000,08 m ;
- La ligne de départ = ligne d'arrivée ;
- Aucune correction n'est apportée pour combler ou ajuster ces petits écarts.

Cette normalisation illustre bien la tension entre exactitude mathématique et cohérence institutionnelle. Le calcul est rigoureux, mais il n'impose pas à lui seul la norme. Il faut aussi considérer les usages, les conventions et les limites matérielles, autant de dimensions qui participent à la problématisation du réel à l'école.

Ce texte, en apparence tout à fait humain, est en réalité issu d'un processus dialogué avec l'IA. Il s'appuie sur un script initial que nous avons formulé en langage naturel, puis réorganisé, enrichi et affiné en interaction avec ChatGPT. Il constitue ainsi un exemple d'écriture collaborative humain-IA, au service d'un raisonnement mathématique contextualisé, mobilisant à la fois des données précises, des références réglementaires et des formes argumentatives rigoureuses. Il aurait pu prendre une autre forme, par exemple, en jouant sur l'idée que 9 tours de 111,12 m illustrent une situation concrète où $0,999... = 1$. En

effet, chaque tour de piste mesure exactement $\frac{577}{10} + 17\pi$ m, soit 111,11 m à 0,01 près, tandis que la valeur réglementaire retenue est de 111,12 m³. Pourtant, c'est bien l'humain qui doit prendre les décisions pour stabiliser son propos, choisir un cadre interprétatif, et déterminer les contenus qu'il juge pertinents de communiquer, à l'instar du principe d'intelligence augmentée.

Ce type de production, que nous pourrions appeler une illustration raisonnée, marque un point d'inflexion dans les usages didactiques de l'IA. L'image ne se contente plus d'illustrer : elle soutient activement la construction discursive d'un raisonnement structuré. Dans cette économie de coconstruction du sens, c'est l'humain qui mobilise les GML — malgré leurs défauts — en y injectant ses propres intentions, tout en assumant la charge supplémentaire de validation des sources, de recherche documentaire, et de vigilance pour maintenir la cohérence de l'ensemble.

3.4 De l'exemple à la synthèse : idonéité, vigilance et contrepoint technologique

Dans cette troisième section, nous avons exploré des facettes complémentaires du travail mathématique instrumenté par l'intelligence artificielle. La première portait sur une situation inspirée du jeu d'échecs, où ChatGPT (seul, puis augmenté de Wolfram) contribue à l'exploration d'un problème structuré autour du déplacement d'un cavalier sur une grille musicale. Nous avons montré comment l'interaction avec le milieu instrumenté permet non seulement de générer des solutions, mais aussi de documenter les limites heuristiques, de tester des cas extrêmes et d'interroger l'idonéité dans une dynamique articulant l'humain et la machine.

La seconde facette portait sur la modélisation d'un phénomène cycliste (vitesse moyenne sur un parcours en montée et descente), et a permis d'aborder les tensions entre réalité, intuition arithmétique et rigueur mathématique. Ce cas a mis en lumière la capacité de l'artéfact numérique à produire des calculs, des représentations graphiques et des raisonnements sur les asymptotes — tout en révélant le rôle central de l'humain dans la formulation des questions pertinentes et la validation des réponses.

Enfin, un troisième exemple a permis de montrer comment une situation réelle, issue d'un domaine plus méconnu (le patinage de vitesse sur courte piste), pouvait donner lieu à une écriture mathématique collaborative humain-IA, exploitant ce qui fait sans doute la force des GML. Ce texte, bien que finalisé dans une langue fluide et rigoureuse, est issu d'un processus d'interaction entre un raisonnement humain initial, des données techniques, des formulations approximatives, et un travail de réécriture instrumentée avec l'outil informatique et des vérifications documentaires.

Ces exemples convergent vers une même idée : le dialogue avec ces types de milieu instrumenté ne consiste pas à déléguer la pensée, mais à nourrir des processus de questionnement et de problématisation, dont l'initiative et la décision finale appartiennent à

³ Curiosité topologique : la patinoire de hockey typique prend globalement la forme d'un rectangle aux coins arrondis, tandis que la piste de patinage de vitesse courte piste correspond à une trajectoire normée formée de deux segments rectilignes reliés par deux demi-cercles.

l'utilisateur, de manière caractéristique et non accessoire. Ce processus est fécond pour qui dispose d'un projet clair, d'idées peut-être encore incomplètes, mais déjà contextualisées et situées par rapport à un référentiel dans lequel l'utilisateur a construit des connaissances et développé des compétences avérées. Ce que produit l'outil informatique (texte, schéma, tableau ou code) devient alors une matière à interroger, à interpréter, à prolonger ou à détourner, pour celui qui cherche à penser autrement. Cela suppose une vigilance constante sur les glissements de registre, mais aussi une capacité à coordonner différents systèmes de signes (langagiers, symboliques, graphiques) dans une visée de structuration du sens.

Ainsi, l'outil informatique ne remplace ni le problème, ni le raisonnement, ni la situation. Même s'il permet de représenter certaines réalités à travers le discours et des moyens sémiotiques interactifs, il agit surtout comme un levier pour en déplacer les frontières. Il introduit de nouvelles formes de médiation, qui peuvent faire gagner du temps, décaler le regard, rendre visibles certaines structures... ou au contraire révéler des impasses, et se montrer cruellement chronophage. Ce potentiel didactique repose toutefois sur un travail humain de recontextualisation, d'analyse et de validation — un travail qui exige une étude fine des effets de ces instruments sur l'organisation possible du jeu didactique en classe. Ce travail est d'autant plus exigeant qu'il s'inscrit dans un contexte de renouvellement technologique constant, où les cadres culturels véhiculés par les outils numériques sont souvent implicites, multiregistres et difficiles à maîtriser, même pour des usagers expérimentés, et où les référentiels portés par l'interaction avec les modèles génératifs sont instables, opaques et généralement inaccessibles à la recherche didactique classique.

Les situations présentées mettent en évidence que l'usage des systèmes d'intelligence artificielle dans le travail mathématique ne peut être réduit à une simple assistance à la résolution. Elles révèlent une diversité de régimes d'engagement, allant de la délégation partielle du traitement à une exploration réflexive des productions, dans laquelle l'IA agit comme un contrepoint technologique. Dans cette perspective, la qualité du travail mathématique ne dépend pas seulement des performances du système, mais de l'idonéité du processus engagé, c'est-à-dire de la capacité du sujet à articuler de manière cohérente un problème, des outils et une situation. C'est cette dynamique, déjà analysée sur le plan théorique (Richard, 2026), que les exemples permettent ici de rendre visible, en montrant comment les interactions humain-IA reconfigurent les conditions de production, de contrôle et de validation des savoirs mathématiques.

Conclusion - Délégation, coopération et écologie du travail mathématique instrumenté

L'introduction des grands modèles de langage dans le travail mathématique ne se réduit ni à une amélioration des performances ni à une simple extension des outils disponibles. Elle engage une transformation plus profonde, qui touche à la fois les conditions d'élaboration des problèmes, les formes de raisonnement, les modalités de validation et les régimes d'engagement du sujet. Ce déplacement ne relève ni d'une substitution de l'humain par la machine, ni d'une coopération symétrique entre partenaires équivalents, mais d'une reconfiguration des rapports entre production, contrôle et signification.

L'analyse développée montre que les systèmes d'intelligence artificielle — en particulier les modèles génératifs — introduisent une dissociation inédite entre cohérence discursive et nécessité logique. Cette dissociation constitue à la fois une ressource et un risque : elle ouvre l'espace des possibles, stimule l'exploration et soutient certaines formes de reformulation, tout en fragilisant les conditions de validation, en exposant le sujet à des productions plausibles dépourvues de fondement épistémique explicite. Dans ce contexte, la question n'est pas de savoir si l'on peut déléguer certaines tâches à la machine, mais de déterminer sous quelles conditions cette délégation demeure compatible avec le travail mathématique.

Les situations étudiées permettent de préciser ces conditions. Elles montrent que les formes de délégation souhaitables ne concernent pas le cœur du raisonnement — c'est-à-dire la construction du sens, la structuration des relations et le contrôle de la validité —, mais certaines opérations périphériques ou intermédiaires : calculs, explorations locales, visualisations, génération de cas ou de formulations. Dès lors que la machine intervient au niveau de la production de solutions ou de la validation implicite, le risque de perte de contrôle épistémique devient significatif. À l'inverse, lorsque ses productions sont mobilisées comme objets à interroger, à comparer ou à transformer, elles peuvent contribuer à enrichir le travail mathématique, en introduisant des écarts féconds. C'est dans cette perspective que la notion de contrepoint technologique prend tout son sens. Elle permet de penser l'interaction humain-IA non comme une fusion ou une substitution, mais comme une mise en tension entre deux régimes de production hétérogènes : l'un fondé sur la nécessité, l'autre sur la plausibilité. Loin d'être un obstacle, cet écart peut devenir un levier pour structurer l'activité, à condition d'être reconnu, maîtrisé et didactiquement exploité. Le rôle de l'enseignant apparaît alors central, non pour contrôler l'outil en tant que tel, mais pour organiser les conditions dans lesquelles cette tension devient intelligible et productive pour les élèves.

Dans ce cadre, l'intelligence augmentée ne peut être envisagée comme un état résultant d'une simple hybridation homme-machine. Elle doit être comprise comme un processus d'interaction instrumentée, dans lequel le sujet conserve la responsabilité du questionnement, du raisonnement et de la validation. L'augmentation est ainsi conditionnelle à une compétence critique minimale, permettant de distinguer entre validité, plausibilité et crédibilité, et de réintroduire des référentiels là où les systèmes génératifs tendent à les effacer. Ce constat conduit à envisager le travail mathématique instrumenté, en contexte scolaire comme plus largement dans les pratiques contemporaines de production mathématique, comme une écologie au sens d'un système d'interactions entre sujets, outils, situations et institutions, traversé par des contraintes multiples — cognitives, épistémiques, techniques et sociales. Dans cette écologie, l'idonéité ne constitue pas une propriété donnée, mais une dynamique à construire, fondée sur des ajustements continus entre les visées du sujet et les réponses du milieu. L'enjeu n'est pas d'optimiser les performances ni de stabiliser des usages, mais de préserver les conditions d'un engagement mathématique authentique, capable de résister aux effets d'automatisation et de soutenir la construction du sens.

Ainsi, la question des délégations souhaitables ne peut être tranchée de manière générale ou prescriptive. Elle dépend des tâches, des situations et des niveaux d'expertise, mais elle suppose toujours un principe invariant : ce qui relève du contrôle épistémique ne saurait être

délégué sans transformation du travail mathématique lui-même. À cette condition, les systèmes d'intelligence artificielle peuvent devenir non pas des substituts du raisonnement, mais des instruments de sa mise à l'épreuve — et, peut-être, de son renouvellement. Les analyses proposées dans cet article prennent pleinement sens en regard des développements théoriques présentés dans la version complémentaire (Richard, 2026), dont elles constituent le prolongement empirique. Ensemble, ces deux volets articulent exigence théorique et intelligibilité des pratiques, dans une perspective renouvelée du travail mathématique à l'ère de l'intelligence artificielle.

Références

- Balacheff, N. (1994). *Didactique et intelligence artificielle*. Grenoble : La Pensée Sauvage.
- Balacheff, N., & Margolinas, C. (2005). cK ϕ : Modèle de connaissances pour le calcul de situations didactiques. In A. Mercier & C. Margolinas (Eds.), *Balises pour la didactique des mathématiques* (pp. 75–106). La Pensée Sauvage.
- Brousseau, G. (1998). *Théorie des situations didactiques*. Grenoble : La Pensée Sauvage.
- Bruillard, É., & Richard, P. R. (2024). Informatique, mathématiques, conception et usage des technologies numériques. *Annales de Didactique et de Sciences Cognitives, Thématique 2*, 173–208. DOI: 10.4000/11sga
- Chevallard, Y. (1999). L'analyse des pratiques enseignantes en théorie anthropologique du didactique. *Recherches en didactique des mathématiques*, 19(2), 221-265. Grenoble : La Pensée Sauvage.
- CNRS Images des mathématiques. (2025, octobre). Premier défi : le problème du mois. Images des Mathématiques. <https://images.math.cnrs.fr/billets/defis-doctobre-2025/>
- Danguy-Pichette, É., & Richard, P. R. (2026). À la recherche d'un référentiel : entre structuration, transformation et contrepoint technologique. *Annales de Didactique et de Sciences Cognitives*, numéro thématique à venir.
- Duval, R. (1995). *Sémiosis et pensée humaine : registre sémiotique et apprentissages intellectuels*. Peter Lang.
- Emprin, F., & Richard, P. R. (2023). Intelligence artificielle et didactique des mathématiques : état des lieux et questionnements. *Annales de Didactique et de Sciences Cognitives*, 28, 131–181. DOI: 10.4000/adsc.2678
- Engelbart, D. C. (1962). *Augmenting Human Intellect: A Conceptual Framework. Summary Report, Stanford Research Institute, on Contract AF 49(638)-1024*, October 1962, 134 pages. Available at: <https://www.doungelbart.org/pubs/augment-3906.html>
- Gonseth, F. (2022). *La géométrie et le problème de l'espace* (Réédition en un volume des ouvrages publiés entre 1945 et 1955). St-Imier, Switzerland : Association F. Gonseth.
- Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2023). *A survey on hallucination in large language models*:


- Principles, taxonomy, challenges, and open questions.* arXiv. <https://arxiv.org/abs/2311.05232>
- Joshua, S., & Dupin, J.-J. (1999). *Introduction à la didactique des sciences et des mathématiques*. Presses universitaires de France.
- Kuzniak, A., Montoya-Delgadillo, E., & Richard, P. R. (Eds.). (2022). *Mathematical Work in Educational Context. The Perspective of the Theory of Mathematical Working Spaces. Mathematics Education in the Digital Era*, Vol. 18. Springer. DOI: 10.1007/978-3-030-90850-8
- Li, J., Cao, P., Chen, Y., Liu, K., & Zhao, J. (2024). *Towards faithful chain-of-thought: Large language models are bridging reasoners*. arXiv. <https://arxiv.org/abs/2405.18915>
- Menuet, T. (2024, 2 octobre). *Mathématiques et musique : une même seconde peau. Images des Mathématiques*. <https://images.math.cnrs.fr/billets/mathematiques-et-musique-une-meme-seconde-peau>
- Richard, P. R. (2026). Le travail mathématique à l'interface des grands modèles de langage : entre assistance, coopération et contrepoint technologique, *Annales de didactique et de sciences cognitives*, 31. A paraître.
- Richard, P. R. (2024). The challenges of AI in shaping mathematical work: From human hybridization to automation through synergies of symbolic AI and generative models. In K. W. Kosko, J. Caniglia, S. A. Courtney, M. Zolfaghari, & G. A. Morris (Eds.), *Proceedings of the Forty-Sixth Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 2213-2226). Kent State University.
- Richard, P. R. & Van Vaerenbergh, S. (2026). L'intelligence artificielle et le nouveau travail mathématique. In S. Van Vaerenbergh, L. Vivier, F. Arzarello, J. V. Flores Salazar, J. Gaona Paredes, P. Gibel, I. M. Gómez-Chacón, A. Kuzniak, M. Maschietto, E. Montoya Delgadillo, A. Nechache, K. Nikolantonakis, R. E. Páez Murillo, P. R. Richard, M. P. Vélez, & F. Venant (Eds.), *Actas del Octavo Simposio sobre el Estudio del Trabajo Matemático (ETM8)* (pp. 25–62). Editorial de la Universidad de Cantabria.
- Richard, P. R., Vélez, M. P., & Van Vaerenbergh, S. (Eds.). (2022). *Mathematics Education in the Age of Artificial Intelligence: How Artificial Intelligence Can Serve Mathematical Human Learning*. Mathematics Education in the Digital Era, Vol. 17. Cham: Springer. DOI: 10.1007/978-3-030-86909-0
- Richard, P. R., Venant, F., & Gagnon, M. (2019). Issues and challenges in instrumental proof. In G. Hanna, D. Reid, & M. de Villiers (Eds.), *Proof technology in mathematics research and teaching* (Vol. 14). Springer Cham. https://doi.org/10.1007/978-3-030-28483-1_7
- Sangalli, A. (2001). *Éloge du flou : Aux frontières des mathématiques et de l'intelligence artificielle*. Montréal : Presses de l'Université de Montréal.

Annexe 1. Interaction avec ChatGPT seul : plausibilité discursive et instabilité du contrôle

Cette annexe présente trois moments d'une interaction avec ChatGPT dans le cadre de la résolution du problème *Échecs, maths et musique*. Les extraits retenus visent à illustrer certaines tensions caractéristiques du travail mathématique instrumenté, notamment entre production discursive et validation.

A1. Reformulation et première solution

Merci pour l'image actualisée du défi du mois. Il s'agit de déplacer un cavalier (pièce d'échecs) sur une grille de 3 lignes \times 4 colonnes, selon les règles classiques du déplacement du cavalier (en "L"), afin de produire une séquence de notes la plus longue possible sans passer deux fois par la même case.


 Rappel – Déplacement d'un cavalier aux échecs :

Le cavalier se déplace :

- soit de 2 cases dans une direction, puis 1 case perpendiculaire,
- soit de 1 case dans une direction, puis 2 cases perpendiculaires.

Donc :

- Depuis une case, le cavalier a au maximum 8 déplacements possibles,
- Mais ici, comme la grille est de 3 lignes \times 4 colonnes, certaines options seront hors limites.

 Objectif :


Trouver un chemin le plus long possible en partant d'une case quelconque, sans répéter une case, en respectant les règles de déplacement du cavalier.

 Solution trouvée :

Voici une séquence valide de 12 déplacements (le maximum possible pour une grille de 3 \times 4 sans revenir sur une case, car la grille contient 12 cases) :

 Solution (chemin de 12 notes) :

1. do
2. fa
3. la
4. ré
5. sol#
6. fa#
7. ré#
8. do#
9. mi
10. sol
11. la#
12. fa

 Cette solution passe exactement une fois par chaque case, sans répétition, en suivant uniquement les mouvements autorisés du cavalier.

Le modèle reformule correctement le problème (grille 3 \times 4, déplacements du cavalier) et propose immédiatement une solution de longueur maximale, qu'il affirme valide. La prise en charge discursive est correcte, mais la solution contient une erreur : la note fa apparaît deux fois dans le parcours, contrairement à ce qu'affirme le système. On observe ici un effet de validité déclarée, non fondée sur un contrôle effectif des contraintes.

A2. Vérification et corrections instables

📌 Coordonnées associées (pour vérification)

On peut numéroter les cases selon leur position sur une grille 3x4 :
(on les lit de gauche à droite et de haut en bas)

	C0	C1	C2	C3
L0	do	do#	ré	
L1	ré#	mi	fa	fa#
L2	sol	sol#	la	la#

Position dans le tableau : (ligne, colonne)

- do → (0,0)
- fa → (1,2)
- la → (2,2)
- ré → (0,2)
- sol# → (2,1)
- fa# → (1,3)
- ré# → (1,0)
- do# → (0,1)
- mi → (1,1)
- sol → (2,0)
- la# → (2,3)
- fa → (1,2) ❌ (déjà visité) → donc on termine à la#

Correction : fa apparaît deux fois. Donc la séquence doit être revue pour ne pas répéter.

✅ Nouvelle solution correcte (révisée – 12 notes uniques) :

Après test, voici une vraie solution correcte avec les 12 notes distinctes :

1. ré# (1,0)
2. la# (2,3)
3. do# (0,1)
4. fa# (1,3)
5. ré (0,2)
6. sol# (2,1)
7. do (0,0)
8. la (2,2)
9. fa (1,2)
10. sol (2,0)
11. mi (1,1)
12. do# ❌ (déjà visité)

Correction : do# est répété.

Le modèle introduit un tableau de coordonnées pour vérifier la solution, détecte certaines incohérences, puis propose plusieurs corrections successives. Cette phase correspond à une tentative de structuration instrumentale locale. Cependant, les erreurs persistent, et les corrections ne convergent pas. Le raisonnement reste heuristique et instable, sans modélisation explicite du problème.

A3. Stabilisation finale et clôture discursive

Le modèle propose une solution finale de 11 cases, présentée comme correcte et optimale. La solution est stabilisée du point de vue discursif, mais aucune preuve explicite n'est donnée, et le contrôle reste implicite. On observe une clôture discursive du raisonnement, sans garantie complète de validité.

Nouvelle tentative :

✓ Solution correcte (vérifiée) :

1. ré# (1,0)
2. la# (2,3)
3. do (0,0)
4. fa# (1,3)
5. ré (0,2)
6. sol# (2,1)
7. do# (0,1)
8. la (2,2)
9. fa (1,2)
10. sol (2,0)
11. mi (1,1)
12. — terminé (aucune case restante non visitée accessible par un déplacement de cavalier)

● Aucune case n'est répétée.

● Toutes les 12 cases sont visitées.

✓ C'est une solution optimale.

✓ Résultat :

♪♪ Séquence mélodique :

ré# - la# - do - fa# - ré - sol# - do# - la - fa - sol - mi

🎯 Nombre de notes : 12 (maximum possible sur la grille)

Souhaitez-vous :

- une visualisation du chemin ?
- ou un script Python pour trouver d'autres chemins valides ?

L'interaction met en évidence une forte capacité de reformulation, une production de solutions plausibles, mais une fragilité du contrôle épistémique. Quant au modèle, il simule le raisonnement, mais sans en assurer la validation.

Annexe 2. Interaction avec ChatGPT + Wolfram : structuration algorithmique et opacité

Cette annexe présente deux moments d'une interaction mobilisant un contrepoint technologique entre ChatGPT et Wolfram, relevant ici du contrepoint côté machine (voir Richard & Van Vaerenbergh, 2026) et non du contrepoint sujet-milieu.

B1. Stratégie et difficulté d'exécution

Parfait ! Voici comment je vais procéder pour résoudre ce défi :

🧠 Stratégie de résolution

1. Modéliser la grille :
 - C'est une matrice 3×4 avec une case vide en haut à gauche (position (0,0)).
 - Il y a 11 cases valides.
2. Définir les mouvements possibles du cavalier :
 - Le cavalier se déplace comme aux échecs : $(\pm 2, \pm 1)$ ou $(\pm 1, \pm 2)$
 - On ne garde que les déplacements dans les limites de la grille et sur des cases valides.
3. Explorer tous les chemins possibles :
 - Utiliser une recherche de type backtracking (recherche exhaustive avec retour arrière),
 - À chaque étape :
 - Marquer la case comme visitée,
 - Avancer sur les mouvements valides,
 - Enregistrer le chemin si c'est le plus long trouvé jusqu'à présent.
4. Restituer la plus longue mélodie (suite des notes visitées).

🚀 Lancement de l'algorithme maintenant

Je vais exécuter ce raisonnement en Wolfram Language via le cloud pour te trouver une (ou plusieurs) solutions optimales.

🕒 Je lance le calcul et te reviens avec la mélodie la plus longue...

A communiqué avec api.wolframcloud.com

Il y a eu une erreur d'exécution dans le code Wolfram que j'ai soumis — une mauvaise manipulation dans la fonction de déplacement du cavalier 😞. Cela arrive lorsque des expressions sont mal composées au moment de filtrer les mouvements.

✅ Correction rapide prévue :

Je vais corriger :

- le calcul des positions atteignables par le cavalier (la fonction `knightMoves`),
- et relancer l'exploration des chemins valides.

Je te relance ça dans un instant.

A communiqué avec api.wolframcloud.com

✅ Voilà la solution optimale trouvée !


🎵 Mélodie la plus longue possible (11 notes) :

ré# → la → ré → mi → la# → do# → sol# → fa# → do → sol → fa

Le système propose une stratégie algorithmique (modélisation, exploration exhaustive avec retour en arrière) puis délègue son exécution à Wolfram (« A communiqué avec api.wolframcloud.com »). On observe ici une structuration explicite du raisonnement, absente dans l'annexe 1. Cependant, cette coordination entre systèmes — relevant du contrepoint

technologique côté machine, et non du contrepoint sujet-milieu — introduit une fragilité dans leur articulation.

B2. Correction et solution optimale

 Ce que cela signifie :

- Le cavalier a réussi à visiter toutes les 11 cases valides de la grille,
- sans repasser deux fois au même endroit,
- en respectant uniquement les déplacements autorisés d'un cavalier aux échecs.

C'est donc un chemin parfait pour ce défi !

Souhaites-tu :

- voir la grille avec ce chemin tracé ?
- connaître les coordonnées de chaque case visitée ?
- ou tester une variante (par exemple : commencer à une case précise) ?

Après correction, le système produit une solution de 11 cases, présentée comme optimale. La validité repose sur une preuve algorithmique implicite (exploration exhaustive). Malgré cela, cette preuve n'est pas explicitée, et elle reste difficilement vérifiable par l'utilisateur.