Planification inverse et stratégies d'optimisation dans les TPS

Dr Frédéric Miéville,
physicien médical SSRPM

HESAV – Bsc-S3
29.10.2020

Plan du cours

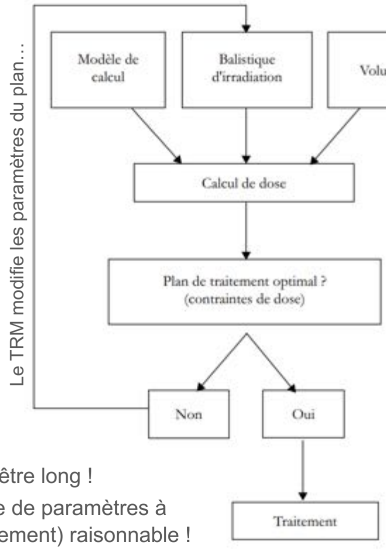
- La planification direct
- Pourquoi sommes-nous arrivé à la planification inverse ?
- Le concept d'optimisation basé sur une fonction de coût
- Deux types de stratégies d'optimisation
    i. Déterministes
    ii. Stochastiques
- Notion de contraintes
- La segmentation IMRT et VMAT
- Pour aller plus loin… *Multi-criteria Optimization* (MCO) et frontière de Pareto
- Résumé
- Exercice et corrigé

Concepts de planifications direct et inverse

La planification direct : concept

- La planification direct est ce que l'on appelle plus couramment la **radiothérapie conformationnelle en 3D (3DCRT)**
- C'est la première technique ayant découlé de l'apparition du CT scanner (imagerie 3D)
- Elle s'applique aussi bien pour les **photons** que pour les **électrons**
- La création d'un plan 3DCRT se déroule généralement de la façon suivante :
  1. Le dosimétriste reçoit :
     a. Les données anatomiques du patient (images CT)
     b. Les volumes du médecin (volumes cibles et organes à risque)
     c. La prescription médicale du médecin
  2. Le dosimétriste définit une balistique de traitement (nombres de champs de traitement, énergie, orientation, MLC, filtre, etc.)
  3. Le TPS calcule la dosimétrie (algorithme de calcul de dose)
  4. La dosimétrie est évaluée avec certains outils (isodoses 2D et DVH)

La planification direct : concept

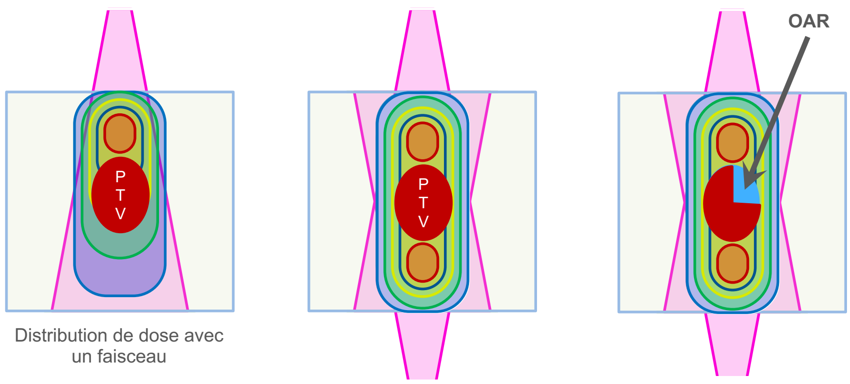


- Processus qui peut être long !
- Il faut que le nombre de paramètres à ajuster soit (humainement) raisonnable !

➢ Il faut que les faisceaux restent *simples* (donc pas de modulation d'intensité…)

La planification direct : distribution de dose

- Quelle est le type de distribution de dose et les caractéristiques engendrées par une planification direct basée sur un faisceau de photons ?



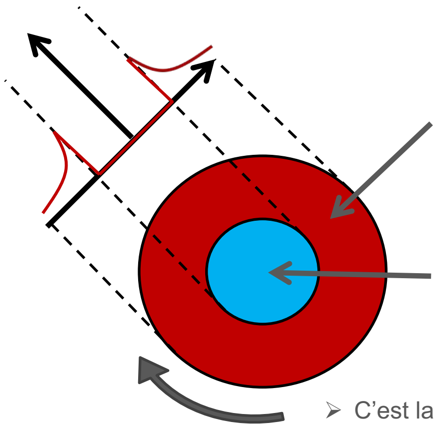
Distribution de dose avec un faisceau

Distribution de dose avec deux faisceaux (tunnel de dose)

➢ Peut-t-on trouver une méthode pour épargner les OARs ?

Origine de la planification inverse : problème inverse

- **Quel profile de fluence** permet d'obtenir la distribution de dose ci-dessous pour une irradiation rotationnelle (le gantry tourne autour du patient) ?



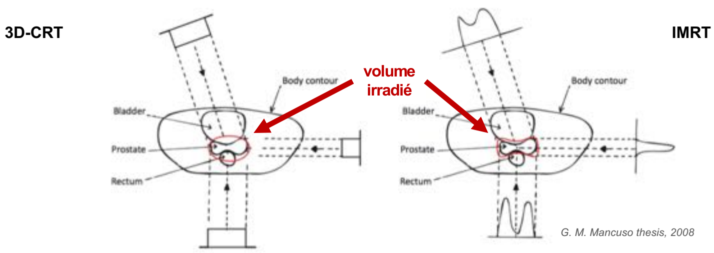
PTV (tumeur) : distribution de dose uniforme

Organe à risque (moelle) : dose 0 Gy

➢ C'est la réponse à cette question qui a débouché sur **l'IMRT/VMAT** et **la planification inverse**

Pourquoi l'IMRT ?

- Motivation : pour avoir une dosimétrie **homogène** et **conformationnelle**, on a besoin de champs avec des **intensités non uniformes**



G. M. Mancuso thesis, 2008

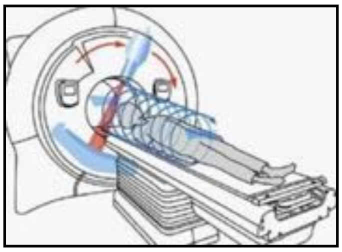
- Définition : L'IMRT est une technique de traitement basée sur l'utilisation de faisceaux d'incidences multiples et dont **l'intensité est modulée***. La dose délivrée au volume cible (PTV) est la superposition de ces faisceaux. **Les degrés de libertés (angle d'incidence, intensité) sont ajustés** afin d'optimiser la couverture au PTV et de diminuer la dose aux organes à risques (OAR)

Idée de A. Brahme (1982)

* IMRT statique = le gantry est statique mais le MLC bouge, VMAT et IMRT rotationnel = le gantry tourne et le MLC bouge

Pourquoi l'IMRT ?

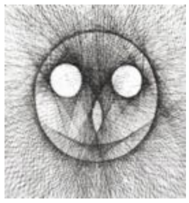
- Alternative : Mackie propose une **solution rotationnelle** appelée Tomotherapy*



Idée de T. R. Mackie (1993)

* le gantry tourne, le MLC bouge et la table avance

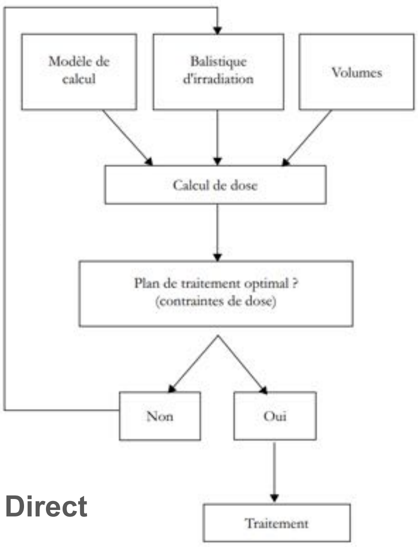
- L'IMRT rotationnel permet une grande flexibilité due au nombreux **degrés de liberté**.
- L'optimisation ne peut plus être «faite à la main» d'où l'introduction du concept de **planning inverse**



G. D. Birkhoff (1940): Dessin composé uniquement de ligne droite. Construire ce type de diagramme est mathématiquement équivalent à résoudre un problème de planning inverse IMRT

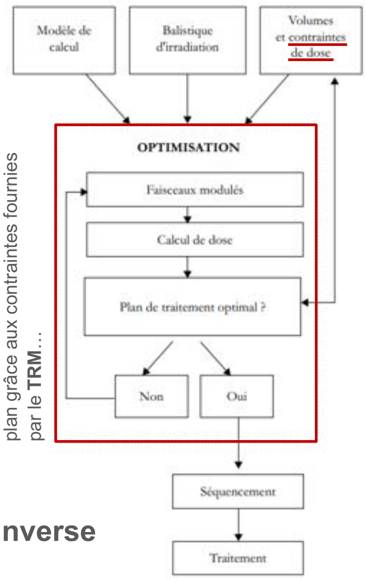
La planification direct et inverse

Le **TRM** modifie les paramètres du plan…



Direct

Le **TPS** modifie les paramètres du plan grâce aux contraintes fournies par le **TRM**…



Inverse

Bases physiques/mathématiques de l’IRMT

- **On peut mathématiquement écrire le problème de la façon suivante :**

$$\boxed{D = \mathbf{A} \cdot b}$$

Valeurs de dose que l’on voudrait

Pondérations des faisceaux

«fonction» qui permet de passer des pondération du faisceaux (en entrée) à la dose (en sortie)

Mathématiquement : Matrice de transfert de l’espace des pondérations à l’espace des doses.

- **Visuellement**

$$b \longrightarrow \boxed{A} \longrightarrow D$$

« Fonction/système »

Bases physiques/mathématiques de l’IRMT

Il suffit d’inverser le problème…

$$\boxed{\mathbf{A}^{-1} \cdot D = b}$$ **ou** $$D \longrightarrow \boxed{A^{-1}} \longrightarrow b$$

« Fonction/système inverse »

… mais pas toujours possible

- Générerait des pondérations négatives
- Matrice trop grande
- Temps de calcul trop long

La matrice est *ill-conditioned (mal conditionnée)*

- Il existe beaucoup de vecteur ***b*** tels qu’appliquée à ***A*** donnent ***D***
- Plusieurs solutions numériques existent

Autre difficulté

- Comment l’utilisateur peut définir ***D*** *dans le TPS* ?

  ➢ Il faut trouver une autre approche car **une inversion de la matrice A pose globalement trop de problèmes**

Concept et algorithmes d'optimisation

Optimisation basée sur le concept de coût (exemple)

Une optimisation basée sur le **concept de coût** est un **processus itératif** qui va permettre de converger vers une solution **«optimum»**

Exemple (simple) : on veut trouver la valeur minimum d'un dé

- On définit une fonction de coût, par exemple :

  Coût = valeur dé * 1000

- On garde la solution qui produit la fonction de coût minimal (minimisation du coût)
- On répète l'opération plusieurs fois.

Essayons :

1. On choisi une **valeur initiale** au hasard, par exemple $v_0 = 4$ (coût = 4000)
2. On jette le dé, $v_1 = 5 \rightarrow$ coût = 5000 $\rightarrow$ on garde donc $v_0 = 4$ (car 4000 < 5000)
3. On jette le dé, $v_2 = 3 \rightarrow$ coût = 3000 $\rightarrow$ on garde donc $v_2 = 3$
4. On jette le dé, $v_3 = 6 \rightarrow$ coût = 6000 $\rightarrow$ on garde donc $v_2 = 3$
5. On jette le dé, $v_4 = 1 \rightarrow$ coût = 1000 $\rightarrow$ on garde donc $v_4 = 1$
6. On jette le dé, $v_5 = 2 \rightarrow$ coût = 2000 $\rightarrow$ on garde donc $v_4 = 1$
7. On remarque que l'on converge **vers la valeur 1** (qui est bien le minimum du dé) !

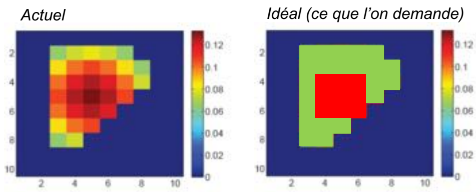
Optimisation basée sur le concept de coût (radio-oncologie)

Processus itératif

i. Le coût est recalculé après chaque itération
ii. Le changement est accepté s'il mène à un coût inférieur
iii. Le but final est de trouver **le minimum de la fonction de coût**[1]

Exemple (radiothérapie) :

$$\boxed{cout = \sum_{i=1}^{n} p_i \left(D_i - D_i^{idéal}\right)^2}$$



Exemple d'un segment de MLC

D_i est la dose que l'on a dans le voxel i
$D_i^{idéal}$ la dose que l'on aimerait dans le voxel i
p_i est un poids qui donne une importance au voxel considéré
n est le nombre total de voxels

[1] Il y a plusieurs sortes de fonction de coût mais la plus utilisée est celle des moindres carrés (*Mean-Squared Error cost function*)

Méthode d'optimisation

Deux grandes familles d'algorithmes

i. Algorithmes **déterministes**
ii. Algorithmes **stochastiques** (ou aléatoires)

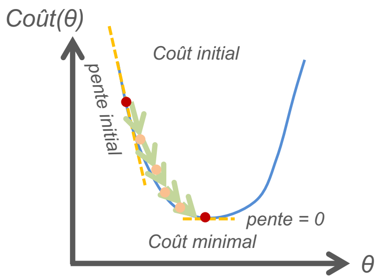
Algorithmes déterministes

- Avantage : rapides
- Désavantage : la solution peut être **bloquée** dans un **minimum local** et donc on a le risque de jamais arriver à la meilleure solution
- Exemple : **méthode du gradient**, méthode du gradient conjugué, méthode de la plus grande pente (*steepest descent*), etc.

Algorithmes stochastiques

- Avantage : moins sensibles aux minimums locaux (théoriquement insensible), permet théoriquement d'atteindre le **minimum global**
- Désavantage : lents, beaucoup de paramètres à régler
- Exemple : **recuit simulé** (*simulated annealing*), algorithmes génétiques, etc.

Méthode du gradient : concept

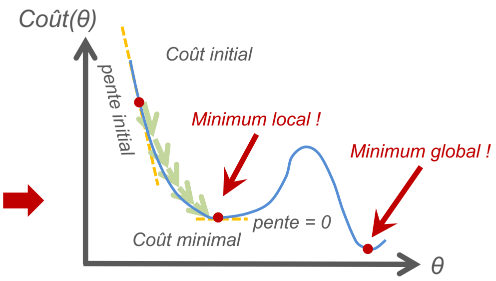


Formule gradient conjugué
$x_{i+1} = x_i - \Delta dy(x_i)$

Méthode :

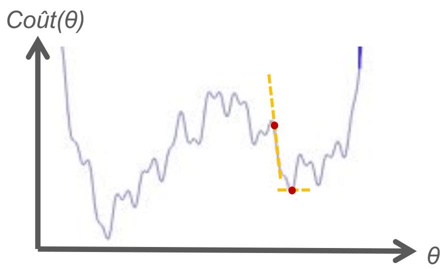
i. On détermine la pente au point où l'on est
ii. On fait «un petit pas» dans la direction de la pente en appliquant la formule du gradient conjugué (calcul de x_{i+1})
iii. On répète l'opération i. et ii.
iv. On s'arrête quand la **pente vaut zéro** (donc quant le coût est minimal localement)

Mais attention !
Il se peut que l'on ait pas le coût minimal et donc pas la meilleure solution…



Méthode du recuit simulé : concept

« Cette méthode vient du constat que le refroidissement naturel de certains métaux ne permet pas aux atomes de se placer dans la configuration la plus solide. La configuration la plus stable est atteinte en ***maîtrisant le refroidissement et en le ralentissant par un apport de chaleur externe*** *»* *(Mise au point par 3 chercheurs de la société IBM 1983)*



- **Si on accepte un état diminuant le coût**, on tend ainsi à chercher le minimum dans le voisinage de l'état de départ
- **Si on accepte un état augmentant le coût** cela permet d'explorer une plus grande partie de l'espace des états et tend à éviter de s'enfermer trop vite dans la recherche d'un optimum local

- Progressivement, on diminue l'amplitude « des sauts possibles » donc on réduit l'espace que l'on explore
- On a trouvé le **minimum global** de notre fonction de coût

➢ La dose trouvée est le plus proche possible de la dose idéale !

Comment spécifier $D_i^{idéal}$? Utilisation de contraintes

$$cout = \sum_{i=1}^{n} p_i \left(D_i - D_i^{idéal}\right)^2$$

D_i est la dose que l'on a dans le voxel i

$D_i^{idéal}$ la dose que l'on aimerait dans le voxel i

p_i est un poids qui donne une importance au voxel considéré

n est le nombre total de voxel

- La dose que l'on aimerait dans chaque voxel **n'est pas directement spécifiable** (512x512x100 images = 26.1 millions de voxels) **!**
- On introduit donc des **métriques intermédiaires** (au lieu de spécifier la dose voxel par voxel)
  - **Dose min**, **max** dans le volume spécifié
  - **DVH min**, **DVH max** sur le DVH du volume spécifié
  - **EUD** (Equivalent Uniform Dose) du volume spécifié
  - Etc.

Tableau de contraintes

- Ces métriques sont regroupées dans un tableau, c'est le **tableau de contrainte**.

coût

ROI	Type	Constrain	Gy	% Volume	% Variation	Weight	Objective Value	a	gEUD
PTV 2 Prostate+VSp	Min Dose		17.5			90	0.00511689		
PTV 2 Prostate+VSp	Max DVH		18	50		40	5.03862e-05		
PTV 2 Prostate+VSp	Min DVH		18	50		40	0		
PTV 2 Prostate+VSp	Max Dose		18.5			80	0.000163892		
Vessie	Max EUD		9.5			2	0.000769994	3	9.91681
TF G	Max EUD		4			1	7.63298e-05	3	4.07814
TF G	Max EUD		4			1	7.63298e-05	3	4.07814
Genita	Max EUD		1			3	0.198227	3	1.57479
Bulbe penien	Max EUD		2			2	0.00306859	3	2.17517
Rectum	Max EUD		9.5			2	0.000406835	3	9.80297

Coût total (sommes de coûts) que l'on veut mimiser — Composite objective value 0.207957 — Recompute Values

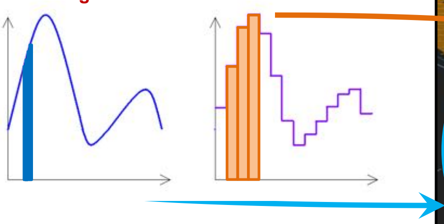
somme de tous les coûts

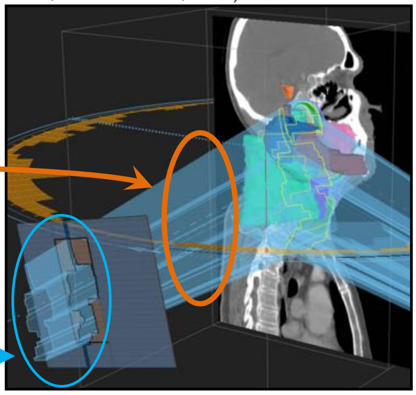
- Ce tableau est déterminé/construit par l'utilisateur, généralement, de façon itérative (il ne faut pas hésiter à le modifier ou même recommencer).
- Ceci peut prendre beaucoup de temps (suivant la complexité du problème) avant d'arriver à un tableau qui va mener à une distribution de dose satisfaisante !

Segmentation

- On a vu l'optimisation de la **fluence théorique** avec les méthodes d'optimisation précédentes (fonction de coût, contraintes, etc.)
- Maintenant, il faut mettre en place la mécanique du linac (MLC, gantry, etc.) pour obtenir la **fluence réelle**

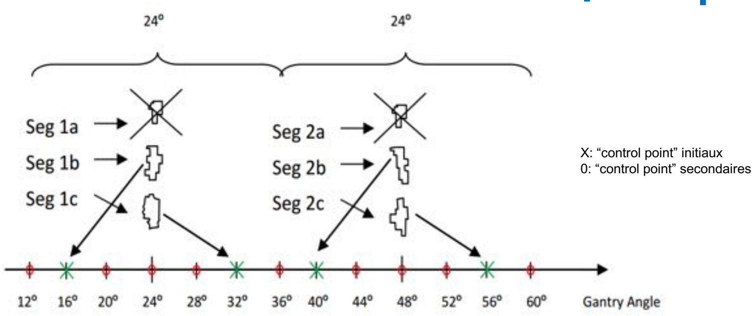
Segmentation = discrétisation





- **Dégrader** la distribution de dose !
- Pour l'IMRT non rotationnelle, il y a la méthode *Direct Machine Optimisation Parameters* (DMPO), qui optimise **directement la fluence réelle**

Création d'une IMRT rotationnel: principe



From RaySearch

- **Création de l'arc**: tous les 24° on segmente la fluence en 3 segments. On retient à chaque fois 2 segments que l'on redistribue autour l'angle considéré (+/- 8°. Ex: angle = 48°: on a un segment à 40° et un segment à 56°)
- Ajout de *control points* par clonage (secondaires)
- Les contraintes mécaniques du MLC (vitesse max, débit de dose max, etc.) sont, en général, directement prises en compte durant le processus d'optimisation (mais la stratégie peut dépendre du TPS…)

Optimisation Multicritères

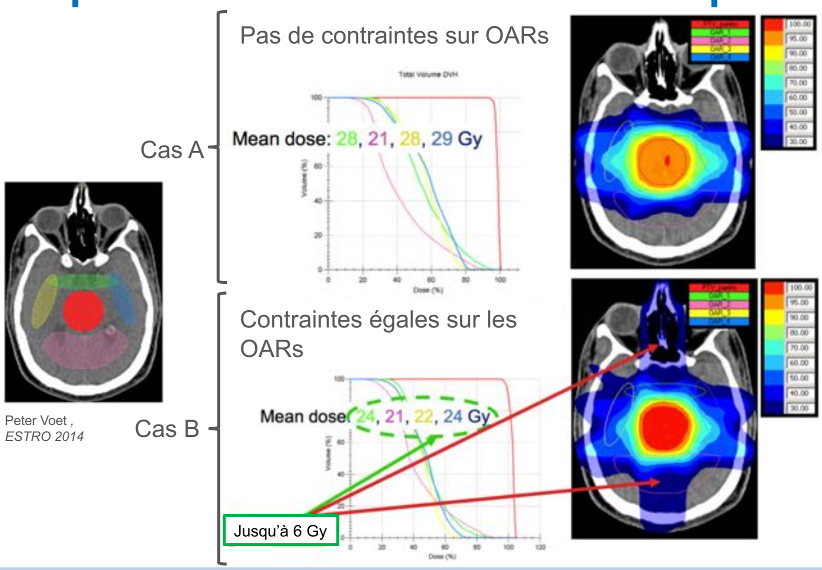
- Optimisation Multicritères = *Multi-criteria Optimization* (MCO)
- Résumé :
  - On a vu que l'on avait des algorithmes pour optimiser la dose du plan. La dose est optimisée via la minimisation d'un fonction de coût. La fonction de coût est basée sur un tableau de contraintes. Le tableau de contraintes est donné par l'utilisateur.
- Problématique :
  - Si le tableau est mal construit la solution ne va pas être optimum…
  - Imaginons que l'on arrive à un plan optimum. Est-ce que ce plan sera aussi optimum pour le médecin ? Que-ce qu'une solution optimum ?
  - Sommes–nous certain d'avoir atteint la solution optimum lorsque l'on est satisfait de notre plan ? Somme-nous loin du plan optimal ?

➢ L'optimisation manuelle (via un tableau de contraintes) **peut être extrêmement chronophage** pour un résultat final qui **n'est pas optimal** !

➢ But de MCO : pouvoir **choisir** parmi un série de plan **tous optimaux** !

Frédéric Miéville, PhD I Service de radio-oncologie I 29.10.2020 I 23

Optimisation Multicritères : exemples



Peter Voet , *ESTRO 2014*

Frédéric Miéville, PhD I Service de radio-oncologie I 29.10.2020 I 24

BONUS = pas évalué à l'examen

Pour aller plus loin…

Quelques considérations pratiques

- Actuellement, les TPS utilisent plutôt des versions d'optimisation améliorées principalement basée sur la méthode du gradient. Ceci par gain de temps.
- Le nombre de paramètres à optimiser est très important. D'où un temps relativement long pour l'optimisation d'un traitement VMAT.
  - Ordre de grandeurs : 1 arc VMAT utilisant 179 points de contrôle (CP) pour un TrueBeam Varian → 179 cp x 120 lames par segments = 21480
- Sur certains TPS (Pinnacle), on alterne une optimisation des points des segments (MU par segment) et une optimisation de la forme des segments (lames du MLC).
- Pour le TPS Tomo (Accuray), l'optimisation est principalement en fluence puis la segmentation vient à la fin.
- Pour une optimisation IMRT (non rotationnel), on a parle souvent de DMPO (*direct machine parameter optimisation*), où la position des segments et les poids sont optimisés ensembles
- Chaque TPS a sa stratégie et il reste difficile d'en connaître tous les secrets !

En résumé :

- Pour avoir une dosimétrie **homogène** et **conformationnelle**, on a besoin de champs d'irradiation avec des intensités non uniformes
- Il faut donc être capable de **moduler la fluence**. Ceci se fait avec le MLC (création de plusieurs segments)
- Le nombre de paramètres à déterminer est extrêmement important, ce qui conduit à la notion de **fonction de coût**
- Le cœur de la planification inverse repose sur **l'optimisation (minimisation) d'une fonction de coût**
- Il y a les méthodes d'optimisation déterministes et stochastiques
  - **i. Déterministe**: rapide mais l'optimisation peut être bloquée dans un minimum local
  - **ii. Stochastique**: lente mais a beaucoup plus de chance d'arriver à trouver le minimum global de la fonction de coût
- Après l'optimisation de la fluence théorique, il faut **ajouter la mécanique** pour obtenir **la fluence réelle** (celle qui sera délivrée par la machine)
- Contrairement à l'IMRT non rotationnelle, **l'IMRT rotationnel** doit encore avoir une **stratégie pour créer les segments de l'arc**

Les objectifs du cours sont :

- Connaître la différence entre la planification direct et la planification inverse
- Savoir expliquer le concept d'optimisation basé sur une fonction de coût
- Connaître et expliquer les deux grandes stratégies d'optimisation
- Savoir la différence entre un plan en fluence et en plan segmenté
- Connaître l'idée de la stratégie *Multi-criteria Optimization* (MCO)

Exercice

Exercice : optimisation méthode du gradient

i. Trouver visuellement le minimum de la fonction $y(x) = - x^3 + x^2 + 60x + 20$

ii. Trouver le minimum par la méthode du gradient sachant que:

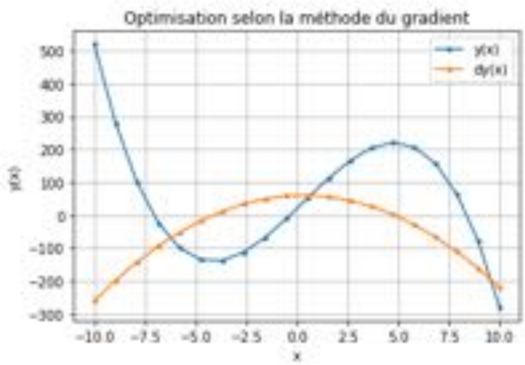
- *Condition initial : $x_0 = -9$ et $\Delta = 0.05$ (pas de convergence de l'algorithme)*

$y(x) = - x^3 + x^2 + 60x + 20$

$dy(x) = -3x^2 + 2x + 60$

Formule gradient conjugué

$$x_{i+1} = x_i - \Delta dy(x_i)$$



Aide

1. Calculer $y(x_0)$, $dy(x_0)$ puis x_1
2. Puis calculer $y(x_1)$, $dy(x_1)$ puis x_2
3. Continuer jusqu'à ce que:

$y(x_{i+1})-y(x_i) < 0.1$

Corrigé de l'exercice

i. Trouver visuellement le minimum de la fonction $y(x) = - x^3 + x^2 + 60x + 20$

ii. Trouver le minimum par la méthode du gradient conjugué sachant que:

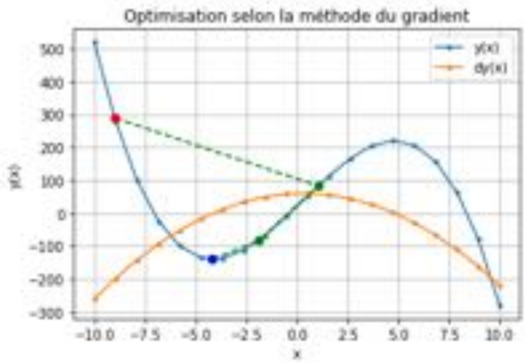
- *Condition initial :* $x_0 = -9$
- *Pas:* $\Delta = 0.05$

$y(x) = - x^3 + x^2 + 60x + 20$

$dy(x) = -3x^2 + 2x + 60$

Formule gradient conjugué

$$x_{i+1} = x_i - \Delta dy(x_i)$$



Réponse

$x_0 = -9.0$	$y_0 = 290.0$	$dy_0 = -201.0$	error = 207.055
$x_1 = 1.05$	$y_1 = 82.945$	$dy_1 = -58.792$	error = 166.004
$x_2 = -1.89$	$y_2 = -83.06$	$dy_2 = -45.509$	error = 57.242
$x_3 = -4.165$	**$y_3 = -140.302$**	**$dy_3 = -0.373$**	**error = 0.002**

En 3 itérations on trouve le minimum…