

Chapitre 2. Le diagramme de Lexis et les conditions de la cohortisation

2.1. Introduction

Le premier chapitre a montré toute l'importance du diagramme de Lexis en analyse démographique classique. Auxiliaire précieux, ce graphique sert notamment à montrer comment les données à analyser s'agencent entre elles, ce qui permet aux démographes non seulement de s'orienter dans le choix des méthodes à utiliser, mais aussi de bien saisir la nature de l'analyse en cours (transversale ou longitudinale, par exemple).

Le deuxième chapitre portera sur ce diagramme. Dans un premier temps, nous en signalerons les particularités au sein de la famille des diagrammes cartésiens (point 2.2). Ensuite, pour rendre compte de ces spécificités, sa construction sera explicitée sur base graphique (point 2.3). Le point suivant servira à dégager les conditions mathématiques précises qui conduisent à l'émergence d'une figure du type « diagramme de Lexis » et, en conséquence, à une possible cohortisation de l'analyse des phénomènes en cause (point 2.4).

Son utilisation sera, dans un premier effort de généralisation, étendue à d'autres variables temporelles que l'âge et le temps (comme, par exemple, la durée écoulée depuis le mariage ou une migration) ainsi qu'à d'autres présentations (comme des diagrammes à plus de deux axes et 3 variables) (point 2.5). Enfin, le dernier point abordera rapidement la question de l'historique de l'élaboration du diagramme de Lexis (point 2.6).

Par rapport à l'objectif de ce travail, qui vise à voir s'il est possible d'exporter les méthodes de l'analyse démographique classique à d'autres domaines, ce chapitre se révélera en fait crucial. En effet, après avoir identifié dans le premier chapitre ce qu'il y a à exporter dans l'analyse démographique classique, ce chapitre permettra de mettre en évidence les conditions que doivent remplir les observations pour donner lieu à ce transfert de méthodes.

2.2. Spécificité du diagramme de Lexis

Le diagramme de Lexis appartient à la famille des diagrammes cartésiens : une observation peut s'y localiser en fonction de ses coordonnées d'âge et de temps, variables occupant respectivement les axes des ordonnées et des abscisses. Toutefois, le diagramme de Lexis ne se résume pas à un banal diagramme cartésien classique. En fait, deux caractéristiques l'en distinguent :

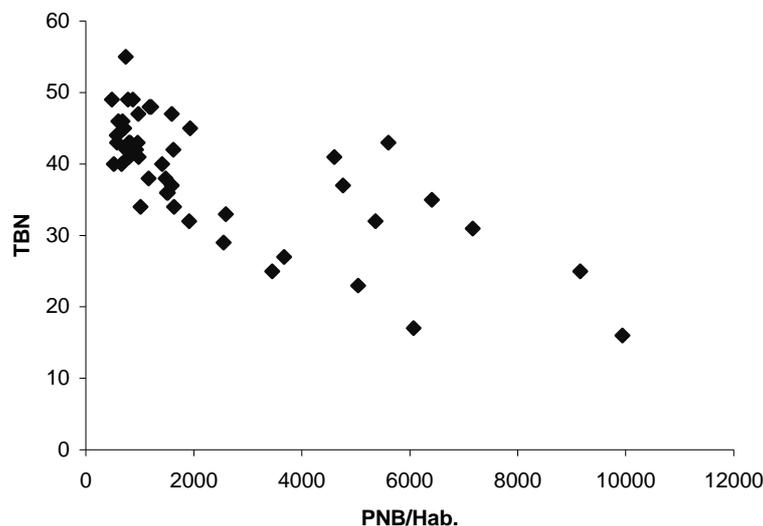
- si le diagramme de Lexis comporte deux axes, il n'en permet pas moins l'utilisation de trois coordonnées distinctes ;
- par ailleurs, sur le diagramme de Lexis, chaque individu sous observation se voit cantonné à une trajectoire forcée dont il ne peut s'échapper, à savoir, sa ligne de vie.

Comme nous allons le montrer, l'apparition de la troisième coordonnée représente une caractéristique certes nécessaire mais non suffisante pour obtenir une représentation graphique méritant d'être qualifiée de « Lexis ». La spécificité du diagramme de Lexis lui vient de la présence des trajectoires forcées, indispensables pour définir des cohortes et, en définitive, pour recourir à la cohortisation afin d'analyser les phénomènes.

2.2.1. Deux axes et trois coordonnées

Prenons comme exemple de diagramme cartésien classique un graphique montrant la relation entre l'évolution du produit national brut par habitant (PNB/H) et le taux brut de natalité (TBN) pour différents pays d'Afrique en 2002 (cf. figure 2.1). Sur une telle figure, il n'est pas possible de glisser une troisième variable, à moins d'ajouter un troisième axe qui transformera le diagramme en plan initial en un autre en volume.

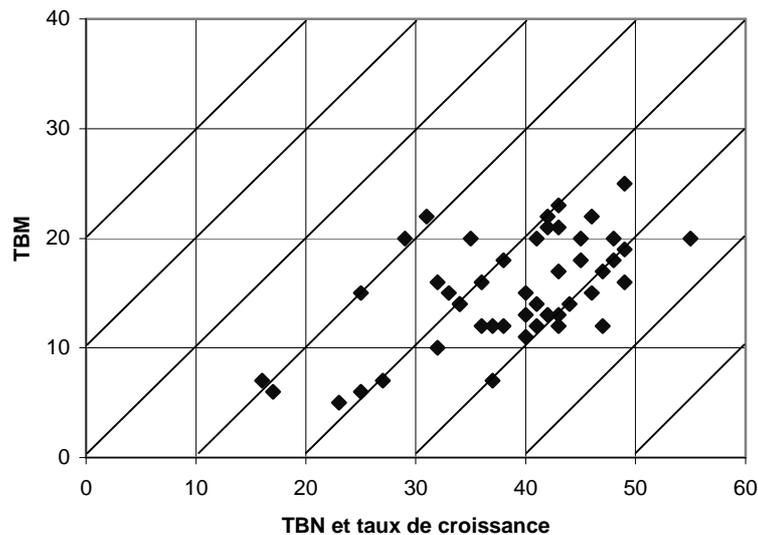
Figure 2.1. Relation entre le PNB/H et le TBN (en %) – Afrique 2002



(Source : Population Reference Bureau, 2002. World Population Data Sheet)

Dans certaines circonstances, des graphiques très proches de la figure 2.1 offrent en fait la possibilité d'utiliser trois variables, sans que ce soit nécessairement exploité. Ainsi, la figure 2.2, illustrant la relation entre le TBN et le taux brut de mortalité (TBM) pour l'Afrique en 2002, peut se compléter par un réseau de diagonales correspondant à des lignes d'iso-taux de croissance. L'échelle du TBN est également valable pour la troisième variable, le taux de croissance (TC), qui correspond à la différence entre les deux taux bruts (TBN – TBM).

L'oblique démarrant au point (0,0) correspond à un taux de croissance de 0 ‰, soit 0 ‰ – 0 ‰, à son point initial ; si le TBN et le TBM valent chacun 10 ‰, le taux de croissance sera de nouveau de 0 ‰. L'oblique commençant au point (10,0) matérialise sur le graphique le taux de croissance de 10 ‰, correspondant à des TBN et TBM valant, par exemple, respectivement 10 ‰ et 0 ‰ ou 20 ‰ et 10 ‰. Pour les TC négatifs, la valeur est à lire soit sur le prolongement à gauche de l'axe des abscisses, soit sur l'axe des ordonnées, mais en prenant le négatif des valeurs du TBM.

Figure 2.2. Relation entre le TBN et le TBM (en %) – Afrique 2002

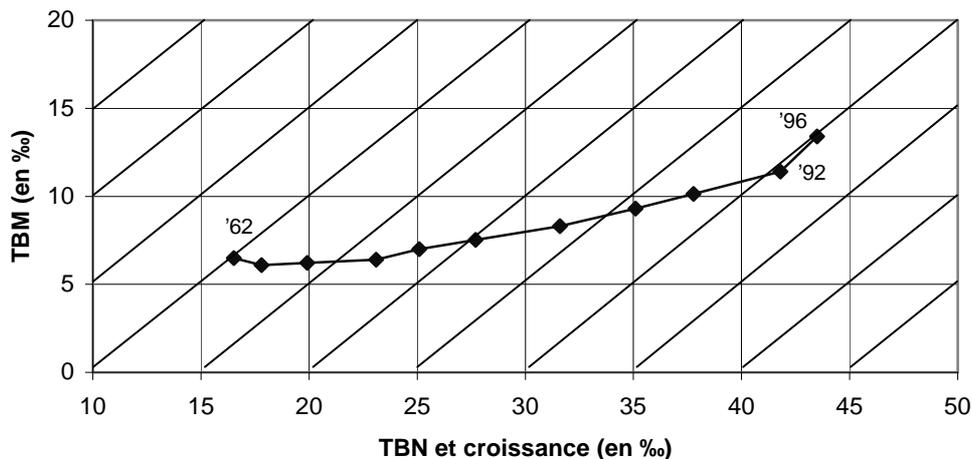
(Source : Population Reference Bureau, 2002. World Population Data Sheet)

La coordonnée *taux de croissance* peut s'utiliser en remplacement de l'une des deux autres pour localiser une observation, tout comme la coordonnée *moment de naissance* peut se substituer à l'âge ou au temps dans le diagramme de Lexis classique. Si la figure 2.2 remplit bien la condition des trois variables représentées avec seulement deux axes, elle n'en devient par pour autant un diagramme de Lexis ; en effet, les diagonales qui s'y dessinent ne sont nullement des trajectoires forcées.

2.2.2. Les trajectoires forcées

Sur un diagramme de Lexis, un individu ne peut s'écarter d'une trajectoire entièrement définie par son moment de naissance : tout ce qui affectera cet individu ne peut que se localiser sur sa ligne de vie. La figure 2.3 ne partage pas cette caractéristique ; elle montre la relation entre le TBN et le TBM pour la Thaïlande entre 1962 et 1996 selon les données de la banque Mondiale ; comme sur la figure 2.2, les diagonales matérialisent la coordonnée « taux de croissance ». Trois réseaux de droites peuvent donc, comme sur le diagramme de Lexis classique, servir pour localiser un point.

La Thaïlande peut très bien se trouver sur une diagonale d'iso-taux de croissance à un moment donné, éventuellement y rester (ce qui se produit pratiquement entre les deux derniers points à droite, ceux des années 1992 et 1996) ou la quitter, comme le montre le reste de la trajectoire qui est largement transversale par rapport aux diagonales. Les diagonales de cette figure n'ont donc rien de commun avec les lignes de vie du diagramme de Lexis classique. Par ailleurs, aucun des deux autres réseaux de droites présents sur la figure ne pourrait tenir ce rôle de trajectoires forcées. Ajoutons qu'aucun autre réseau, quel qu'il soit, ne pourrait le faire.

Figure 2.3. Relation entre le TBN et le TBM – Thaïlande 1962-1996

(Source : World Bank, 1998, World Development Indicators)

Enfin, si sur ce diagramme un troisième réseau de droites apparaît bien et montre la possibilité de recourir à une troisième coordonnée pour localiser une unité d'observation, il n'en devient pas pour autant un diagramme de Lexis classique à trajectoires forcées. Sur cette trame, il ne sera pas envisageable d'établir l'équivalent des cohortes et donc de cohortiser l'étude d'un phénomène au contraire de ce qui se passe avec un diagramme de Lexis classique.

En reprenant la figure 2.1, si la régression du TBN sur le PNB/H devait se révéler parfaite, une oblique pourrait y apparaître avec une inclinaison en fonction de la pente de la droite de régression. Toutes les unités d'observation y trouveraient place. S'agirait-il d'une oblique comparable à une ligne de vie ? Non. En effet, loin de représenter une trajectoire forcée individuelle, il s'agirait d'une trajectoire certes forcée, mais surtout collective ; une fois de plus, sur cette base, il ne serait pas possible de constituer des cohortes et donc de cohortiser l'étude d'un phénomène.

À notre sens, ce qui précède montre clairement que le diagramme de Lexis tire sa spécificité, par rapport aux diagrammes cartésiens, de la possibilité d'utiliser trois coordonnées dans un système à deux axes combinée avec la présence de trajectoires forcées individuelles. D'où viennent ces caractéristiques du diagramme de Lexis classique ? Pour répondre à cette question, nous allons expliquer, dans le point suivant, la construction d'un diagramme de Lexis classique.

2.3. Construction du diagramme de Lexis

Historiquement, pour expliquer son diagramme, Lexis partait d'un diagramme à un axe pour aboutir à la construction finale à deux axes (cf. point 2.6). À notre sens, il est plus logique de partir d'un diagramme à trois axes, de réduire le nombre d'axes et d'aboutir à la représentation en deux axes. C'est le chemin qui sera suivi dans ce point. Cette explication graphique sera suivie, au point 2.4, d'une explication mathématique. Cette dernière pourrait éventuellement suffire. Toutefois, nous avons préféré commencer par une réflexion graphique qui a pour avantage de donner chair en quelque sorte au raisonnement mathématique qui suivra. Cette présentation graphique sera par ailleurs bien utile pour certaines parties du point

2.5, notamment pour établir la distinction entre les diagrammes de Lexis à deux axes et trois coordonnées et ceux comportant trois axes et six coordonnées.

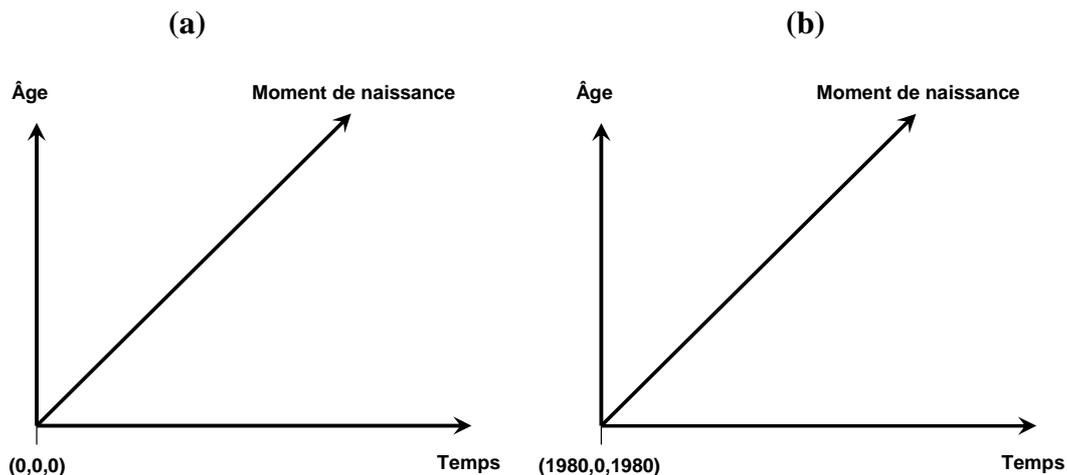
2.3.1. Trois coordonnées et trois axes

En analyse démographique classique, les observations à localiser sur un diagramme de Lexis peuvent se classer en fonction de trois coordonnées définissant le cadre de référence de l'analyse, à savoir en prenant l'exemple des décès :

- la **date** de survenance du décès (variable *temps*, désignée par T) ;
- l'**âge** du décédé au moment du décès (variable *âge*, désignée par A) ;
- le **moment de naissance** du décédé (variable *moment de naissance*, désignée par M).

La solution la plus immédiate pour représenter des données organisées selon un triple critère de classement suppose de dessiner un diagramme à trois axes ; chacun de ces axes supportera une des trois variables en jeu. Sur la figure 2.4.a, l'axe vertical concerne l'âge ; l'axe horizontal à l'avant-plan, le temps et finalement l'axe horizontal fuyant vers l'arrière, le moment de naissance.

Figure 2.4. Trois coordonnées et trois axes



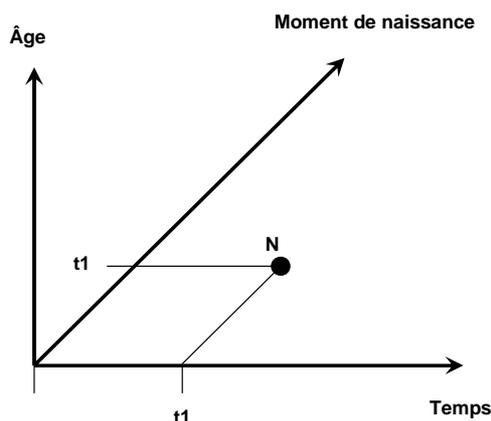
Les unités sur les trois axes sont identiques : il s'agit à chaque fois d'années. Par ailleurs, les échelles des axes du moment de naissance et du temps sont identiques : une succession d'années. À la limite, si la croisée des axes correspond au point (0,0,0), les trois échelles deviennent identiques. Habituellement, ce n'est pas le cas ; ainsi, si les données à représenter proviennent d'une observation ayant débuté en 1980, la croisée des axes aurait pour coordonnées (1980,0,1980), avec le premier 1980 pour le temps, le 0 pour l'âge et le second 1980 pour le moment de naissance (cf. figure 2.4.b). Il y aura donc un décalage de 1980 entre les échelles du temps ou du moment de naissance et celle de l'âge.

Il est à noter que, dans les figures qui suivront, nous avons supprimé la parenthèse contenant les coordonnées à la croisée des axes.

2.3.2. Localisation des lignes de vie dans le diagramme à trois axes

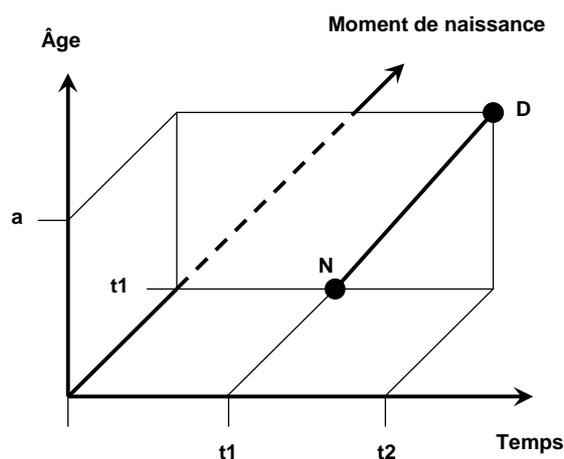
Comme expliqué dans le premier chapitre, la ligne de vie d'un individu montre l'évolution de son âge en fonction du temps. Supposons un individu né en $t1$ et mort en $t2$ à l'âge a . La localisation de son point de naissance sur le diagramme en trois axes (soit N) se fera en fonction des coordonnées suivantes (cf. figure 2.5) : $t1$ pour le temps ainsi que pour le moment de naissance et 0 pour l'âge.

Figure 2.5. Localisation du point de naissance



Comme indiqué sur la figure 2.6, le point de décès (D) se localisera aussi en fonction de ses trois coordonnées, à savoir : $t2$ pour la date ; $t1$ pour le moment de naissance (qui, bien évidemment, pour un individu reste constant durant toute sa vie) et a pour l'âge¹. La ligne de vie de cet individu relie les points de naissance et de décès selon une inclinaison à 45° , vu l'identité des unités sur les trois axes. L'individu en cause ne pourra pas quitter cette droite, qui représente donc pour lui une « trajectoire forcée ». C'est donc exclusivement sur cette ligne de vie que se localiseront tous les événements affectant cet individu.

Figure 2.6. Localisation du point de décès et de la ligne de vie



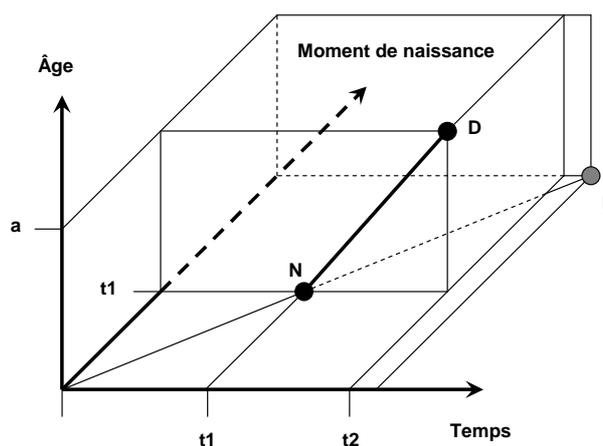
¹ Théoriquement $t2 - t1$ devrait être égal à a . Maintenir cette caractéristique compliquait de beaucoup la perspective de certaines figures qui suivront. Nous avons donc opté pour une inégalité ne respectant pas la relation mathématique au profit, nous l'espérons, de plus de confort dans certaines perspectives. Par ailleurs, quand certains éléments des graphiques sont utilisés dans le texte, ils sont désignés par des caractères italiques (a , $t1$, $N...$), même si sur les figures, ce n'est pas le cas ; nous espérons ainsi faciliter la lecture du texte.

2.3.3. Alignement des points de naissance et des lignes de vie

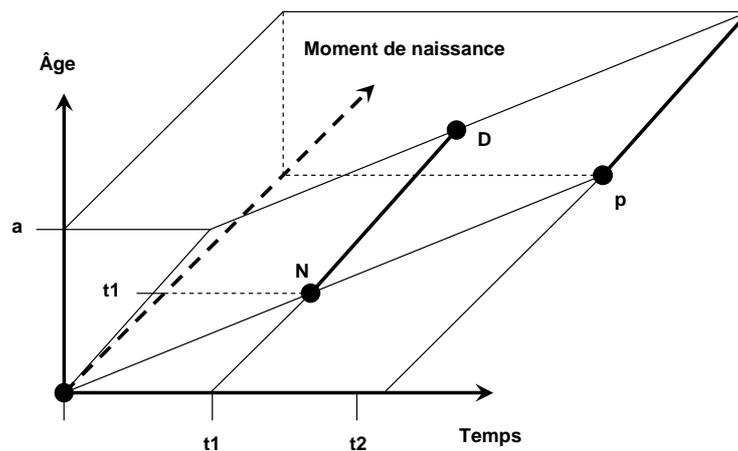
Le moment de la naissance correspond, pour chaque individu, à l'âge de 0 an exact. Sur le graphique à trois axes, les naissances devront se situer sur le plan *temps-moment de naissance* dont la coordonnée *âge* vaut 0. Par ailleurs, même si l'un des deux axes de ce plan porte une variable à valeur constante pour un individu (son moment de naissance) et l'autre pas (le temps), ils présentent tout deux la particularité d'être rigoureusement identiques. Ils se constituent de la même succession d'années (pour autant qu'ils se croisent à une date identique, ce que nous supposerons toujours par la suite, sauf éventuelles indications contraires).

En conséquence, les points de naissance des individus ne pourront que s'aligner le long d'une droite sur le plan *temps-moment de naissance* d'équation $T = M$. En fait, il s'agit de la bissectrice de l'angle formé par les axes du temps et du moment de naissance, soit la ligne partant de la croisée des axes et aboutissant à p et comprenant le point N (cf. figure 2.7). Le reste du plan *temps-moment de naissance* demeurera vierge de tout point de naissance.

Figure 2.7. Alignement des points de naissance



Une relation mathématique simple relie les trois variables *temps*, *âge* et *moment de naissance* : $A = T - M$. Vu que ces trois variables s'expriment dans la même unité (l'année), dans le graphique, les lignes de vie s'inclineront toutes de façon identique, dessinant un angle de 45° par rapport au plan *temps-moment de naissance*. Ce parallélisme combiné avec l'alignement des points de naissance sur la bissectrice fait en sorte que les lignes de vie occupent uniquement un plan et négligent le reste du volume défini par les trois axes (cf. figure 2.8). Ce plan N,D,P est lui-même incliné à 45° et recoupe le plan *temps-moment de naissance* en suivant la bissectrice des points de naissance.

Figure 2.8. Alignement des lignes de vie

En définitive, même si trois variables interviennent pour localiser les observations, un plan suffit pour représenter tout ce qui doit l'être. Dans ces conditions (qui seront affinées ultérieurement), le diagramme en volume peut se résumer dans un diagramme en plan par « réduction » d'un axe, tout en conservant la faculté de recourir au triple critère de classement des évènements à localiser. C'est ce qui va être montré maintenant.

2.3.4. Réduction du nombre d'axes et émergence d'un diagramme en plan

D'une manière générale, le passage d'un diagramme à trois axes à un diagramme à seulement deux axes s'effectue en projetant les lignes de vie sur un des trois plans délimitant le volume de la figure initiale. La variable dont l'axe disparaît ne disparaît pas elle-même. En fait elle va « parasiter » un des deux axes restants, dans le sens où cet axe « hôte » supportera en fait deux variables, la sienne et celle dont l'axe a disparu. Au lieu de se traduire dans le graphique par une horizontale ou une verticale, comme c'est le cas habituellement dans un diagramme cartésien, la variable parasite se matérialisera sur le plan par des obliques soit croissantes, soit décroissantes.

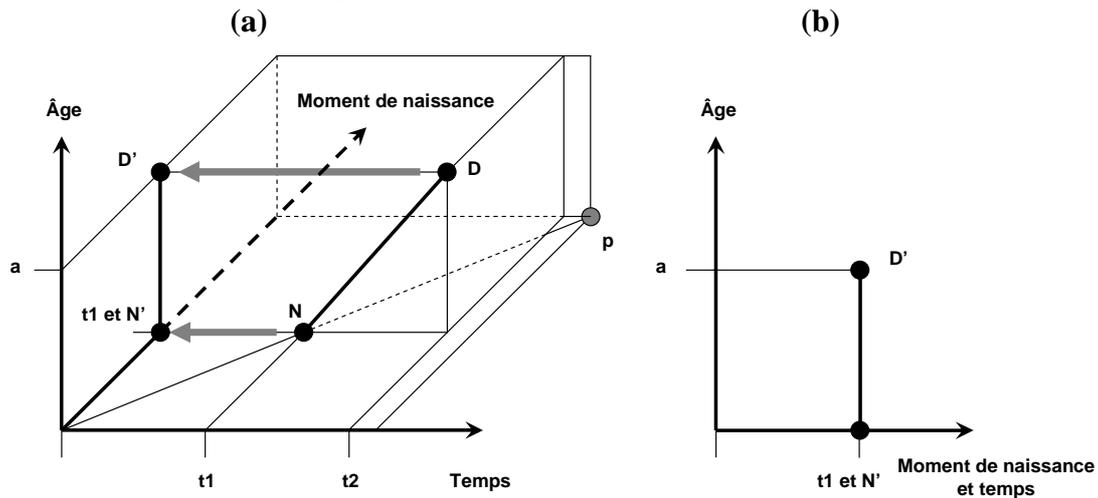
La projection peut se faire indifféremment sur l'un des trois plans du volume initial². Les trois diagrammes plans obtenus via cette opération de réduction seront dénommés selon le plan sur lequel s'effectue la projection. Par ailleurs, N' désignera le positionnement du point de naissance N après projection et D' , celui du point de décès D . Enfin, chacune des constructions obtenues sera illustrée par un graphique muet montrant la trame de droites servant au repérage des observations.

² En intervertissant les axes horizontaux et verticaux, ce sont en fait six versions du diagramme qui sont disponibles. Nous avons décidé ici de n'aborder pour chaque plan de projection qu'une seule des deux versions possibles.

2.3.4.1. Diagramme *moment de naissance-âge*

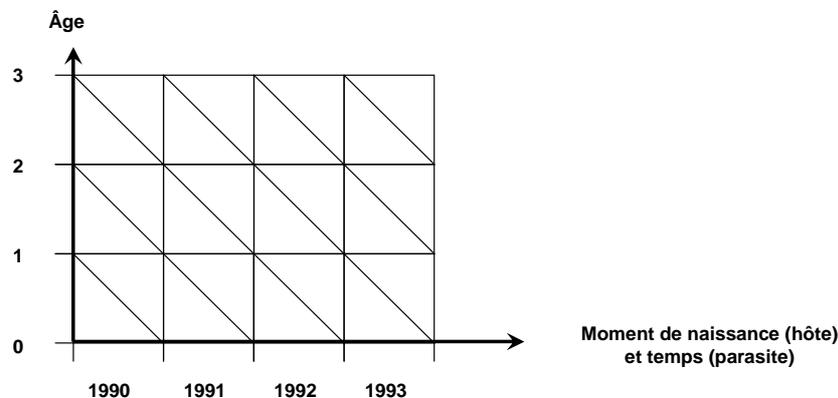
Sur la figure 2.9.a, la projection de la ligne de vie s'effectue perpendiculairement sur le plan *moment de naissance-âge*. En conséquence (comme montré également sur la figure 2.9.b illustrant le résultat de la projection), le point N' se confondra avec $t1$, la coordonnée de la naissance sur l'axe du moment de naissance, et D' se déplacera à la verticale de N' , à hauteur de a . Les lignes de vie apparaissent donc en position verticale. La coordonnée parasite est le temps ; il s'établit sur l'axe du moment de naissance.

Figure 2.9. Projection sur le plan *moment de naissance-âge*



Pour une utilisation pratique, le graphique de droite de la figure 2.9.b se complétera de réseaux de parallèles indiquant des coordonnées remarquables, comme expliqué dans le chapitre 1. Les verticales de la figure 2.10 correspondent à des dates de naissance tombant sur une délimitation d'année et les horizontales, à des âges exacts. Les obliques pour leur part matérialisent les délimitations des années. Une oblique doit s'interpréter comme suit : évolution de l'âge en fonction du moment de naissance pour une date donnée. Vu que les coordonnées hôte et parasite suivent exactement la même échelle, une seule suffira pour les deux au contraire de ce qui se produira pour le diagramme *temps-moment de naissance*.

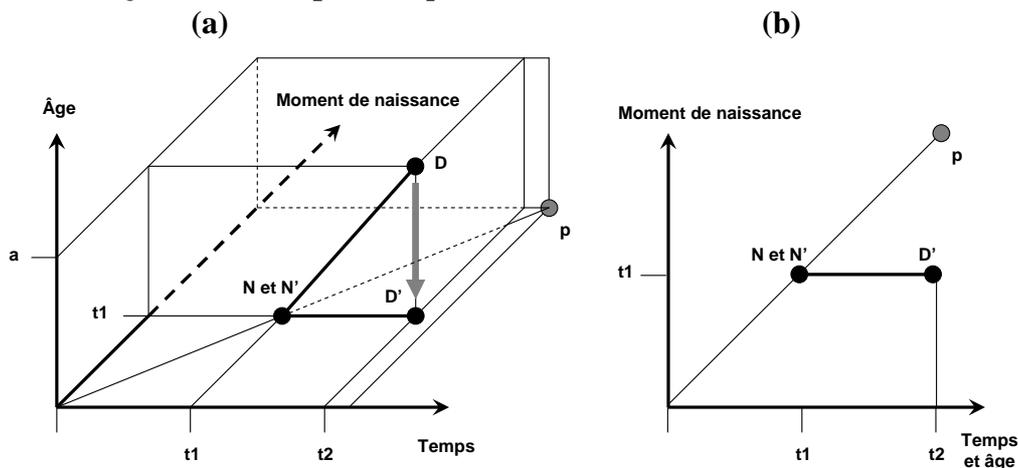
Figure 2.10. Trame sur un diagramme *moment de naissance-âge*



2.3.4.2. Diagramme *temps-moment de naissance*

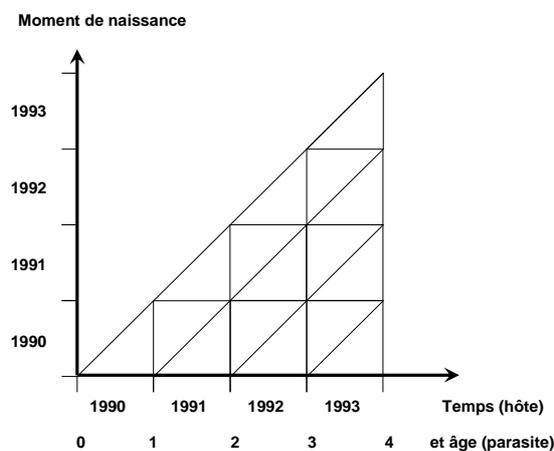
Dans ce cas, la projection de la ligne de vie s'effectue perpendiculairement sur le plan *temps-moment de naissance* (cf. figure 2.11.a). En conséquence, N et N' se confondent ; pour sa part D est projeté sur le plan *temps-moment de naissance* à hauteur de la coordonnée $t1$ sur l'axe du moment de naissance et $t2$ sur l'axe du temps (cf. figure 2.11.b). Les lignes de vie apparaissent donc en position horizontale. La coordonnée parasite est l'âge qui s'établit sur l'axe du temps. La bissectrice comprenant tous les points de naissance apparaît sur le graphique ; elle délimite les lignes de vie, si bien que sa gauche restera libre de toute observation.

Figure 2.11. Projection sur le plan *temps-moment de naissance*



Pour une utilisation pratique, le graphique de droite de la figure 2.11.b se complétera de réseaux de parallèles indiquant des coordonnées remarquables. Les verticales de la figure 2.12 correspondent aux délimitations des années et les horizontales, à des dates de naissance tombant sur une délimitation d'année. Les obliques, pour leur part, matérialisent les âges exacts et doivent s'interpréter comme suit : évolution du moment de naissance en fonction du temps pour un âge donné. Vu que les coordonnées hôte et parasite ne suivent pas la même échelle, il faudra ajouter, sous l'échelle de temps, une autre pour l'âge de façon à pouvoir interpréter les diagonales. Ainsi, la diagonale la plus à droite commence au point 3 sur l'échelle des âges ; il s'agit de l'oblique correspondant l'âge exact 3 ans.

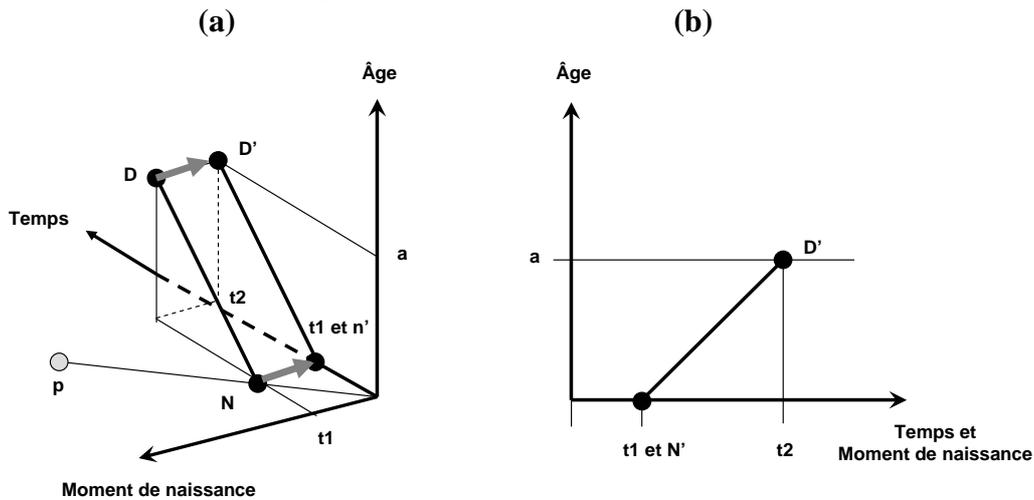
Figure 2.12. Trame sur un diagramme *temps-moment de naissance*



2.3.4.3. Diagramme temps-âge

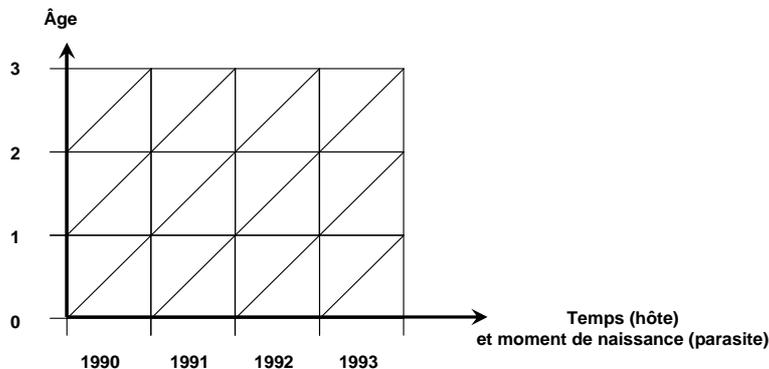
Dans ce troisième cas, la projection de la ligne de vie s'effectue perpendiculairement sur le plan *temps-âge* (cf. figure 2.13.a³). En conséquence, N' se confond avec $t1$ sur l'axe du temps ; pour sa part D est projeté sur le plan *temps-âge* à l'horizontale de a sur l'axe de l'âge et à la verticale de $t2$ sur l'axe du temps (cf. figure 2.13.b). Les lignes de vie apparaissent donc en position oblique. La coordonnée parasite est le moment de naissance qui s'établit sur l'axe du temps. C'est en suivant cette projection que l'on obtient le diagramme de Lexis dans sa version actuellement la plus utilisée et telle qu'elle a été expliquée dans le chapitre 1.

Figure 2.13. Projection sur le plan *temps-âge*



Pour une utilisation pratique, le graphique de droite de la figure 2.13.b se complétera de réseaux de parallèles indiquant des coordonnées remarquables. Les verticales de la figure 2.14 correspondent aux délimitations des années et les horizontales, à des âges exacts. Les obliques pour leur part matérialisent les dates de naissance tombant sur des délimitations d'années et doivent s'interpréter comme suit : évolution de l'âge en fonction du temps pour un moment de naissance donné. Vu que les coordonnées hôte et parasite suivent la même échelle, il ne faudra pas ajouter, sous l'échelle de temps, une autre pour le moment de naissance. Ainsi, la diagonale la plus à droite commençant début 1993 correspond à la ligne de vie d'un individu né le 1^{er} janvier 1993.

Figure 2.14. Trame sur un diagramme *temps-âge*



³ Cette figure 2.13 offre une autre perspective sur le volume initial. C'est maintenant le plan *moment de naissance-âge* qui est à... l'avant-plan et plus le plan *temps-âge*. C'est pour des raisons de confort en matière de perspective que ce choix a été posé.

2.3.4.4. Trois solutions : laquelle choisir ?

La réduction du nombre d'axes débouche donc sur trois versions concurrentes. Sur le plan mathématique, elles sont strictement équivalentes, de même que sur le plan de l'efficacité : toutes trois, elles autorisent la localisation univoque des observations en fonction de trois coordonnées démographiques ; toutes trois montrent des trajectoires forcées que les individus sous observation ne peuvent jamais quitter.

Sur le plan du détail, des différences se marquent entre les trois versions. Par exemple, dans la première méthode et dans la troisième, les points N' et D' sont distincts des points N et D alors que dans la deuxième N et N' se confondent. Autre exemple, dans la deuxième méthode, les analyses transversale et longitudinale s'opposent de manière plus radicale que dans les deux autres : une analyse transversale s'y fera en suivant un couloir vertical et une longitudinale, un couloir horizontal alors que dans les deux autres cas, des couloirs obliques (croissants pour la troisième et décroissants pour la première) interviennent en concurrence avec des couloirs verticaux.

Par ailleurs et plus important, lors de la projection sur le plan *temps-âge*, la ligne de vie conserve sa longueur initiale du diagramme en trois axes ; par contre, dans les deux autres versions, la longueur de la ligne de vie se modifie lors de la projection (longueur finale = longueur initiale * cosinus 45°), et en conséquence, la ligne de vie projetée a la longueur exacte de la vie de l'individu exprimée en unités figurant sur les axes : si un individu est mort à 3 ans et qu'une année est représentée par un centimètre, sa vie sera représentée par un ligne de 3 centimètres. Dans le cas de la projection *temps-âge*, la longueur sur le plan est équivalente à la longueur initiale dans le volume, mais ne correspond pas à l'âge au décès multiplié par la longueur d'une année : pour l'individu mort à 3 ans, la ligne de vie sur le plan mesure dans ce cas 3 centimètres * (1/cosinus 45°), soit 4,24 centimètres.

Dans les deux premiers cas, il s'agit d'une projection de la ligne de vie avec transformation et dans le troisième, d'une translation. On pourrait en conséquence différencier les appellations : « durée de vie » pour les deux premières méthodes et « ligne de vie », pour la troisième.

Ces considérations ne permettent pas de dégager un critère pour privilégier une méthode plutôt qu'une autre. Par contre, l'interprétation des obliques apparaissant dans les différentes constructions nous paraît à même de guider le choix. Pour rappel, dans les trois méthodes, une oblique s'interprète comme suit :

- évolution de l'âge en fonction du moment de naissance pour une date donnée, pour la projection sur le plan *moment de naissance-âge* ;
- évolution du moment de naissance en fonction du temps pour un âge donné, pour la projection sur le plan *temps-moment de naissance* ;
- évolution de l'âge en fonction du temps pour un moment de naissance donné, pour la projection sur le plan *temps-âge* ;

Même si ces trois expressions sont strictement équivalentes sur le plan théorique, et s'il fallait vraiment choisir une des versions, nous marquerions une préférence pour la troisième qui nous paraît « plus naturelle », pour reprendre une expression de R. Pressat⁴. En définitive, hors questions d'habitudes, il n'y a aucune raison de choisir une version plutôt qu'une autre. Toutefois, c'est la troisième version qui semble bien être actuellement la plus utilisée, ce qui n'empêche que certains recourent malgré tout préférentiellement à l'une des deux autres⁵.

⁴ Lettre à l'auteur, du 25 mars 1991.

⁵ À ce sujet, cf., par exemple, la publication suivante : W. Winkler (1969), pp. 90-91.

2.4. Conditions mathématiques de l'élaboration d'un diagramme de Lexis

Dans le point 2.3, l'élaboration du diagramme de Lexis vient d'être explicitée sur base de graphiques. Afin de progresser en matière de généralisation de l'utilisation des outils de la démographie, nous allons aborder cette question sur un plan purement mathématique. Comme déjà signalé, le diagramme de Lexis présente la double caractéristique suivante :

- bien que comportant seulement deux axes, trois variables peuvent y être utilisées pour la localisation des observations ;
- dans ce type de figure, chaque individu suit une trajectoire forcée dont il ne peut s'écarter en aucun cas.

Quelles conditions mathématiques doivent remplir les variables en jeu pour faire émerger ces deux caractéristiques ?

2.4.1. Conditions à la réduction du nombre d'axes

Si des observations se classent dans un système de référence à trois coordonnées, le plus immédiat, pour en illustrer la localisation, consiste à construire un diagramme à trois axes, un pour chacune des variables. Le diagramme en volume ainsi obtenu présente un grave défaut par rapport à un classique diagramme en plan : la difficulté de son utilisation à cause du recours à la perspective dans la troisième dimension. Le point 2.3.4 a montré via un système de projection comment passer de l'utilisation d'un volume à celle d'un plan pour la localisation des données démographiques. Nous allons maintenant aborder cette question d'un point de vue purement mathématique : quelle(s) condition(s) doivent remplir les variables en jeu pour opérer la réduction du nombre d'axes ? Nous allons d'abord traiter cette question de manière générale avant application au cas spécifique des variables du diagramme de Lexis classique.

Pour opérer la réduction, une seule condition suffit, à savoir que les trois variables concernées forment une combinaison linéaire⁶ du type :

$$z = (a*x) + (b*y)$$

où x , y et z symbolisent les trois variables et a et b , des coefficients de valeur quelconque. Il s'agit en fait de l'équation d'un plan. Autrement dit, les observations localisées en fonction de ces trois variables prendront place sur un plan à l'intérieur du volume défini par les axes dévolus aux trois variables (comme elles le font d'ailleurs dans le cas spécifique du diagramme de Lexis ; cf. point 2.3.3).

Supposons que l'on décide de localiser les observations sur le plan xy , défini par les axes dévolus à ces deux variables et donc de supprimer l'axe de la coordonnée z . Les coordonnées en x et en y se conservent sans aucune transformation ; par ailleurs, l'information contenue dans la coordonnée z est également conservée vu la relation entre les trois variables. Si les

⁶ En cas de relation de type multiplicatif, au lieu d'obliques, ce sont des hyperboles, avec comme asymptotes les axes des variables conservées, qui traduiront les valeurs de la variable dont l'axe a disparu. Dans ce cas, on ne peut obtenir une construction du type « diagramme de Lexis » (cf. annexe 2.1)

coordonnées x et y se traduisent de manière classique par des perpendiculaires à leur axe respectif, comment se traduit une valeur particulière de la coordonnée z , soit k , sur le plan xy ?

$$z = (a*x) + (b*y)$$

$$\Rightarrow k = (a*x) + (b*y)$$

$$\Rightarrow y = (k/b) - ((a/b)*x)$$

Une valeur particulière de la coordonnée z se traduira donc sur le plan xy par une droite de pente $(-a/b)$. La valeur de l'abscisse à l'origine de cette oblique se détermine en annulant y :

$$k = (a*x) + (b*y)$$

$$\Rightarrow k = (a*x) + (b*0)$$

$$\Rightarrow k = a*x$$

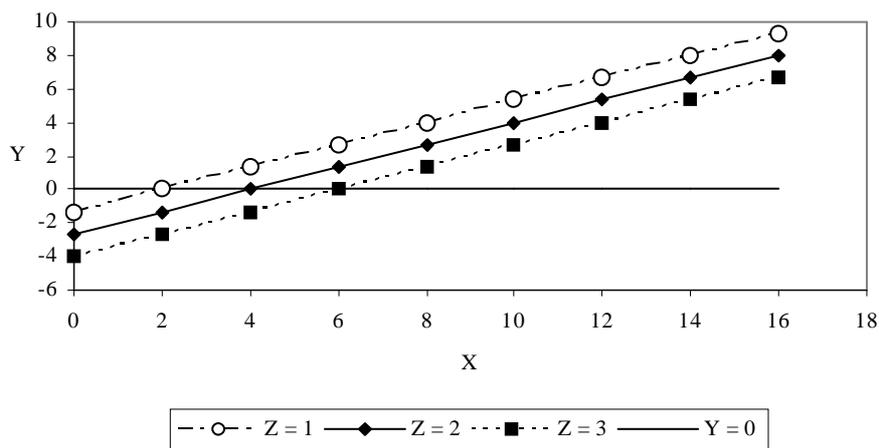
$$\Rightarrow k/a = x$$

Dès lors, si $a = 0,5$, $b = -0,75$ et $k = 2$, alors :

$$y = (2/-0,75) - ((0,5/-0,75)*x) = ((2/3)*x) - (8/3)$$

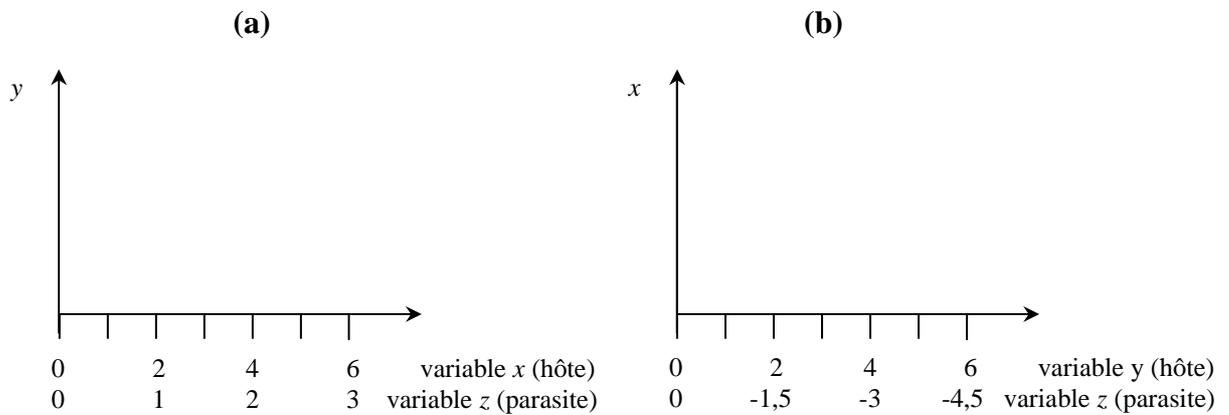
La valeur 2 de la coordonnée z se traduira donc par une oblique de pente croissante $2/3$ sur le plan xy . Par ailleurs, elle interceptera l'axe des abscisses à la valeur k/a , soit $2/0,5 = 4$. La figure 2.15 montre les obliques correspondant aux valeurs 1, 2 et 3 de la coordonnée z . Vu l'existence de ces obliques, cette dernière coordonnée peut donc remplacer une des deux autres pour localiser une observation sur le plan xy , et ce, sans perte d'information.

Figure 2.15. Suppression de l'axe z et report sur l'axe x



L'axe des abscisses de la figure 2.16.a comporte une double échelle⁷ : tout d'abord celle indiquant les valeurs de la coordonnée x et ensuite celle dévolue à la coordonnée z , obtenue en multipliant les valeurs x par $0,5$.

Figure 2.16. La double échelle de l'axe des abscisses



En bref, l'axe de la coordonnée z a été supprimé du volume initial ; cette coordonnée a « parasité » l'axe de la coordonnée x qui de ce fait s'affuble d'une double échelle. Une valeur particulière de la coordonnée z se traduit sur le plan xy par une oblique. Le volume initial a été réduit à un plan tout en conservant la possibilité d'utiliser les trois coordonnées. Cette réduction du nombre d'axes est possible simplement parce que les trois variables en jeu forment une combinaison linéaire.

On peut procéder au même exercice de réduction du volume en décidant de parasiter l'axe y . Dans ce cas, une valeur particulière z se traduira sur le plan xy par une oblique d'équation :

$$\begin{aligned} z &= (a*x) + (b*y) \\ \Rightarrow k &= (a*x) + (b*y) \\ \Rightarrow x &= (k/a) - ((b/a)*y) \end{aligned}$$

L'abscisse à l'origine de cette oblique s'obtiendra en annulant x :

$$\begin{aligned} k &= (a*x) + (b*y) \\ \Rightarrow k &= (a*0) + (b*y) \\ \Rightarrow k &= b*y \\ \Rightarrow k/b &= y \end{aligned}$$

Dès lors, si $a = 0,5$, $b = -0,75$ et $k = 2$, alors :

$$x = (2/0,5) - ((-0,75/0,5)*y) = 4 + 1,5*y$$

⁷ Plutôt qu'une double échelle sur l'axe des abscisses, on pourrait préférer introduire un troisième axe, mais cela nous a paru moins pratique, ne fût-ce que du point de vue de l'espace supplémentaire que nécessite la représentation de telles constructions. À ce sujet, cf. notamment G. Barys (1958), p. 71 ou G. Zeuner (1885), pp. 10-11.

Les obliques matérialisant les valeurs de z auront une pente de 1,5, soit l'inverse de la pente obtenue précédemment. Par ailleurs, la deuxième échelle à faire figurer en parallèle à celle des y pour repérer les valeurs de z se calculera en multipliant les valeurs de y par $-0,75$ (cf. figure 2.16.b).

La situation qui vient d'être décrite est équivalente à la première ; la figure 2.15 illustre le report de z aussi bien sur l'axe des x que sur celui des y . Toutefois, il nous a paru intéressant, pour explorer les différentes possibilités de réduction, de toujours mettre l'axe hôte (qui accueille la coordonnée parasite) en abscisse ; ainsi, sur la figure 2.16.b, l'axe y avec sa double échelle est en abscisse.

Le tableau 2.1 synthétise les six possibilités de réduction, chacun des trois axes pouvant alternativement se reporter sur l'un des deux axes restants. Pour rappel, l'axe hôte sera toujours mis en abscisse, même si ce n'est pas à strictement parler indispensable. Le tableau 2.2 reprend les six possibilités de construction de diagramme à deux axes au départ des trois variables utilisées dans le diagramme de Lexis classique, à savoir les variables *temps*, *âge* et *moment de naissance*. Dans ce cas, les coefficients a et b de la combinaison linéaire valent 1 ou -1 :

$$\text{âge} = \text{temps} - \text{moment de naissance}$$

En conséquence, la pente des obliques vaudra toujours 1 ou -1 ; autrement dit, la coordonnée parasite se traduira par des obliques croissantes ou décroissantes inclinées à 45° . La figure 2.17 illustre les six diagrammes ainsi obtenus. Chacun met en évidence par un grisé le début de la génération 1990.

Dans le tableau 2.2, la formule des abscisses à l'origine des obliques se définit en supposant que l'intersection des axes a pour coordonnées l'origine (la valeur 0) des deux variables dont les axes sont maintenus. Si ce n'est pas le cas, les formules doivent être modifiées pour tenir compte de cette translation d'axe(s). Ainsi, dans le cas 2, l'intersection des axes se produit à l'âge 0 et à la date du 1^{er} janvier 1990 pour le temps. La formule pour trouver l'abscisse à l'origine des obliques matérialisant le moment de naissance devient :

$$k = 1/1/90 - A$$

soit le temps à l'origine du graphique moins l'âge. L'échelle des moments de naissance est donc décroissante de gauche à droite alors que celle des âges est, de manière habituelle, croissante.

Même si la mention ne se retrouve que dans les cas 3 et 4, chacune des figures comporte une partie « vide », dans le sens où aucune observation ne pourra s'y localiser, à moins de vouloir utiliser des âges négatifs ! La valeur 0 pour la coordonnée *âge* marque donc la limite de cette zone vide. Dans les cas 1 et 6, cette zone vide se situe sous l'axe horizontal, prolongé à droite et à gauche⁸. Dans les cas 2 et 5, c'est toute la partie à gauche de l'axe vertical – prolongé haut et bas – qui restera obligatoirement vierge de toute observation. Dans le cas 3, c'est la zone sous la diagonale de l'âge 0 qui sera concernée ; et dans le 4, la partie au-dessus de cette ligne. À chaque fois, bien sûr, cette limite diagonale peut se prolonger.

⁸ Ces prolongements indiqueraient simplement que les données s'étendraient à une gamme plus vaste de temps et/ou de moment de naissance que la période reprise sur l'axe.

Tableau 2.1. La réduction des axes : six possibilités

Cas (1)	Ord. (2)	Abs. (3)	Axe supprimé (4)	Équation de l'oblique pour la valeur k de l'axe supprimé (5)	Pente des obliques	Abscisses à l'origine des obliques (6)
1	y	x	z	$k = (a*x) + (b*y)$ $\Rightarrow y = (k/b) - ((a/b)*x)$	$-(a/b) = -(0,5/(-0,75))$ $= 2/3$	$k = a*x = 0,5*x$
2	x	y	z	$k = (a*x) + (b*k)$ $\Rightarrow x = (k/a) - ((b/a)*y)$	$-(b/a) = -(-0,75/0,5)$ $= 3/2$	$k = b*x = -0,75*y$
3	x	z	y	$z = (a*x) + (b*k)$ $\Rightarrow x = (z/a) - ((b/a)*k)$	$(1/a) = (1/0,5)$ $= 2$	$z = b*k$ $\Rightarrow k = z/b = z/-0,75$
4	z	x	y	$z = (a*x) + (b*k)$	$a = 0,5$	$a*x = -b*k$ $\Rightarrow k = (a/-b)*x$ $= (2/3)*x$
5	z	y	x	$z = (a*k) + (b*y)$	$b = -0,75$	$a*k = -b*y$ $\Rightarrow k = (-b/a)*y$ $= 1,5*y$
6	y	z	x	$z = (a*k) + (b*y)$ $\Rightarrow y = (z/b) - ((a/b)*k)$	$1/b = 1/-0,75$ $= -4/3$	$z = a*k$ $\Rightarrow k = z/a$ $= 2*z$

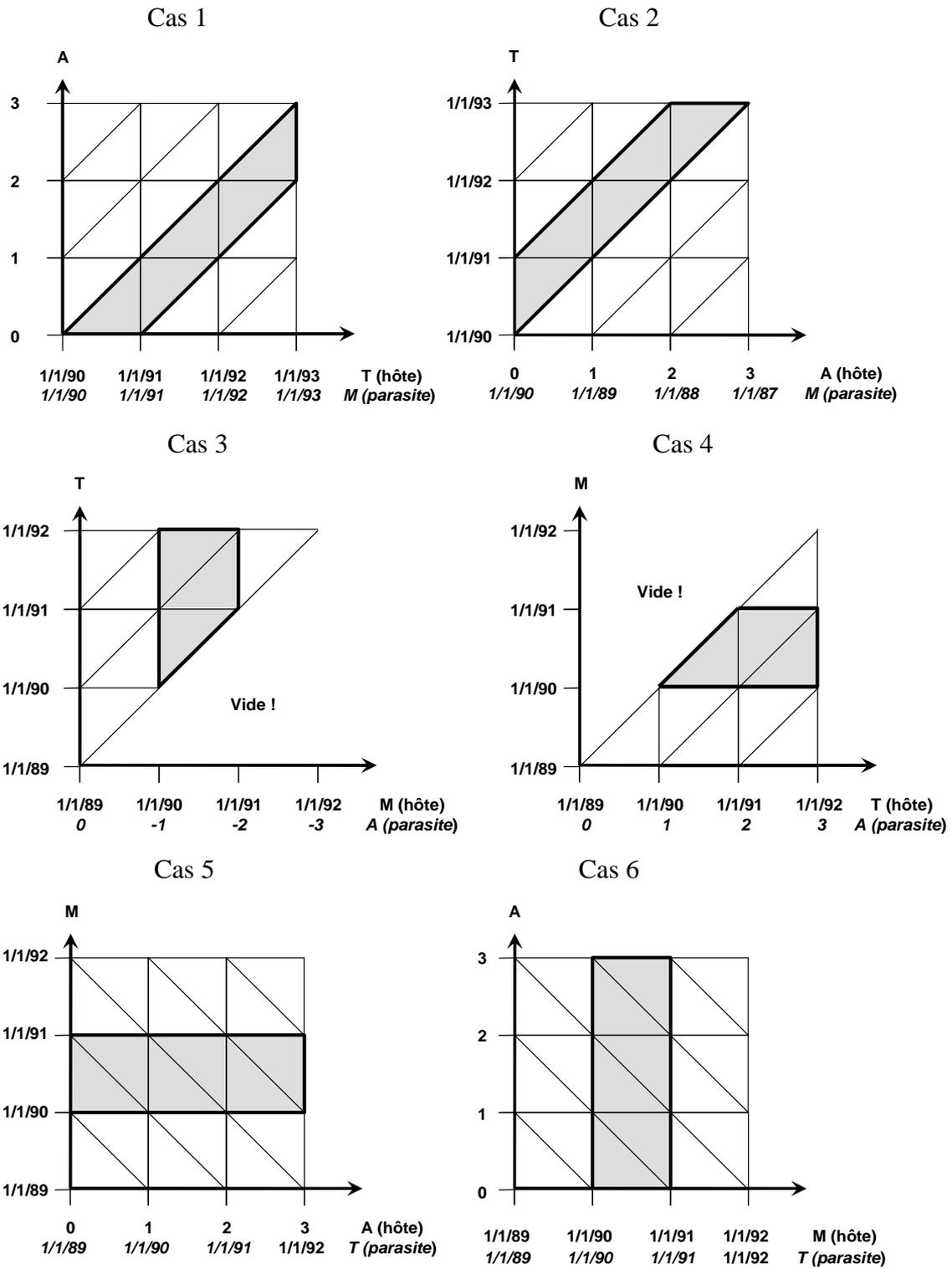
- (1) : équation générale : $z = (a*x) + (b*y)$, avec $a = 0,5$ et $b = -0,75$;
- (2) : axe des ordonnées ;
- (3) : axe des abscisses ;
- (4) : axe supprimé et reporté sur l'axe des abscisses ;
- (5) : k = une valeur constante de la variable supprimée et se matérialisant par une oblique ;
- (6) : étiquettes pour la variable dont l'axe est supprimé ; à faire figurer sous celles de la variable en abscisse.

Tableau 2.2. Les six possibilités avec les variables du diagramme de Lexis classique

Cas (1)	Ord. (2)	Abs. (3)	Axe supprimé (4)	Équation de l'oblique pour la valeur k de l'axe supprimé (5)	Pente des obliques	Abscisses à l'origine des obliques (6)
1	A	T	M	$k = (a*T) + (b*A)$ $\Rightarrow A = (k/b) - ((a/b)*T)$	$-(a/b) = -(1/(-1))$ $= 1$	$k = a*T = T$
2	T	A	M	$k = (a*T) + (b*k)$ $\Rightarrow T = (k/a) - ((b/a)*A)$	$-(b/a) = -(-1/1)$ $= 1$	$k = b*T = -A$
3	T	M	A	$M = (a*T) + (b*k)$ $\Rightarrow T = (M/a) - ((b/a)*k)$	$(1/a) = (1/1)$ $= 1$	$M = b*k$ $\Rightarrow k = M/b = -M$
4	M	T	A	$M = (a*T) + (b*k)$	$a = 1$	$a*T = -b*k$ $\Rightarrow k = (a/-b)*T$ $= T$
5	M	A	T	$M = (a*k) + (b*A)$	$b = -1$	$a*k = -b*A$ $\Rightarrow k = (-b/a)*A$ $= A$
6	A	M	T	$M = (a*k) + (b*A)$ $\Rightarrow A = (M/b) - ((a/b)*k)$	$1/b = 1/-1$ $= -1$	$M = a*k$ $\Rightarrow k = M/a$ $= M$

- (1) : équation générale : $A = T - M$, avec A, T et M désignant respectivement les variables *âge*, *temps* et *moment de naissance* et avec $a = 1$ et $b = -1$;
- (2) : cf. tableau précédent ; *idem* pour les autres remarques.

Figure 2.17. Les six diagrammes de Lexis possibles (cas en référence au tableau 2.2)



Certaines des figures proposées peuvent sembler inutilement compliquées, par exemple celle du cas 2 avec ses moments de naissance décroissants en regard des âges croissants, alors qu'il suffirait de reporter les moments de naissance sur l'axe du temps pour éviter cette présentation pour le moins inconfortable. En fait, une telle option nous ramènerait au cas 1, moyennant une inversion des axes. Notre but étant ici d'explorer sur le plan théorique toutes les possibilités disponibles, nous avons donc conservé des figures contre intuitives au risque de proposer des solutions sans portée utilitaire évidente.

Le cas 1 de la figure 2.17 représente le diagramme *temps-âge* dont la construction a été explicitée au point 2.4.3 ; le cas 4, le diagramme *temps-moment de naissance* (point 2.4.2) et le cas 6, le diagramme *moment de naissance-âge* (point 2.4.1). Les trois autres cas ne représentent que des formes dérivées des trois premiers.

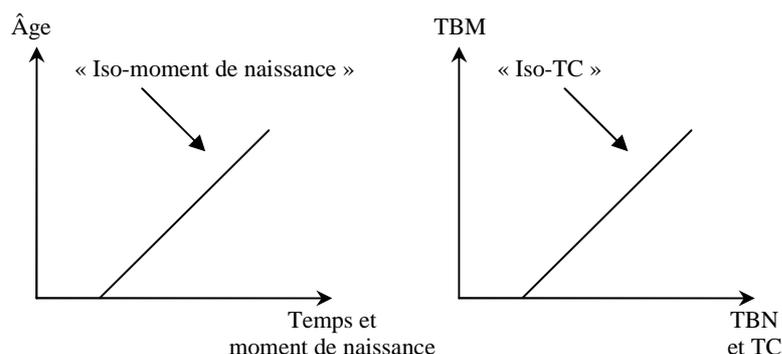
2.4.2. Conditions d'émergence des trajectoires forcées

Les derniers commentaires du point précédent pourraient laisser accroire que la réduction du nombre d'axes débouche inmanquablement sur un diagramme ayant toutes les caractéristiques d'un diagramme de Lexis classique. Il n'en est rien, comme l'a montré l'exemple du diagramme mettant en relation le *TBN* et le *TBM* et où une troisième coordonnée émergeait, à savoir *le taux de croissance* (cf. figure 2.2). Si ces trois variables respectent bien la condition de la combinaison linéaire ($TC = TBN - TBM$), les obliques – ni aucun autre réseau de droites – ne constituent pas des trajectoires forcées. En fait, pour obtenir une trajectoire forcée, il faut, en plus de la combinaison linéaire, qu'une des trois variables reste rigoureusement constante pour chaque individu sous observation, comme le fait le moment de naissance. Pour montrer ce point, nous allons utiliser en concurrence les deux exemples suivants :

- les variables *temps*, *âge* et *moment de naissance* avec suppression de l'axe du moment de naissance ;
- les variables *taux brut de natalité (TBN)*, *taux brut de mortalité (TBM)* et *taux de croissance naturelle (TC)* avec suppression de l'axe de ce dernier.

Selon l'explication développée au point précédent, ces deux cas sont semblables : ils se résument dans une relation du type $z = x - y$ si z désigne successivement le *moment de naissance* et le *TC* ; x , le *temps* et le *TBN* et y , l'*âge* et le *TBM*. La suppression de l'axe z entraînera l'apparition d'obliques croissantes de pente unitaire sur le plan xy pour matérialiser la coordonnée parasite (cf. figure 2.18).

Figure 2.18. Diagramme de Lexis classique et diagramme en taux



Cette apparente similitude cache en fait une différence fondamentale. Dans le diagramme de Lexis, les obliques représentent des « trajectoires forcées », dans le sens où tout ce qui concernera un individu est à localiser sur cette trajectoire ; à aucun moment, il ne pourra s'en écarter. Mathématiquement, cette notion de trajectoire forcée se traduit comme suit :

- la coordonnée *âge* se calcule comme suit : $A = T - M$.
- vu la constance du moment de naissance pour un individu, l'équation devient :

$$A = T - k,$$

soit l'équation d'une droite de pente 1. A est donc une fonction linéaire de T (et de M) du type $y = a*x + b$, avec $y = A$; $x = T$; $a = 1$ et $b = -M$. Et donc, tout ce qui concerne cet individu se localisera sur cette oblique.

Qu'en est-il des obliques du diagramme en taux ? La situation peut se décrire comme suit :

- la coordonnée TBM se calcule selon la formule suivante : $TBM = TBN - TC$
- malgré cette relation, comme aucune des trois variables n'est constante pour une unité d'observation, le TBM n'est pas une fonction linéaire des deux autres variables.

En conséquence, les obliques de ce diagramme en taux ne correspondent nullement à des trajectoires forcées : à un moment donné, un pays peut se trouver sur une oblique (correspondant à une valeur de TC) et à un autre moment, sur une autre. Le même raisonnement peut se tenir pour les horizontales ou les verticales. Il n'y a, sur ce diagramme, rien de « forcé » dans la trajectoire suivie pour un pays (*cf.* figure 2.3).

Sur le diagramme de Lexis, les trajectoires forcées apparaissent à cause de la coordonnée *moment de naissance*, qui est, par définition, invariable pour un individu. C'est la présence de ces trajectoires forcées qui rend le diagramme de Lexis particulier par rapport aux autres diagrammes ayant subi la réduction d'axes. Bien plus, c'est la présence de cette coordonnée qui génère le processus de cohorte et subséquemment celui de cohortisation lors de la phase d'analyse.

En conclusion, pour avoir un diagramme de Lexis, il faut trois variables reprises dans une combinaison linéaire (ce qui assure la réduction du nombre d'axes) et qu'une des trois soit constante pour une unité d'observation (ce qui assure la présence des trajectoires forcées). Les caractéristiques précises des figures obtenues via la réduction du nombre d'axes et, singulièrement, celles des obliques se modifieront selon la variable dont l'axe disparaît et selon le signe et les valeurs des coefficients a et b affectant les variables dans la relation qui les unit.

Nous avons montré que six diagrammes de Lexis pouvaient se déduire d'un jeu de trois variables bénéficiant des caractéristiques voulues. Parmi ces six variantes, trois types principaux se dégagent selon le plan sur lequel se projettent les lignes de vie. Toutes ces versions s'équivalent dans le sens où elles permettent la localisation des observations avec la même précision. Pour la suite de l'exposé, en cas de circonstances favorables à l'établissement d'un diagramme de Lexis, nous nous contenterons d'en explorer une seule version, les autres pouvant assez facilement s'en déduire. Par ailleurs, autant que possible, la version retenue pour l'exposé sera celle qui se rapprochera le plus d'un diagramme de Lexis classique, avec les variables jouant le rôle du temps, de l'âge et du moment de naissance se positionnant sur les mêmes axes.

2.5. Temps, âge et autres durées

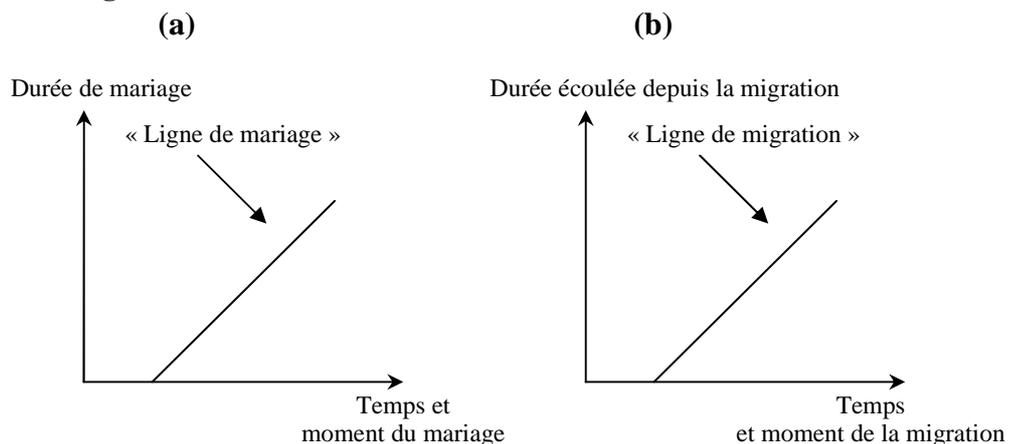
L'application du diagramme de Lexis se limite-t-elle au seul cas classique de l'âge, du temps et du moment de naissance ? D'autres combinaisons de variables peuvent-elles donner lieu à l'établissement de diagrammes ayant subi la réduction de trois à deux axes et présentant des trajectoires forcées ? Pour répondre à ces questions, il faut donc repérer d'autres ensembles de trois variables respectant les conditions mathématiques identifiées au point précédent⁹.

Dans un premier élan de généralisation, les exemples porteront exclusivement sur des variables temporelles au sens strict (s'exprimant en secondes, jours, mois, années...) ; ce n'est qu'ensuite que la généralisation portera sur des variables de nature quelconque (cf. chapitre 3). Par ailleurs, la démarche se fera de manière progressive dans le sens où la généralisation portera d'abord sur l'axe des ordonnées (point 2.5.1) puis, dans un deuxième temps, sur l'axe des abscisses (point 2.5.2). Ensuite, conséquence de cette généralisation du diagramme de Lexis, un troisième point servira à généraliser la notion de cohorte en cas de variables temporelles (point 2.5.3). Le dernier point opérera une ouverture en direction des diagrammes à plus de deux axes et trois variables (point 2.5.4).

2.5.1. Remplacement de l'âge par une autre durée

L'âge représente la durée écoulée depuis la naissance. En démographie, d'autres durées peuvent intervenir pour organiser les observations. Ainsi, en matière de fécondité, des analyses se basent sur un classement des naissances en fonction du temps et de la durée du mariage des mères concernées (cf. figure 2.19.a¹⁰ montrant uniquement les axes et une trajectoire forcée). En fait, une construction de ce type comporte une troisième coordonnée et des trajectoires forcées, ce qui le rend comparable à un diagramme de Lexis classique.

Figure 2.19. Diagrammes de Lexis avec durée de mariage ou durée écoulée depuis une migration



⁹ Pour des exemples, cf. notamment C. Blayo (1991), R. Pressat (1973), G. Wunsch et M. Termote (1978) ou C. Vandeschrick (1994).

¹⁰ Dans ce graphique et ceux qui suivent, nous avons opté pour la solution donnant le résultat le plus proche possible du diagramme de Lexis classique, notamment dans le choix de la variable dont l'axe est supprimé et dans celui de l'axe parasité. Par ailleurs, les coefficients a et b valant toujours l'unité, les obliques seront toutes inclinées à 45° et par surcroît croissantes, vu les choix opérés.

La troisième coordonnée, le moment du mariage, y tient le rôle joué par le moment de naissance dans le diagramme de Lexis classique. Par ailleurs, sur ce diagramme, un individu sous observation ne peut que suivre une trajectoire oblique croissante inclinée à 45° et commençant à la date de son mariage ; cette « ligne de mariage » représente bien pour lui une trajectoire forcée. Elle traduit l'évolution de la durée de son mariage (et non plus de son âge comme dans le diagramme de Lexis classique) en fonction du temps. Sur un tel graphique, une cohorte rassemblera non plus l'ensemble des individus nés la même année, mais celui des individus s'étant mariés durant la même année (soit une promotion de mariage).

Les trois variables en jeu respectent bien les deux conditions émises pour avoir la possibilité de construire un diagramme de Lexis. Tout d'abord, elles forment bien une combinaison linéaire :

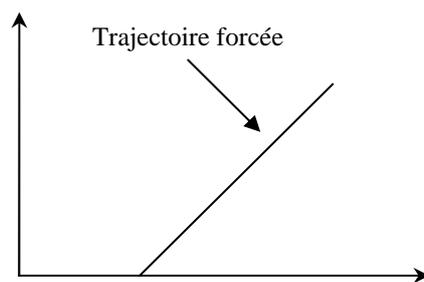
$$\text{durée du mariage} = \text{temps} - \text{moment du mariage}$$

Par ailleurs, la coordonnée *moment du mariage* demeure effectivement invariante individu par individu. Il en va de même pour les trois variables *temps*, *durée écoulée depuis une migration* et *moment de la migration*. La figure 2.19.b pourrait donner lieu à un commentaire semblable à celui qui vient d'être proposé pour la figure 2.19.a en remplaçant simplement *durée du mariage* par *durée écoulée depuis une migration* ; *moment du mariage* par *moment de la migration* et *ligne de mariage* par *ligne de migration*.

On l'aura compris, toute durée écoulée depuis un évènement quelconque peut suppléer l'âge. Son emploi simultané avec le temps entraîne automatiquement l'apparition d'une troisième coordonnée – le moment de l'évènement depuis lequel la durée se compte – dont la valeur demeure constante individu par individu. *De facto*, ces trois variables respecteront les deux conditions nécessaires à l'établissement d'un diagramme de Lexis (*cf.* figure 2.20)

Figure 2.20. Diagramme de Lexis général en cas d'une durée combinée avec le temps

Durée depuis un évènement



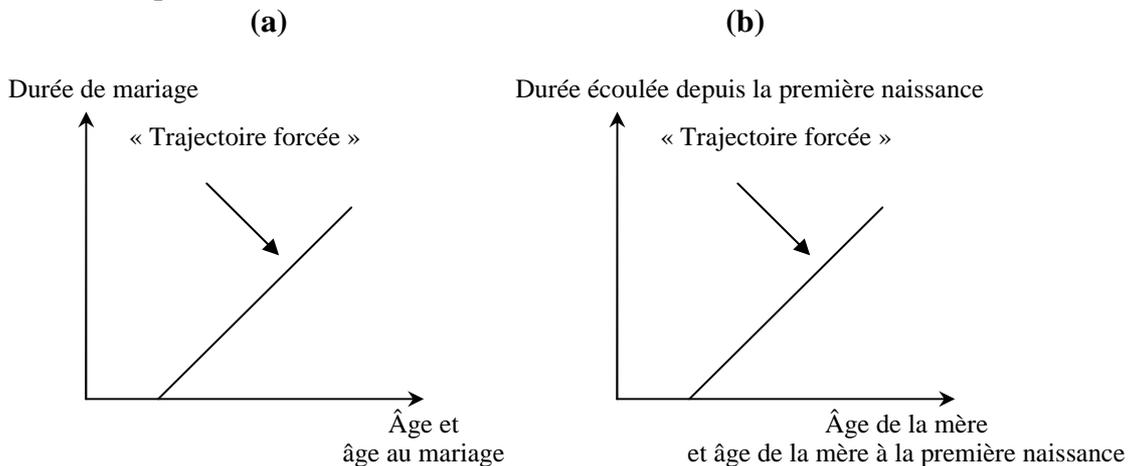
Temps (hôte)
et moment de l'évènement à l'origine de la durée (parasite)

2.5.2. Remplacement du temps par une autre durée

Le point précédent a mis en évidence une première généralisation dans l'utilisation du diagramme de Lexis en montrant que toute durée écoulée depuis un évènement quelconque pouvait prendre le relais de l'âge, variable occupant dans le cas le plus habituel l'axe des ordonnées. Qu'en est-il de la variable figurant sur l'axe des abscisses ?

Le temps-calendrier n'est, lui-même, rien d'autre qu'une durée et plus précisément, la durée écoulée depuis la naissance de Jésus¹¹. Le temps-calendrier peut très bien être remplacé par une autre durée, et ceci sans faire référence à d'autres calendriers (musulman ou copte, par exemple). Ainsi, sur la figure 2.21.a, l'axe des ordonnées est dévolu à la durée de mariage et celui des abscisses, à l'âge. Ce graphique comporte aussi une trajectoire forcée traduisant la coordonnée *âge au mariage* pour un individu. Une cohorte regroupera ici l'ensemble des individus qui se sont mariés au même âge révolu. Des données de fécondité, par exemple, pourront être localisées sur une telle construction (cf. tableau 2.3 et figure 2.22).

Figure 2.21. Diagrammes en âge et durée de mariage ou durée écoulée depuis la première naissance



Les trois variables en jeu respectent bien les deux conditions émises pour avoir la possibilité de construire un diagramme de Lexis. Tout d'abord, elles se réunissent bien dans une combinaison linéaire :

$$\text{durée du mariage} = \text{âge} - \text{âge au moment du mariage}$$

Par ailleurs, la coordonnée *âge au moment du mariage* demeure effectivement invariante individu par individu. La figure 2.21.b propose un autre exemple de diagramme de Lexis sans axe du temps. Dans cette figure¹² :

- l'axe des abscisses reçoit à la fois l'âge de la mère (coordonnée hôte) et l'âge de la mère à la naissance de son premier enfant (coordonnée parasite) ;
- la coordonnée parasite se traduit par des obliques qui sont en fait des trajectoires forcées ;

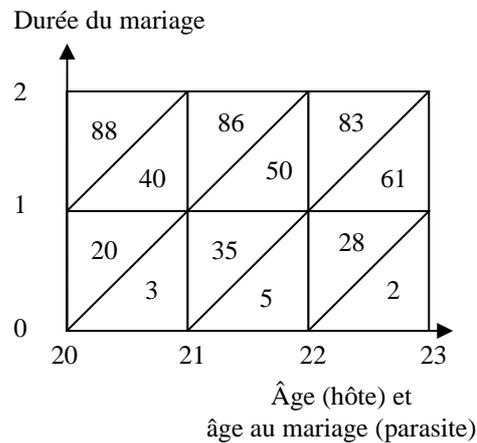
¹¹ Même si ce point est sans incidence par rapport au sujet de la thèse, notons qu'il est historiquement établi qu'une erreur a été commise à propos de l'estimation de la date de naissance de Jésus. Selon certains, il serait en fait né quatre ans avant le début de l'ère... chrétienne. À ce sujet, cf. par exemple le site <http://eocf.free.fr/text-calendriers.htm>.

¹² Cet exemple est repris de la publication suivante : C. Blayo (1991), p. 1382.

- l'axe des ordonnées reçoit les durées écoulées depuis la naissance du premier enfant ;
- une cohorte rassemble l'ensemble des femmes ayant eu un premier enfant à un âge révolu identique ;
- les trois variables en jeu respectent les conditions d'élaboration d'un diagramme de Lexis (combinaison linéaire et invariance de l'une d'entre elles).

Tableau 2.23 et figure 2.22. Naissances : triple classement

Âge à la naissance	Âge au mariage	Durée du mariage	
		0 an	1 an
20	20	3	-
	19	20	40
	18	-	88
21	21	5	-
	20	35	50
	19	-	86
	22	2	-
22	21	28	61
	20	-	83



Rem. : âges et durées en années révolues ; source : données fictives.

En conclusion, toute une série de durées peuvent jouer le rôle que le temps tient dans le diagramme de Lexis classique. La figure 2.24, à la fin du point suivant, montre un diagramme de Lexis à portée générale dans le sens où il indique clairement que le temps et l'âge de la construction classique peuvent se remplacer par d'autres durées.

2.5.3. Généralisation du concept de « cohorte » en cas de variables temporelles

En conformité avec le diagramme de Lexis dans sa version la plus habituelle, la cohorte se définit comme « l'ensemble des individus qui ont expérimenté un même évènement durant la même année ». Or, comme les points précédents l'ont montré, ce diagramme peut se construire au départ d'une grande variété de variables temporelles, et éventuellement même sans axe du temps-calendrier. Cette généralisation de l'utilisation du diagramme de Lexis entraîne *de facto* celle du concept de « cohorte »¹³.

Pour ce faire, différentes notions se révéleront fort utiles. Dans un premier temps, nous allons en préciser la définition et surtout en illustrer l'usage au moyen de deux exemples, à savoir, d'abord, celui du diagramme de Lexis classique avec le temps, le moment de naissance et l'âge et, ensuite, un autre où l'âge de la mère, l'âge de la mère à la naissance de son premier enfant et la durée écoulée depuis la première naissance se substituent respectivement aux trois variables du premier exemple.

¹³ Pour un exposé fouillé de la généralisation de la notion de cohorte, cf. la publication suivante : C. Blayo (1991). On pourra aussi consulter C. Vandeschrick (1994).

Définition des différents concepts :

- point initial : évènement de nature quelconque à partir duquel le temps peut être décompté ; il s'agit donc d'un instant conventionnel, à l'origine du décompte du temps ;
- durée de référence : durée (depuis un point initial) servant à la localisation des données ;
- évènement-origine¹⁴ : évènement que doit avoir subi un individu pour faire partie de la population sous observation ;
- cohorte : ensemble des individus ayant subi l'évènement-origine durant la même durée révolue de référence.

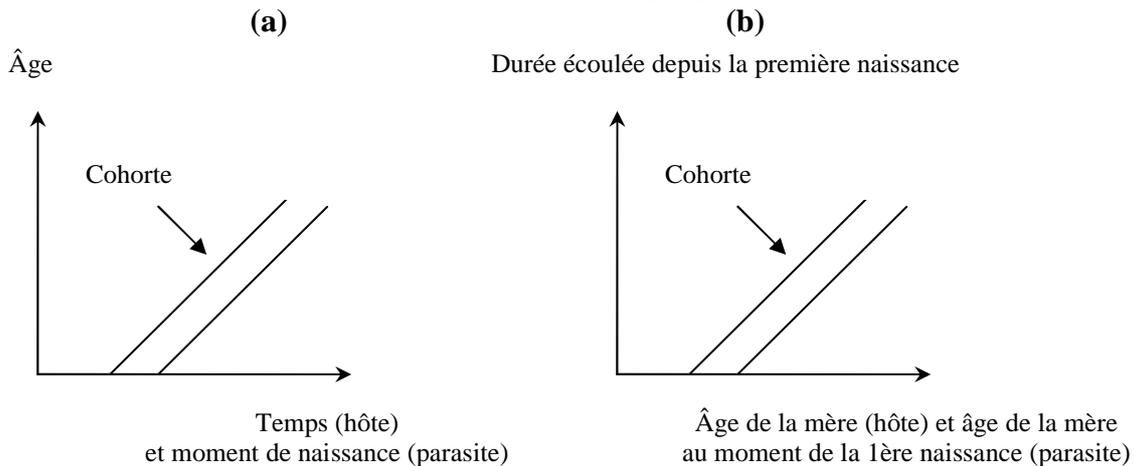
Dans nos exemples, l'évènement-origine est donc soit la naissance des individus, soit le fait pour les femmes d'avoir eu une première naissance (cf. tableau 2.4). Par ailleurs, pour entrer dans une cohorte précise, les individus doivent avoir subi cet évènement-origine soit pendant une année donnée, soit à un âge révolu donné. Le point initial de ces deux durées de référence diffère : le début de l'ère chrétienne d'un côté, la naissance des mères de l'autre. Enfin, l'âge du premier exemple cède le relais à la durée écoulée depuis la première naissance dans le second en tant que durée écoulée depuis l'évènement-origine.

Tableau 2.3. Les notions utilisées et leur application aux deux exemples

Notion	Symbole	Premier exemple	Deuxième exemple
Durée de référence (variable en abscisse)	<i>dr</i>	le temps	l'âge de la mère
Point initial de la durée de référence	<i>pi</i>	le début de l'ère chrétienne	la naissance des mères
Évènement-origine (fondateur de la cohorte)	<i>eo</i>	la naissance des individus sous observation	la naissance d'un premier enfant
Durée écoulée depuis l'évènement-origine (variable en ordonnée)	<i>deo</i>	l'âge	la durée écoulée depuis la première naissance
Durée de référence à l'évènement-origine (coordonnée parasite sur l'axe des abscisses)	<i>dreo</i>	le moment de naissance des individus sous observation	l'âge de la mère à sa première naissance

Dans le premier exemple, une cohorte regroupe l'ensemble des individus nés la même année et dans le second, celui des femmes ayant eu une première naissance au même âge révolu. Comme le montre la figure 2.23, une cohorte est délimitée par deux trajectoires forcées, c'est-à-dire par deux obliques matérialisant la coordonnée invariante individu par individu. Dans le premier exemple, la cohorte correspondant à une année donnée est limitée par les lignes de vie de deux individus nés respectivement en début et en fin de l'année concernée ; dans le second, la cohorte correspondant à un âge révolu donné est limitée par les trajectoires forcées de deux femmes ayant eu une naissance aux âges exacts encadrant l'âge révolu concerné.

¹⁴ Nous aurions préféré utiliser les expressions « évènement-origine » pour désigner le point initial et « évènement commun » pour désigner l'évènement-origine. En effet, le point initial correspond à un évènement à partir duquel la durée de référence est comptée ; c'est bien l'origine de cette dernière. L'évènement-origine désigne l'évènement que des individus doivent avoir connu à la même durée pour entrer dans une même cohorte ; c'est donc bien un évènement qui leur est commun. Toutefois, vu que le vocable « évènement-origine » est couramment employé en démographie pour désigner l'évènement qui fait entrer un individu dans une cohorte, nous lui avons conservé ce rôle et opté pour « point initial » pour désigner l'origine de la durée de référence.

Figure 2.22. Les deux exemples sous forme de graphiques

La définition d'une cohorte comprend deux éléments essentiels, à savoir l'évènement-origine qui la fonde et la durée de référence qui sert à le repérer. Pour être complète, la désignation d'une cohorte devrait donc comporter une double référence explicite :

- à son évènement-origine désignant l'évènement fondateur de la cohorte expérimenté par l'ensemble des individus la composant. En ne précisant que ce premier élément, on parlera, pour le premier exemple, d'une « cohorte de naissance » (ou « génération ») ; pour le deuxième exemple, d'une « cohorte de parité 1 » ; ou, autres possibilités en dehors des deux exemples, d'une « cohorte de mariage », d'une « cohorte de migration », d'une « cohorte de divorce »...
- à sa durée de référence désignant la durée qui a servi à définir l'intervalle pendant lequel l'ensemble des individus composant la cohorte doit avoir expérimenté l'évènement-origine. En ne précisant que ce deuxième élément, une « cohorte de calendrier » désignerait une cohorte pour laquelle l'évènement-origine s'est produit durant la même année (cas du premier exemple) ; une « cohorte d'âge », comme dans le deuxième exemple, désignerait une cohorte pour laquelle l'évènement-origine s'est produit au même âge (révolu) ; autre exemple encore, une « cohorte de durée du mariage » désignerait une cohorte pour laquelle l'évènement-origine s'est produit à la même durée (révolue) du mariage...

Un même évènement peut aussi bien jouer le rôle de point initial que celui d'évènement-origine, comme ce sera souligné dans les deux exemples de cohortes qui vont suivre :

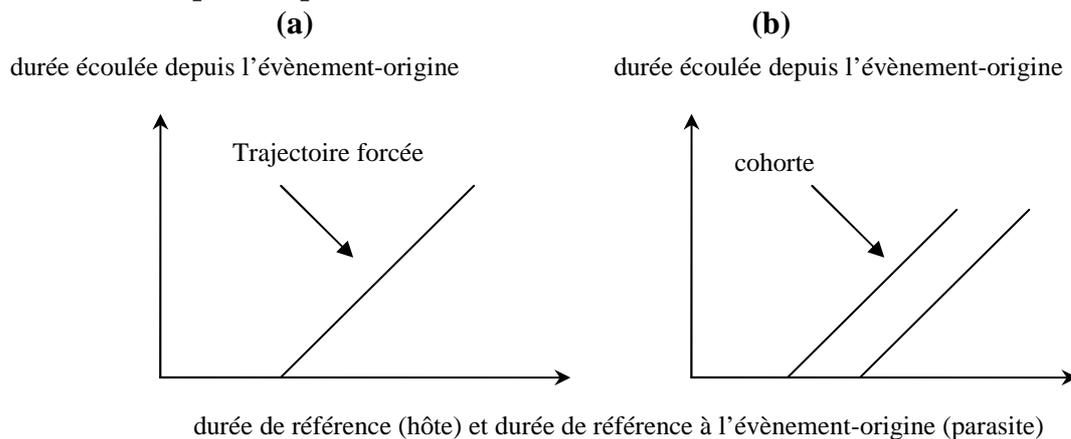
- ensemble des individus qui ont divorcé durant la troisième année après une migration (l'évènement-origine étant le divorce et une migration constituant le point initial de la durée de référence) ;
- ensemble des individus qui ont migré durant la deuxième année après un divorce (l'évènement-origine étant, cette fois, la migration et un divorce constituant le point initial de la durée de référence).

Finalement, une « cohorte de naissance de calendrier » ou « génération de calendrier » désignerait l'ensemble des individus nés durant la même année ; une « génération d'âge (des mères) », l'ensemble des individus nés de mères d'âge révolu identique et une « génération d'âge de mariage », l'ensemble des individus nés à une même durée révolue du mariage de leur mère. De même, une « cohorte de mariage de calendrier » ou « promotion de mariage de calendrier » désignerait l'ensemble des individus s'étant mariés la même année, et une « promotion de mariage d'âge », l'ensemble des individus mariés au même âge.

En théorie, pour définir complètement une cohorte, il faudrait toujours recourir à l'expression cohorte d'évènement-origine eo de durée de référence dr ¹⁵. Dans la pratique, comme la durée de référence correspond souvent au temps-calendrier, la mention la concernant est en règle générale purement et simplement omise. Une cohorte sera définie uniquement via le critère d'évènement : « cohorte de eo ». On parlera donc de cohorte de naissance, de mariage, de migration, de divorce, étant sous-entendu qu'il s'agit de cohorte de calendrier en termes de durée de référence. Pour les plus courantes, les cohortes de naissance, les démographes recourent à un seul mot : génération. On applaudira au passage le sens de la concision des démographes : « cohorte de naissance de calendrier » se traduit par le raccourci « génération » !

D'une manière générale, comment représenter sur un diagramme de Lexis une cohorte de eo de dr ? La solution la plus simple, car produisant un diagramme facile à mettre en parallèle avec un diagramme de Lexis classique, consiste à mettre la durée de référence en abscisse en tant que coordonnée hôte, en remplacement du temps (*cf.* figure 2.24) ; le moment de cette durée de référence qui a vu les individus subir l'évènement-origine en tant que coordonnée parasite (en remplacement du moment de naissance) et la durée écoulée depuis l'évènement-origine en ordonnée (en remplacement de l'âge).

Figure 2.23. Diagramme de Lexis avec durée de référence et évènement-origine quelconques



Des coordonnées ainsi définies ne pourront que respecter les conditions d'établissement d'un diagramme de Lexis :

- durée écoulée depuis l'évènement-origine = durée de référence – durée de référence à l'évènement-origine ;
- durée de référence à l'évènement-origine ne peut qu'être invariable individu par individu.

Sur un tel graphique, des trajectoires forcées (sous forme d'obliques croissantes inclinées à 45°) apparaîtront pour matérialiser la coordonnée invariante (*cf.* figure 2.24). Finalement, une cohorte rassemblera tous les individus ayant subi l'évènement-origine eo durant la même

¹⁵ C. Blayo désigne par « variable » la durée écoulée depuis l'évènement-origine ; par ailleurs notre durée de référence ne trouve pas de correspondant car toujours équivalente dans ses exemples au temps-calendrier (*cf.* : C. Blayo (1991), pp. 1394-1399). Il est à noter que cette dernière publication ne poursuit pas le même objectif que la nôtre.

durée révolue de référence dr ; sur le diagramme, elle sera délimitée par les deux trajectoires forcées encadrant la durée révolue concernée (cf. figure 2.24).

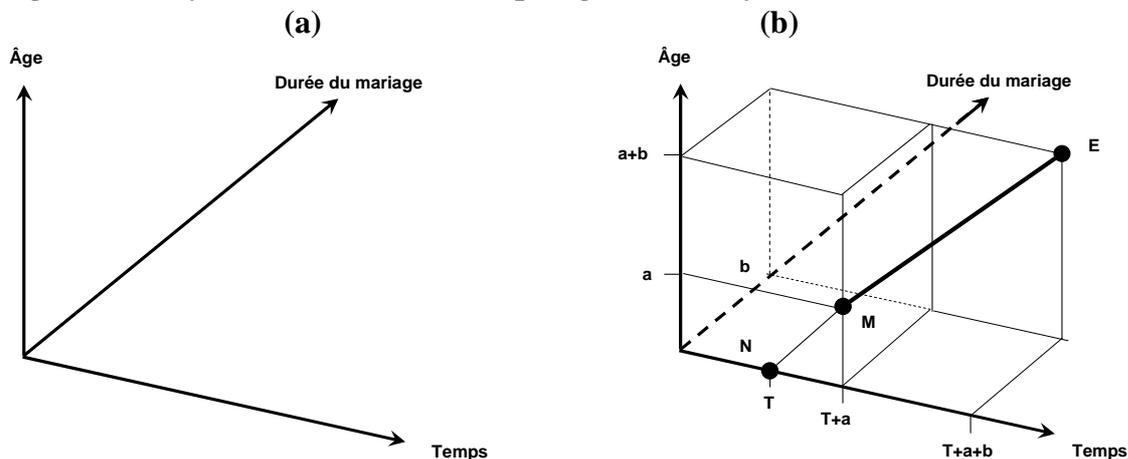
2.5.4. Des diagrammes de Lexis temporels à plus de deux axes

Tous les diagrammes de Lexis proposés jusqu'ici comptent deux axes tout en permettant d'utiliser trois coordonnées pour la localisation des observations. Est-il possible d'envisager l'élaboration de constructions à plus de deux axes et trois variables qui conserveraient les caractéristiques d'un diagramme de Lexis, dont la présence de trajectoires forcées ? La réponse à cette question est résolument affirmative¹⁶. Par ailleurs, bien que, sur un plan purement théorique, le nombre d'axes – égal ou supérieur à deux – peut être quelconque, nous restreindrons l'exposé qui va suivre au seul cas des graphiques à trois axes étant donné que, sur le plan pratique de la représentation, une telle limitation s'impose tout naturellement,

2.5.4.1. Construction des diagrammes de Lexis à trois axes

La figure 2.25.a représente un système à trois axes : le temps, l'âge et la durée écoulée depuis le mariage. C'est par facilité que nous avons décidé d'attribuer les axes à des durées de référence, selon la nomenclature adoptée précédemment ; le choix aurait très bien pu se porter sur des variables à valeur constante individu par individu (comme le moment de naissance, le moment du mariage ou l'âge au moment du mariage) sans modifier la suite du raisonnement, sauf dans le détail de sa présentation, qui en deviendrait en outre un peu moins aisée, du moins à notre avis.

Figure 2.24. Système à trois axes et repérage d'une trajectoire



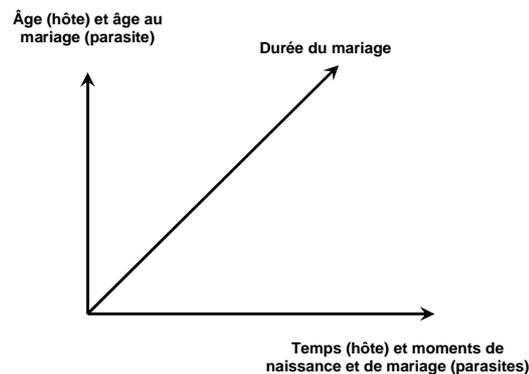
¹⁶ Keiding signale que Lexis a introduit, dès 1875, ce type de construction pour représenter des populations touchées par plusieurs changements d'état successifs, comme, par exemple, le mariage et le divorce (cf. N. Keiding (1990), p. 490). Il ne s'agissait donc pas de multiplier les coordonnées servant au repérage des observations comme ci-après, mais simplement de scinder en tronçons les lignes de vie, chaque tronçon correspondant à des états précis. Plus récemment, Wunsch a proposé des figures en volume allant exactement dans le sens de ce qui sera explicité ci-après (cf. G. Wunsch (1968), pp. 408-409). Tous les essais de Wunsch à ce sujet n'ont pas été publiés (cf. C. Vandeschrick (1994), p. 304).

La figure 2.25.B montre la ligne de vie d'une femme née au temps T ; mariée au temps $T+a$, soit à l'âge a et ayant subi un évènement quelconque E au temps $T+a+b$, soit à l'âge $a+b$ et à une durée du mariage b ¹⁷. La ligne de vie NME est en fait brisée en deux : avant le mariage, la durée du mariage vaut continuellement 0 et donc cette partie de la ligne de vie reste sur le plan *temps-âge* (segment NM). Après le mariage, la troisième coordonnée devient pleinement active et croîtra au même rythme que les deux autres ; en conséquence, la ligne de vie va progressivement pénétrer à l'intérieur du volume défini par les trois axes (segment ME).

En fait, une construction de ce type permettra de localiser des observations en fonction de six coordonnées. Tout d'abord, les trois premières sont simplement les variables figurant sur les axes. Par ailleurs, pour se localiser sur le plan *temps-âge* (en fait un diagramme de Lexis classique), la coordonnée du moment de naissance pourrait suppléer le temps ou l'âge. Cette coordonnée, parasite sur l'axe du temps, peut se déduire des deux autres : le moment de naissance se calcule en retirant l'âge du temps.

Les variables des deux autres plans peuvent aussi se combiner pareillement pour faire émerger une autre coordonnée du même type : premièrement, en retirant la durée du mariage du temps, on obtient la date du mariage ; en la retirant de l'âge, l'âge au mariage. Vu la nature des équations générant ces nouvelles variables, elles se traduiront chacune par des obliques croissantes sur le plan des deux variables les ayant engendrées.

Figure 2.25. Diagramme à trois axes et coordonnées parasites



Comme indiqué sur la figure 2.26, les moments de naissance et de mariage parasiteront de conserve l'axe du temps et l'âge au mariage, celui de l'âge¹⁸. Par ailleurs, ces trois nouvelles coordonnées présentent toutes la caractéristique d'être invariantes individu par individu¹⁹. En définitive, les six coordonnées²⁰ disponibles pour repérer un évènement sont le temps, l'âge, la

¹⁷ Si cet évènement n'est pas le décès, le segment pourra se prolonger au-delà du point E .

¹⁸ Les coordonnées parasites pourraient se reporter sur l'axe de l'autre variable ayant permis son calcul, mais cela aurait entraîné l'apparition d'échelles différentes à faire figurer sous les axes, ce que nous avons voulu éviter.

¹⁹ C'est parce que les axes ont été *a priori* attribués aux durées de référence que les trois variables parasites présentent cette caractéristique d'invariance ; tout autre choix initial pour les axes ferait apparaître des variables en variation continue individu par individu.

²⁰ D'une manière générale, pour autant que les n variables figurant sur les axes aient les caractéristiques voulues (possibilité de se combiner deux à deux dans une relation additive pour faire émerger une variable supplémentaire) et si n est égal ou supérieur à 2, le nombre de coordonnées effectivement utilisables en plus de celles figurant sur les axes correspond au nombre de combinaisons de deux éléments à choisir parmi n :

- si $n = 2$ (cas classique), une variable supplémentaire : $(2!/(2!0!))$;
- si $n = 3$, trois variables supplémentaires : $(3!/(2!1!))$;
- si $n = 4$, six variables supplémentaires : $(4!/(2!2!))$.

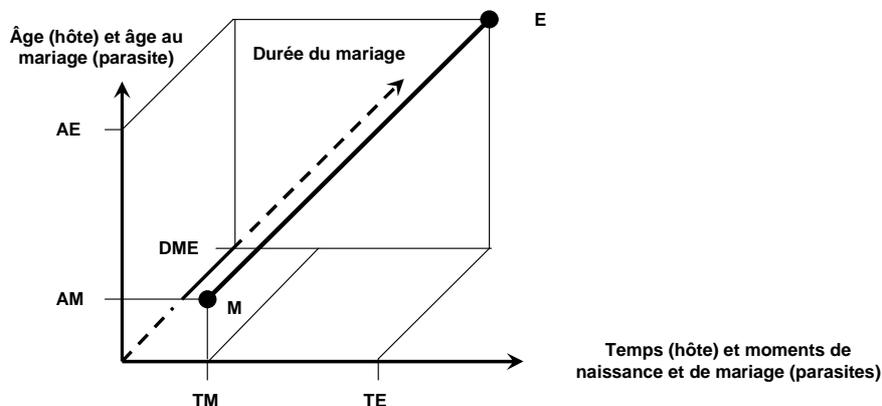
G. Wunsch ((1968), p. 409) a procédé à des calculs de ce type, mais en les faisant porter sur le nombre de plans nécessaires en fonction du nombre d'évènements-origines.

durée du mariage, le moment de naissance, le moment du mariage et l'âge au moment du mariage. Toute combinaison de trois coordonnées choisies parmi les six disponibles permettra de localiser une observation. En effet, chacune se traduira dans le volume par un plan soit horizontal, soit vertical, soit oblique (selon qu'elle se traduit elle-même par une horizontale, une verticale ou une oblique sur le plan de ses axes) ; l'intersection de ces plans définira l'emplacement de l'observation.

2.5.4.2. Localisation de trajectoires et d'observation sur des diagrammes à trois axes

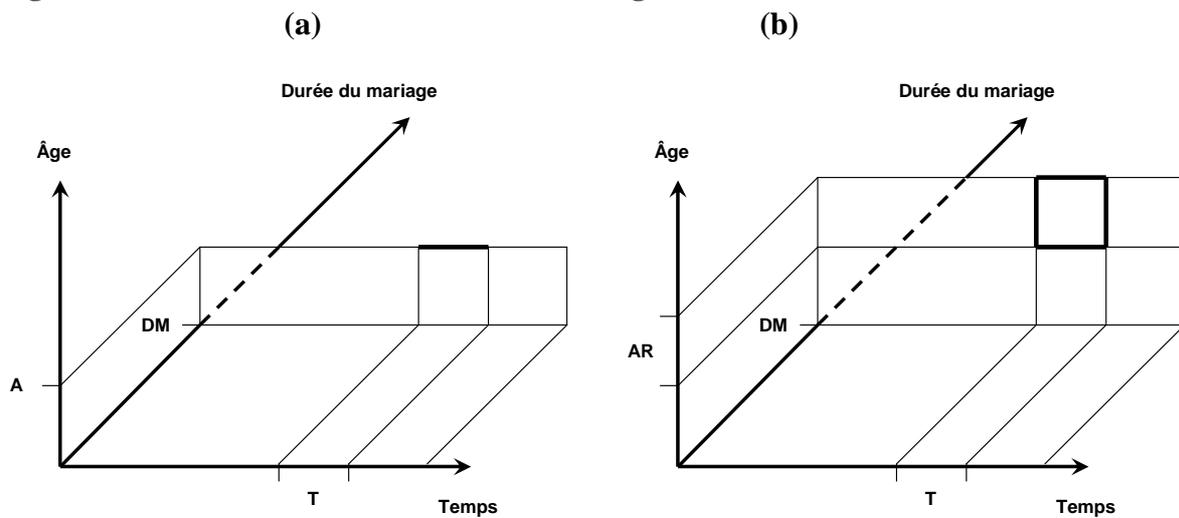
La figure 2.27 est constituée d'un extrait de la figure 2.25 ; seule la partie de la ligne de vie postérieure au mariage a été prise en compte, soit le segment ME . Le moment du mariage y est désigné par TM (et non plus $T+a$) ; le moment de l'évènement E , par TE (et non plus $T+a+b$) ; les âges au mariage et à l'évènement E , respectivement par AM et AE (et non plus a et $a+b$) ; finalement, la durée du mariage à l'évènement par DME (et non plus b). Le trait qui unit les points M (le mariage) et E (un évènement quelconque) représente une trajectoire forcée : tout ce qui affectera l'individu en cause entre les moments du mariage (TM) et de l'évènement (TE) devra se localiser sur cette oblique²¹. Si sur le diagramme de Lexis classique à deux axes, la ligne de vie est déterminée sur base du seul point de naissance, dans un diagramme à trois axes comme celui de la figure 2.27, la position exacte de la ligne de mariage – et singulièrement son point initial – dépend en plus de la date du mariage : que celui-ci soit plus précoce ou plus tardif et la trajectoire forcée de la vie maritale occupera une position respectivement plus basse ou plus haute, plus à gauche ou plus à droite dans le volume. Il n'empêche que toutes ces trajectoires seront parallèles comme les lignes de vie dans un diagramme de Lexis classique.

Figure 2.26. Localisation d'une ligne de mariage sur un diagramme à trois axes



Outre des évènements individuels et à l'instar du diagramme de Lexis classique, des ensembles d'individus ou d'évènements pourront bien évidemment aussi se localiser sur un diagramme en trois axes. Ainsi, le trait horizontal en gras de la figure 2.28.a localise l'effectif des individus ayant fêté leur anniversaire A et leur anniversaire de mariage DM durant l'année T ; pour sa part, le carré de la figure voisine localise l'effectif des individus ayant fêté leur anniversaire de mariage DM à l'âge révolu AR durant l'année T . Un effectif correspondra à un trait (cas a) s'il se définit à l'aide d'une coordonnée de durée (l'année) et de deux autres ponctuelles (l'âge exact et l'âge exact du mariage), et à une surface (cas b) si sa définition s'appuie sur deux coordonnées de durée (année et âge révolu) et une troisième ponctuelle (âge exact du mariage).

²¹ Par rapport aux figures précédentes, on n'a pris en compte que la partie de la ligne de vie après le mariage, c'est-à-dire celle où toutes les coordonnées sont actives.

Figure 2.27. Localisation d'effectifs sur un diagramme en trois axes

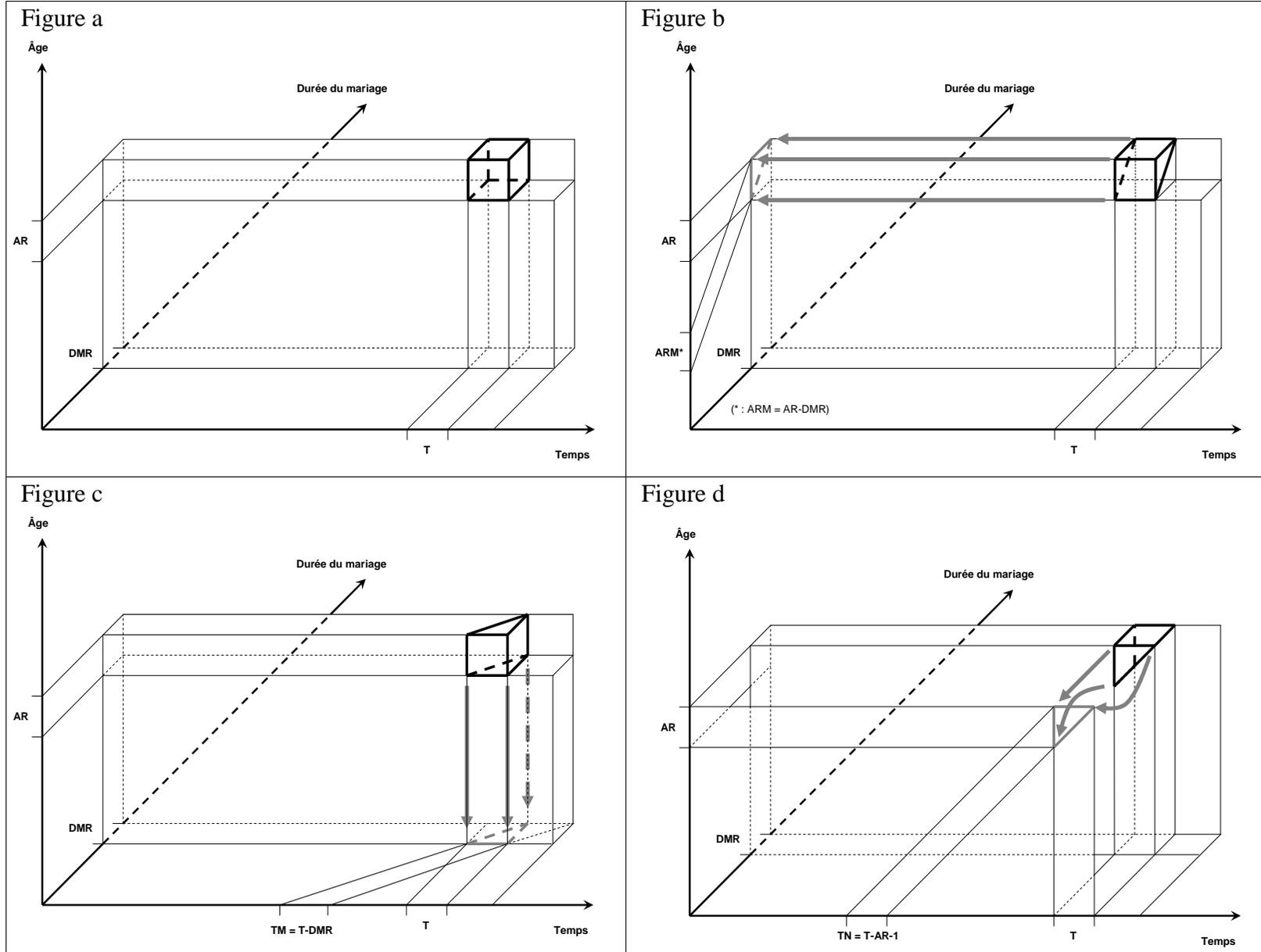
Si, dans le diagramme de Lexis classique, les ensembles d'évènements occupent toujours des surfaces, ils occupent toujours des volumes dans le diagramme à trois axes, car ils se définissent toujours à l'aide de trois coordonnées non ponctuelles. Ainsi, le cube de la figure 2.29.a regroupe les évènements s'étant produits durant l'année T pour des individus d'âge révolu AR et de durée révolue de mariage DMR .

Ces évènements concernent deux cohortes de mariage d'âge : le prisme de la figure 2.29.b localise le sous-ensemble des évènements s'étant produits pour des individus mariés à l'âge révolu ARM (avec $ARM = AR - DMR$)²² ; le deuxième sous-ensemble (non représenté) concerne pour sa part des individus mariés à l'âge révolu $ARM - 1$ (en supposant donc que AR , T et DMR représentent chacun un an).

Les évènements du cube initial concernent deux promotions de mariage. Le prisme de la figure 2.29.c localise parmi le total des évènements du cube ceux survenus pour des individus mariés durant l'année TM (avec $TM = T - DMR$), le solde concernant la promotion $TM - 1$. Finalement, les évènements du cube relèvent de deux générations : la génération TN (avec $TN = T - AR - 1$), comme l'indique la figure 2.29.d ainsi que la génération suivante. Sur cette dernière construction, des lignes courbes apparaissent ; elles sont nécessaires pour éviter de confondre certains traits du prisme ou du triangle avec certaines flèches de projection.

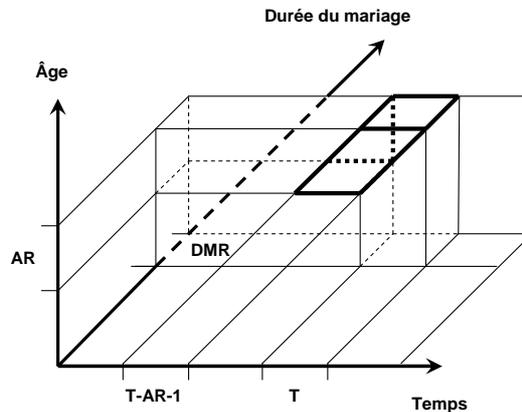
²² Les âges au mariage supposés par ce graphique suggèrent une nuptialité pour le moins précoce. C'était indispensable pour pouvoir exécuter ces figures sans translation d'axes !

Figure 2.28. Ensemble d'événements sur un diagramme en trois axes



Au départ d'un diagramme à trois axes, on pourrait aussi identifier des parallélépipèdes rectangles et autres formes définies au moyen d'une, deux ou trois²³ coordonnées invariantes individu par individu. À titre indicatif (et sans chercher à illustrer tous les cas de figures envisageables...), la figure 2.30 propose le parallélépipède regroupant les événements s'étant produits pour la génération $T-AR-1$, à l'âge révolu AR et pour une durée de mariage DMR . Ces événements se sont produits durant deux années successives $T-1$ et T .

Figure 2.29. Localisation d'un parallélépipède sur un diagramme à trois axes



2.5.4.3. Diagrammes à trois axes et réduction du nombre d'axes

Les graphiques à trois axes du point précédent pourraient se réduire en diagrammes en plan par projection. Toutefois, au contraire de ce qui avait été constaté pour l'élaboration du diagramme de Lexis classique au départ du volume (point 2.4), l'opération de réduction s'accompagnera ici de la perte d'une part de l'information initiale contenue sur le diagramme à trois axes²⁴.

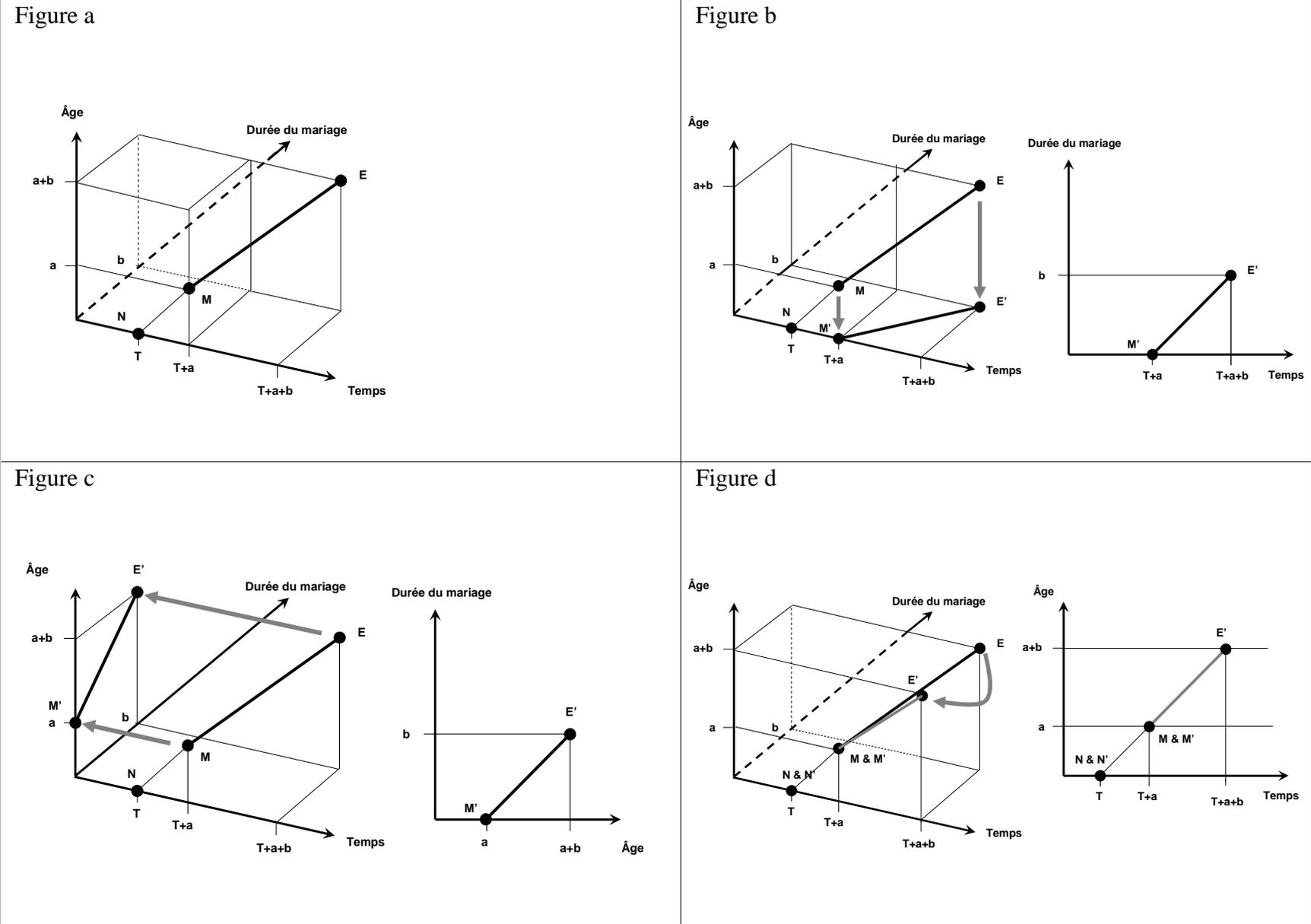
La figure 2.31.a montre la trajectoire d'un individu né en T , marié en $T+a$ et affecté en $T+a+b$ par un événement quelconque E . Sa vie maritale, trajectoire forcée, sera représentée par le segment ME , éventuellement prolongé au-delà de E , si l'événement en cause n'est ni le décès, ni une rupture d'union²⁵. Cette trajectoire peut être projetée sur l'un des trois plans définissant le volume. Ainsi, la figure 2.31.b montre la projection de la ligne de mariage sur le plan *temps-durée de mariage*. Le diagramme à deux axes obtenu présente toutes les caractéristiques d'un diagramme de Lexis ; toutefois, suite à la projection, l'information concernant l'âge est perdue : il est impossible sur cette figure de donner le moindre renseignement à propos de l'âge des individus concernés par un événement.

²³ L'annexe 2.2 reprend l'exercice perspectif si les trois coordonnées invariantes devaient intervenir sous forme unitaire.

²⁴ À ce sujet, cf. G. Wunsch (1968), pp. 408-409.

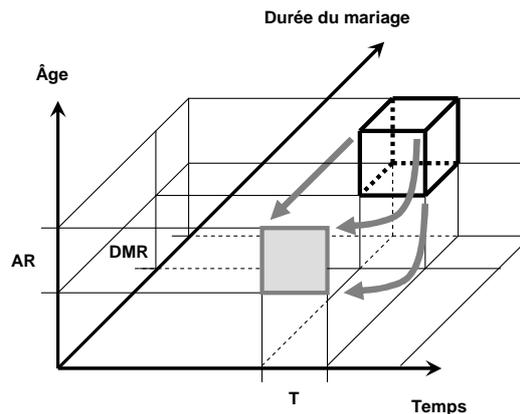
²⁵ Comme nous l'avons déjà précisé, la ligne de vie – que l'on pourrait aussi dénommer ligne d'âge, ce qui serait plus approprié, vu ce qui va suivre – pourrait se prolonger au-delà de la mort. De même, la durée écoulée depuis le mariage peut très bien se prolonger au-delà d'une éventuelle rupture d'union ; la durée écoulée après le mariage continue d'augmenter, même après la fin de l'union. Dans ce cas, il faudrait sans doute donner un autre nom à la trajectoire : ligne de durée écoulée depuis le mariage en remplacement de ligne de mariage. Il est à noter que cette ligne montre aussi l'évolution de l'âge ou du temps !

Figure 2.30. Diagramme à trois axes et projection



La projection de la trajectoire pourra s'exécuter sur les deux autres plans (cf. figure 2.31.c et d). À chaque fois, un diagramme de Lexis à deux axes apparaîtra ; à chaque fois, l'information indiquée initialement par l'axe supprimé disparaît. La projection sur un plan peut aussi concerner des ensembles d'évènements²⁶ ou d'individus. Ainsi, sur la figure 2.32, le cube contenant les évènements survenus durant l'année T pour des individus d'âge révolu AR et pour une durée de mariage DMR est projeté sur le plan *temps-âge*, autrement dit sur un diagramme de Lexis classique. Comme il sied sur ce type de figure, les évènements d'une année et d'un âge révolu s'y localisent dans un carré, mais sans plus aucune possibilité de distinction à propos de la durée du mariage.

Figure 2.31. Projection d'un volume sur un des plans d'un diagramme à trois axes



Ce type de projection pourra se faire au départ de circonstances différentes (sur un des deux autres plans ; au départ de prismes, de parallélépipèdes...), mais toujours avec la perte d'une partie de l'information initiale. Cette perte d'information souligne une différence fondamentale entre la réduction du nombre d'axes telle qu'opérée ici par rapport à celle utilisée lors de l'élaboration du diagramme de Lexis classique. D'où vient cette différence ?

Dans le cas du diagramme de Lexis classique, les trois variables du volume initial (temps, âge et moment de naissance) formaient une combinaison linéaire ; par contre, dans le cas qui nous occupe maintenant, il n'en va plus de même : le temps, l'âge et la durée de mariage ne peuvent se combiner dans une relation de ce type. Cette différence se reflète dans la distribution des trajectoires forcées : au contraire du diagramme de Lexis classique où les lignes de vie n'occupaient qu'un plan dans le volume initial, dans le cas de la figure 2.32, les trajectoires forcées peuvent se localiser n'importe où dans le volume²⁷, ce qui empêche de réduire le volume à un plan sans perte d'information. La condition de combinaison linéaire n'étant pas remplie, la réduction d'axe ne pourra s'opérer, du moins sans perte d'information.

Certaines circonstances précises d'analyse pourraient donner lieu à l'établissement de diagrammes de Lexis à trois axes (ou plus), éventuellement simplifiés, dans une deuxième étape, par réduction du nombre d'axes. Ainsi, Blayo propose notamment d'analyser les avortements de femmes mariées selon la durée de leur mariage tout en distinguant en plus des sous-cohortes selon leur âge au mariage²⁸. Pour représenter ces données, il faut utiliser un diagramme de Lexis à trois axes. Parmi les différentes formes que peut prendre ce

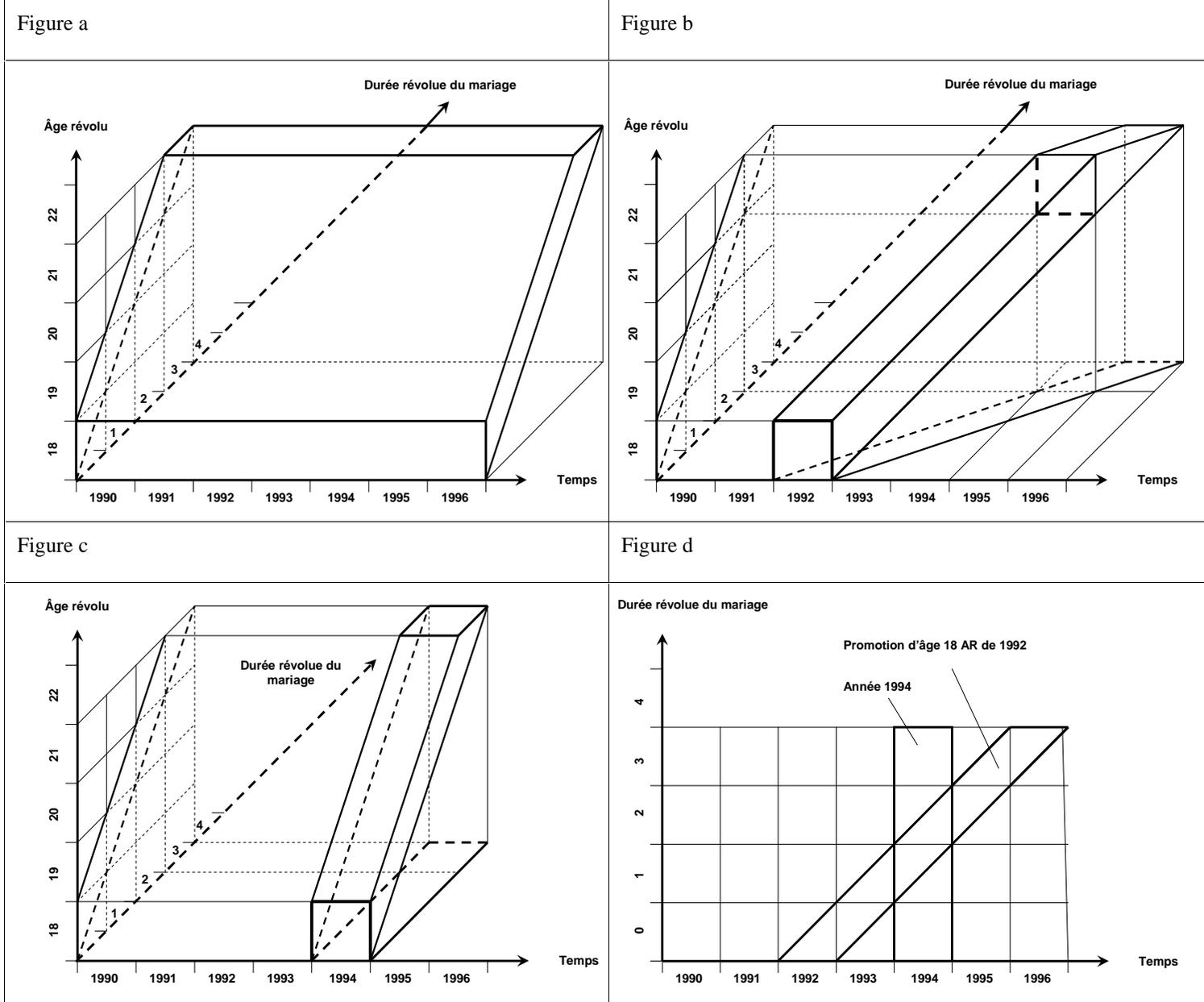
²⁶ La figure 2.29 recourait déjà à ce type de projection, mais dans un autre but : déterminer les cohortes concernées par un ensemble d'évènements. Il ne sera donc pas nécessaire de fortement développer ce point.

²⁷ « N'importe où », c'est trop dire dans le cas qui nous sert d'exemple : on voit mal quelqu'un se marier avant d'avoir atteint un certain âge : toute la partie du volume sous le plan correspondant à l'âge minimum au mariage restera libre de toute trajectoire de mariage.

²⁸ Cf. tableau 6, exemple 3 dans C. Blayo (1991), pp. 1394-1395.

diagramme, optons pour celle où le temps, l'âge et la durée du mariage occupent les trois axes en tant que coordonnées hôtes.

Figure 2.32. Diagramme à plus de deux axes et sous-cohortes



Dans un tel volume, les avortements survenus pour des femmes s'étant mariées à 18 ans révolus occupent une tranche oblique (cf. figure 2.33.a). À l'intérieur de cette tranche, on peut isoler la zone dévolue aux seules femmes s'étant mariées à 18 ans révolus durant l'année 1992, soit la promotion d'âge 18 ans révolus de l'année 1992 (cf. figure 2.33.b). Cette zone se termine en biseau car les observations y présentées ne dépassent pas l'âge de 22 ans exacts. La figure 2.33.c montre, pour sa part, la zone de la tranche initiale dévolue à tous les événements s'étant produits durant la seule année 1994 ; ces données concernent les promotions d'âge 18 ans révolus de différentes années. Ces deux dernières figures permettent d'opposer les optiques transversale et longitudinale : les données de la première déboucheront sur une analyse longitudinale et celles de la seconde, sur une analyse transversale.

En fait, cette situation d'analyse peut se représenter d'une manière plus simple après réduction d'axes, après suppression de celui de l'âge. La figure 2.33.d a été obtenue par projection sur le plan *temps-durée du mariage*, projection déjà suggérée sur les figures précédentes. La promotion d'âge 18 ans révolus de 1992 y occupe un couloir oblique, selon une perspective longitudinale et l'année 1994, un couloir vertical selon une perspective transversale.

Dans le cas présent, la réduction d'axes s'opère sans perte d'information car, dès le départ, la variable dont l'axe va disparaître a été en partie neutralisée : ne sont prises en considération que les seules femmes s'étant mariées à 18 ans ; celles qui se sont mariées plus ou moins jeunes restent exclues de la tranche initiale. En conséquence, un indice synthétique mesuré dans la cohorte de la figure 2.33.d quantifiera l'activité du phénomène le long de la promotion d'âge 18 ans révolus en 1992 ; de même, celui qui le serait dans le couloir vertical, la quantifierait pour l'année 1994, en recoupant les promotions d'âge 18 ans révolus de différentes années.

Prétendre que la réduction d'axes s'opère sans perte d'information, c'est trop dire, sauf si la précision à propos des générations est sans importance. En effet, les femmes de la promotion d'âge 18 ans révolus de 1992 relèvent de deux générations, celles de 1973 et de 1974. Cette distinction pourrait être maintenue à conditions de multiplier les axes et donc les niveaux de sous-cohortes.

Dans son article, Blayo propose à titre d'exemple une analyse portant sur les naissances légitimes de rang 2 parmi les femmes nées en 1960, mariées en 1980 et ayant eu une première naissance en 1982²⁹. Pour illustrer ces données, il faudrait imaginer un diagramme à quatre axes pour figurer le temps, l'âge, la durée du mariage et la durée écoulée depuis la première naissance. En fixant la génération, la promotion de mariage et l'année de la première naissance, les observations pourraient se projeter sur un diagramme de Lexis classique *temps-âge*.

Ces deux exemples montrent, à notre avis, le caractère relativement inopérant des diagrammes de Lexis à plus de deux axes qui, soit restent difficiles à lire (alors que la vocation du diagramme de Lexis est de faciliter la compréhension de l'agencement des données !), soit se révèlent purement et simplement impossibles à représenter. La solution consiste à revenir à un diagramme à deux axes en fixant la valeur de certaines coordonnées.

Techniquement réalisable en cas de trois axes, ce type de construction peut se révéler utile pour bien montrer, par exemple, que les femmes mariées à un âge révolu durant une année

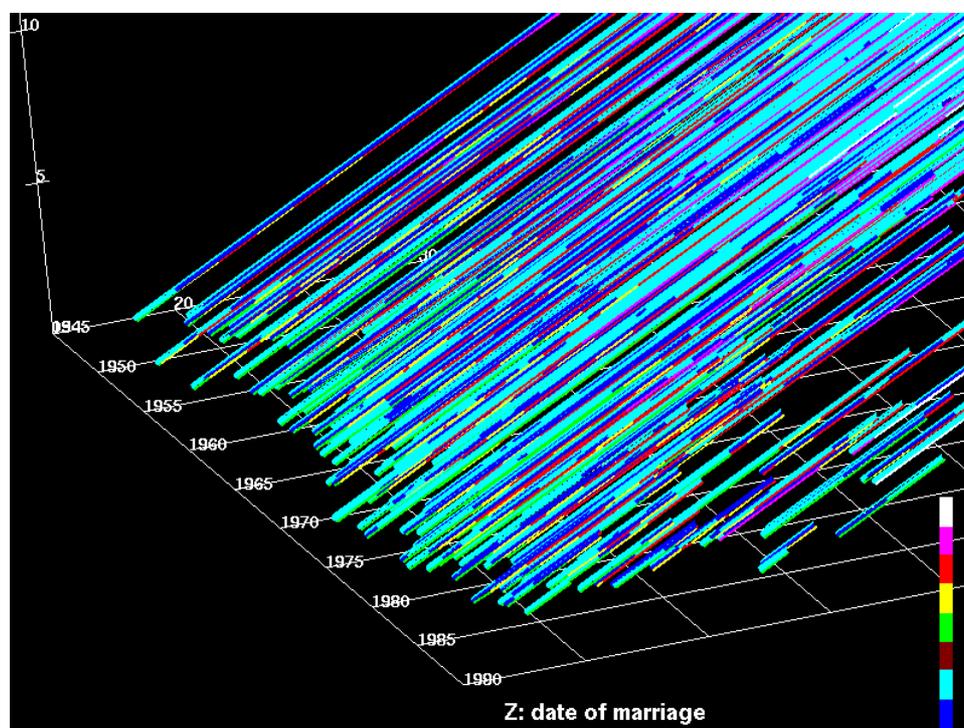
²⁹ Cf. tableau 2, dans C. Blayo (1991), p. 1384. Il est à noter que, dans son exemple, Blayo ne retient que les naissances féminines et les femmes analphabètes. Ces deux caractéristiques « qualitatives » sont hors de propos dans notre cas ; elles ont été abandonnées pour la suite de l'exposé.

appartiennent à deux générations ou que les mariages d'une génération à un âge révolu s'étalent sur deux générations, mais il s'agit là de considérations par ailleurs bien établies sur la base de simples diagrammes à deux axes...

2.5.4.4. « Lexis pencils »

Les « crayons de Lexis » ou « Lexis pencils »³⁰ se localisent le plus souvent³¹ dans des diagrammes de Lexis à trois axes (cf. figure 2.34). L'axe Z supporte la date du mariage ; l'axe vertical, la durée du mariage et l'axe en arrière plan l'âge. Les lignes de mariage s'y représentent non pas par de simples lignes, mais par des crayons dont chacune des trois faces visibles montre, par un système de couleurs³², le statut par rapport à l'emploi de la femme et de l'homme formant le couple concerné ainsi que l'âge de l'enfant le plus jeune au sein de la famille. Et c'est bien l'objectif de ce type de figuration : « a way of visualizing a moderate number of individual time changes in a number of categorical variables »³³.

Figure 2.33. Diagramme à trois axes et Lexis pencils



Source : B. Francis et J. Pritchard (1997, b)

Cette figure représente des lignes de mariage. En quoi diffère-t-elle de la figure 2.27 (indépendamment du fait que, d'un côté, les lignes sont nombreuses et, de l'autre, il n'y en a qu'une seule) ? Sur la figure 2.34, toutes les trajectoires commencent sur le plan horizontal et

³⁰ À ce sujet, cf. notamment B. Francis et J. Pritchard (1997, a) ou B. Francis et J. Pritchard (1997, b). Il est à noter que les crayons de Lexis s'utilisent éventuellement sur des diagrammes à deux axes. Par ailleurs, il existe un logiciel pour dessiner les crayons de Lexis, « Lexis plot » (adresse web : http://www.cas.lancs.ac.uk/alcd/visual/lexis_plot.html).

³¹ Parfois sur des diagrammes à deux axes (cf. http://www.cas.lancs.ac.uk/alcd/visual/lexis_pencil.html).

³² Initialement, sur la page web, les faces du crayon sont colorées et non en grisés, ce qui rend la figure plus attractive, si pas plus lisible.

³³ B. Francis et J. Pritchard (1997, a).

non pas à des hauteurs différentes. En fait, il suffirait de déposer le volume de la figure 2.27 sur le plan *temps-âge* et de l'orienter adéquatement pour obtenir le même effet.

Plus significatif comme différence : dans la figure 2.34, les trajectoires montent parallèlement au plan *âge-durée de mariage* alors que dans la figure 2.27, elles ne sont parallèles à aucun des plans définissant le volume initial. Il aurait suffi que, pour l'axe, Z on opte pour le temps et non pour la date de mariage pour que les deux figures ne s'opposent pas sur ce point. La coordonnée figurant sur l'axe Z de la figure 2.34 est invariante individu par individu, ce qui force les trajectoires à suivre une orientation parallèle au plan formé par les deux autres axes, alors que dans la figure 2.27, les trois coordonnées invariantes parasitent des axes et se traduisent donc par des obliques.

Bref, les crayons de Lexis sont bien un exemple d'application où un diagramme de Lexis à trois axes peut intervenir.

2.5.4.5. De l'utilité des diagrammes à plus de deux axes

En général, le démographe utilise le diagramme de Lexis pour visualiser les données et surtout comprendre comment elles s'organisent par rapport à la trame de la figure 2.14, trame que constituent les droites symbolisant les limites d'années et de générations ainsi que les âges exacts. Le plus souvent, les lignes de vie – ou autre trajectoires forcées, s'il échet – n'apparaissent pas sur le graphique ; en effet, pour peu que les observations se multiplient, ces lignes occupent tout l'espace et font disparaître la trame, ce qui rend la figure inutilisable et finalement sans grand intérêt pratique. Si l'une ou l'autre ligne de vie (en dehors de celles qui délimitent les générations) apparaît, c'est pour expliquer la logique de la construction et de l'utilisation de l'outil.

Dès 1880, Lexis épouse ce point de vue en défendant son diagramme³⁴ :

« De plus, cette construction nous permet très facilement de nous dispenser entièrement des lignes de vie, en ne conservant que leurs sommets, c'est-à-dire les *points mortuaires*... Il n'est d'ailleurs pas absolument nécessaire que ces points soient dessinés ou que leur densité plus ou moins grande soit représentée graphiquement ; la chose principal est toujours le réseau de lignes entre lesquelles on se figure mentalement les points mortuaires ».

Que le diagramme soit en plan ou en volume ne change rien à ce propos comme le suggère, à notre avis, la figure 2.34 : comment arriver à lire, comprendre ou organiser les observations sur ce diagramme où la trame disparaît derrière les crayons de Lexis et où des observations en dissimulent d'autres³⁵, et ce malgré l'apport de la couleur sur le diagramme original ?

³⁴ W. Lexis (1880), p. 301. Si nous sommes d'accord avec les considérations de Lexis pour utiliser concrètement son diagramme sans y dessiner les lignes de vie ou les points mortuaires, par contre on ne peut considérer qu'il s'agisse d'un point assurant la prétendue supériorité de sa construction par rapport à d'autres, comme celle de Becker (qui sera présentée dans le point 2.6), qui en fait présentent toutes les qualités de la sienne.

³⁵ À ce sujet, il faut signaler la possibilité parfois offerte de faire pivoter le diagramme de Lexis, comme indiqué sur le site web « <http://www.cas.lancs.ac.uk/alcd/visual/> » via l'option : « Animated gif file as an alternative for browsers which cannot view VRML files ». Cette possibilité est éventuellement intéressante à certains égards, mais pour autant que les observations restent peu nombreuses. Le facteur limitatif du nombre reste donc d'actualité.

Tant que le diagramme de Lexis reste à deux axes, son usage demeure aisé. Par contre, dès que la perspective dans la 3^e dimension devient nécessaire pour la représentation d'un troisième axe, il devient nettement moins convivial. Ainsi, les figures 2.29 ou 2.31 demanderont une attention des plus soutenues pour apprécier correctement la nature de l'information représentée. À notre sens, l'élaboration de diagramme à plus de deux axes est une voie sans grand intérêt : autant le diagramme en deux axes aide l'utilisateur, autant l'apparition d'un troisième axe en annule la plupart des bienfaits. Pour la suite de l'exposé, nous nous en tiendrons donc aux diagrammes à deux axes.

2.6. Historique du diagramme de Lexis

2.6.1. Introduction

Vu l'importance accordée au diagramme de Lexis dans ce travail, il nous a paru opportun d'aborder brièvement l'historique de sa mise au point au dix-neuvième siècle³⁶. Ce sujet se révélera d'autant plus intéressant que, d'une part, pour expliquer la construction de son diagramme, Lexis suivait un cheminement différent de celui que nous avons adopté et que, d'autre part, la question de la paternité de cette figure à été le sujet d'une controverse pendant une période relativement longue (plus de vingt ans ; *grosso modo* entre 1880 et le début du vingtième siècle, en prenant les dates extrêmes des documents de Lexis à ce sujet et dont nous avons pu prendre connaissance).

Le premier point abordera donc la façon dont Lexis expliquait sa version du diagramme ainsi que d'autres, notamment celle de Becker (point 2.6.2). Le point suivant traitera de la question de la paternité du diagramme et remettra en cause l'appellation du diagramme tant il semble évident que Lexis ne peut en revendiquer l'invention (point 2.6.3).

2.6.2. Le diagramme de Lexis expliqué par... Lexis

Pour expliquer sa construction³⁷, Lexis part d'une figure ne comportant qu'un seul axe représentant le temps (*cf.* figure 2.35.a). Sur cet axe, prennent place les points de naissance (soit respectivement pour trois individus, a, b et c) et de décès (soit respectivement A, B et C) ; l'espace entre les points de naissance et de décès pour un individu figure sa durée de vie. Si les observations se multiplient, le graphique devient rapidement inutilisable par profusion de points. Il fallait donc « employer une méthode quelconque pour séparer les lignes de vie superposées dans l'axe »³⁸.

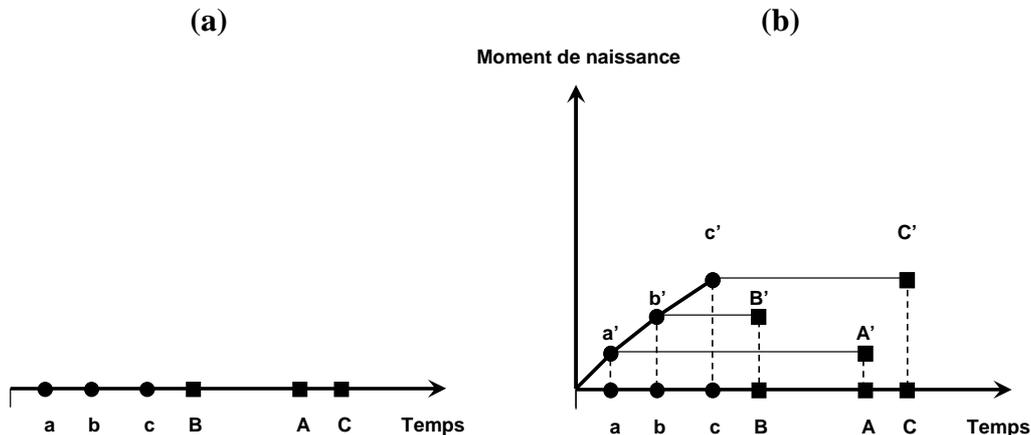
³⁶ Nous avons publié trois textes sur le sujet (C. Vandeschrick (1992, 1994 et 2001). Nous ne reprendrons ici que quelques éléments de ces publications, tout en ajoutant quelques compléments d'information. Par ailleurs, on trouvera un bon résumé en langue russe de l'histoire du diagramme dans la publication suivante : Grande Encyclopédie Soviétique (éditeur) (1994), notamment aux pages 440-441.

³⁷ Lexis a présenté pour la première fois sa figure dans une publication en allemand (W. Lexis (1875)). Par confort linguistique, nous nous référons toutefois préférentiellement à une publication en français parue 5 ans après (W. Lexis (1880)).

³⁸ W. Lexis (1880), p. 299.

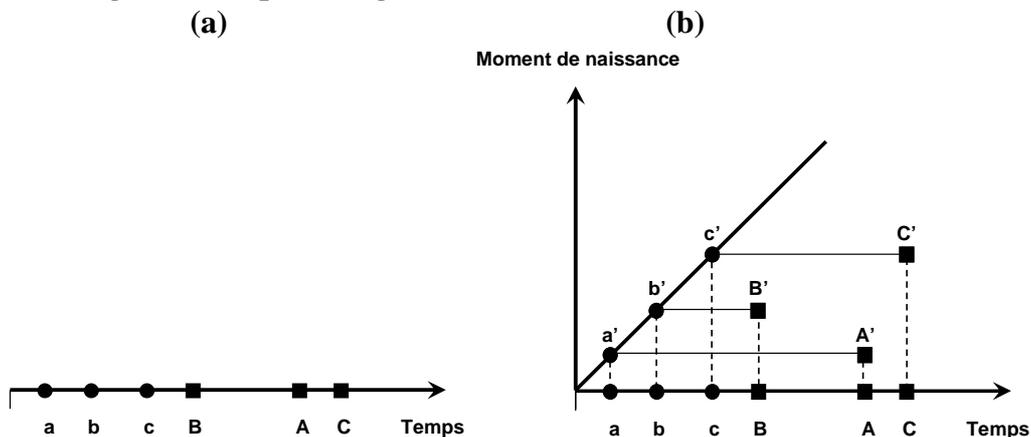
Une première solution consiste à déplacer verticalement les durées de vie en laissant entre chacune un espace constant, ce qui conduit à la figure de Knapp³⁹. Cette figure marque un progrès, mais souffre aussi d'un grave défaut : les âges et les générations ne peuvent s'y représenter de façon systématique ; l'importance graphique d'une génération est fonction du nombre d'individus qui la composent et un âge se représente par une ligne brisée et non par une simple oblique inclinée à 45° (cf. figure 2.35.b).

Figure 2.34. Ligne du temps et diagramme de Knapp



Becker reprit cette idée, mais en respectant, pour le déplacement vertical des lignes de vie, l'écart horizontal entre les points de naissance des différents individus (cf. figure 2.36.b)⁴⁰. Il obtient ainsi un diagramme en temps et moment de naissance tel que défini au point 2.3.4.2. Ce diagramme a toutes les qualités voulues, avec notamment la représentation systématique du temps, de l'âge et du moment de naissance.

Figure 2.35. Ligne du temps et diagramme de Becker

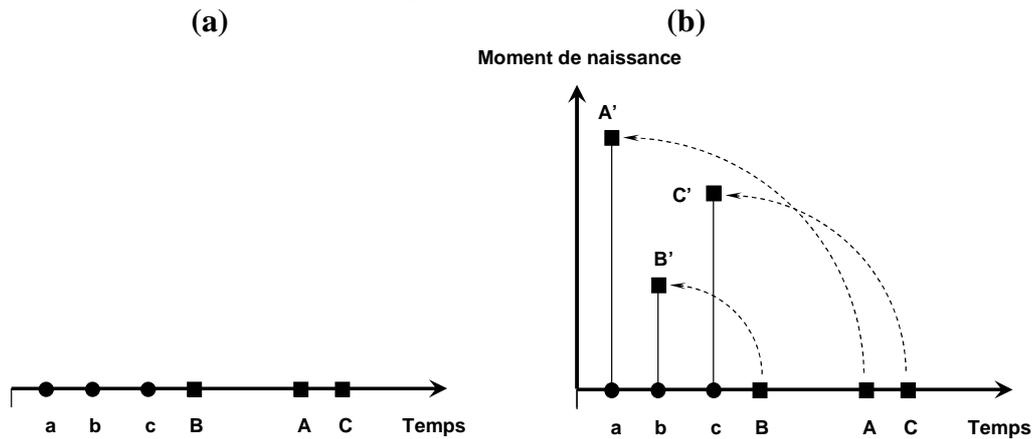


Lexis repart toujours de la figure à un seul axe, mais il fait subir une rotation de 90°⁴¹ aux durées de vie (cf. figure 2.37.b) ; celles-ci se retrouvent donc en position verticale. Son procédé débouche sur le diagramme en moment de naissance et âge (cf. point 2.3.4.1).

³⁹ G. Knapp (1868), p. 27.

⁴⁰ K. Becker (1874), p. 77.

⁴¹ Lexis a hésité à propos de l'angle de rotation (cf. W. Lexis (1880), p. 300). Il a aussi proposé un angle de 60° qui donnait un diagramme triangulaire ; selon Lexis ((1875, pp. 13-14), cette méthode offrait l'avantage de déboucher sur des triangles élémentaires équilatéraux et non isocèles. N. Ryder a aussi proposé une telle représentation triangulaire (à ce dernier sujet, cf. G. Caselli et E. Lombardo (1990), pp. 407-408).

Figure 2.36. Le diagramme de Lexis par Lexis

Lexis estime que sa méthode est la plus commode. Il ajoute :

« De plus, cette construction nous permet très facilement de nous dispenser entièrement des lignes de vie, en ne conservant que leurs sommets, c'est-à-dire les *points mortuaires*. En effet, lorsque les lignes de vie sont verticales par rapport à l'axe de temps, chacune d'elles est parfaitement déterminée dès que son point terminal a été établi, et dans ce cas la ligne elle-même devient superflue »⁴².

On ne peut suivre Lexis sur ce point : la construction de Becker a exactement les mêmes qualités et offre les mêmes possibilités que la sienne⁴³ ; la seule différence vient des variables privilégiées et de la variable parasite (l'âge chez Becker ou le temps chez Lexis). Pour le reste, tout ce que la version de Lexis permet de faire, celle de Becker le permet tout autant, moyennant quelques petites adaptations. Ainsi, sur le schéma de Becker, on peut se contenter de localiser les points mortuaires qui à eux seuls déterminent les durées de vie, même si celles-ci se retrouvent en position horizontale ; le caractère non indispensable du dessin effectif des lignes de vie sur le diagramme ne peut être considéré comme une exclusivité de la version de Lexis !

Par ailleurs, selon sa méthode d'explication, Lexis n'est pas en mesure de faire émerger la troisième forme du diagramme (le diagramme en temps et âge ; cf. point 2.3.4.3), qui a pourtant été proposée par Brasche plusieurs années auparavant⁴⁴. En effet, par simple déplacement ou rotation de la durée de vie depuis l'axe horizontal, il n'est pas possible d'aboutir à cette dernière construction, où la ligne figurant la vie d'un individu est plus longue que l'espace entre sa date de naissance et sa date de décès sur l'axe du temps, ce que Lexis préférerait sans doute éviter.

⁴² W. Lexis (1880), p. 301.

⁴³ Pour une comparaison des constructions de Lexis et Becker, cf. notamment G. Barsy (1958).

⁴⁴ O. Brasche (1870). Bien plus que Lexis, cet auteur doit être considéré comme un précurseur en matière de diagramme de Lexis (cf. point suivant).

2.6.3. Diagramme de Lexis : un déni de paternité

2.6.3.1. La controverse

À notre avis, le diagramme de Lexis ne porte pas le nom de son inventeur. Notre intérêt pour cette question de la paternité du diagramme trouve son origine dans une curieuse note d'une publication de Lexis, à propos de son livre de 1875 :

« Le titre porte selon l'usage de librairie la date de 1875, quoique l'ouvrage eût déjà paru en 1874, et l'impression avait déjà commencé pendant l'été en même temps que paraissait le travail de Becker, dont le mien est entièrement indépendant. Une thèse doctorale de Verwey, en langue hollandaise, traitant du même sujet et de l'application du même principe parut peu de temps après, mais n'avait pas de rapport avec mon travail. L'article de Verwey a été publié en anglais dans le '*Journal of the Statistical Society*', Londres, décembre 1875. On doit encore mentionner la construction stéréométrique de Zeuner (*Abhandlungen zur math. Statistik*, Leipzig. 1869) qui correspond au modèle, construit dernièrement par l'ordre de la Direction de la statistique italienne et qui a été décrit par L. Perozzo dans un intéressant article des '*Annali di Statistica*.' »⁴⁵

Dans cette note, Lexis semble sur la défensive. Ainsi, il insiste sur la date de publication de son livre officiellement paru en 1875, mais dont il soutient que l'impression était déjà terminée en 1874 et indépendamment d'un ouvrage de Becker paru en... 1874. À propos de Verwey⁴⁶, il évoque une thèse en néerlandais défendue en 1874, mais il cite un article en anglais paru en décembre 1875, soit un an après la thèse. Enfin, à propos de Zeuner, Lexis ne signale nullement l'éventuelle filiation de la figure de cet auteur avec la sienne propre, alors qu'elle s'impose et que l'ouvrage de Zeuner date de 1869.

En fait, une controverse à propos de l'invention du diagramme a opposé Lexis et Zeuner. Zeuner entame un texte daté de 1886 en stigmatisant le manque de reconnaissance émanant des spécialistes d'une discipline et dont souffrent certaines recherches menées par des mathématiciens. Il cite l'exemple des travaux de Gossen en économie politique, publiés en 1854 et ignorés jusqu'à ce que des scientifiques étrangers les découvrent vingt ans plus tard. Ensuite, il évoque de manière générale « la lenteur avec laquelle les mathématiciens pénètrent dans d'autres domaines scientifiques, en partie à cause des résistances qu'ils y rencontrent dès l'abord »⁴⁷ ; dans le paragraphe suivant, il signale qu'il en va de même en statistique, dont l'étude de la mortalité constitue un des sujets importants. À ce propos, il cite son ouvrage de 1869 – s'inscrivant dans la foulée de celui de Knapp édité en 1868 – et où il explicite les méthodes à employer en matière d'analyse de la mortalité en s'appuyant notamment sur des graphiques qui ne vont pas sans évoquer le diagramme de Lexis.

C'est sur la base de ce document de 1869 que, en 1886, Zeuner conteste ouvertement la paternité du diagramme que Lexis revendique. Voici deux citations pour le moins explicites :

- « Lexis utilise la même représentation graphique dans son livre '*Einleitung in die Theorie der Bevölkerungstastistik*' Strasbourg 1875, avec des axes à angle

⁴⁵ W. Lexis (1880), pp. 298-299.

⁴⁶ Le nom de cet auteur peut prendre une autre graphie : Verweij ; nous avons opté pour la graphie la plus souvent rencontrée, même si elle ne semble pas la plus authentique.

⁴⁷ G. Zeuner (1886), p. 6.

droit ; il accepte ainsi le plan de ma représentation stéréométrique – remarquons en passant qu’il n’en fait aucune mention. Je reviendrai ultérieurement sur cette question. »⁴⁸ ;

- « D’une façon générale, les figures 2 et 3 représentent exactement le plan de base de ma démonstration stéréométrique de l’année 1869 et c’est cette démonstration que Lexis a revendiquée pour lui en 1875. Toute la différence entre les deux démonstrations tient dans ce que Lexis, utilisant ma démonstration actuelle, veut décompter dans la figure 3 le nombre de points de décès qui tombent dans la courbe fermée abcd, alors que moi, je les déduis de la différence des points d’intersection avec les lignes de vie... »⁴⁹.

Lexis reviendra plus tard sur cette question de la paternité dans une publication de 1903⁵⁰. Dans une longue note de bas de page, il cite différents auteurs ayant aussi réfléchi à la représentation graphique des faits démographiques avec l’intention de démontrer qu’il doit bien être considéré comme l’inventeur du diagramme. Comme en 1880, il répète son argumentation à propos de Becker et Verwey. De Brasche (ignoré en 1880, mais pas en 1875), il cite une construction qu’il qualifie d’ébauche dont la publication remonte à 1870. Pour l’essentiel, la note porte sur son différend avec Zeuner ; Lexis se défend vigoureusement contre les attaques de Zeuner :

« C’est de manière tout à fait injustifiée que Zeuner (dont j’ai par ailleurs présenté les écrits dans mon livre de 1875 en page 33, et aussi dans le texte en français de cet essai) prétend dans un travail ultérieur (...) que ma construction a été reprise de la base de sa présentation stéréométrique. Je suis parti d’une toute autre conception, et, si en faisant cela, c’est la même forme de réseau que dans le plan de base de Zeuner qui en résulte, le contenu de celle-ci est cependant pensé différemment »⁵¹.

La différence de conception dont question repose sur le fait qu’il représente des densités de décédés contre des densités de vivants pour Zeuner. Ce dernier et Lexis semblent donc bien inconciliables !

Pour trancher la question de la paternité du diagramme, nous allons établir un historique aussi précis que possible des différentes étapes en ayant jalonné la mise au point. Pour entrer dans la course à la paternité, un auteur se doit d’avoir proposé un diagramme comportant deux axes, mais permettant le repérage systématique des trois variables âge, temps et moment de naissance et, en sus, avoir intégré – éventuellement sans les dessiner – la notion de trajectoires forcées à travers les lignes de vie. Avoir complété la figure en y ajoutant les réseaux de droites représente un plus appréciable car facilitant de beaucoup l’utilisation effective de l’outil. Par ailleurs, il est utile de rappeler ici que le diagramme de Lexis se décline en fait en trois versions équivalentes, se distinguant simplement selon le choix des variables privilégiées ou parasite (*cf.* point 2.3.4).

⁴⁸ G. Zeuner (1886), p. 8.

⁴⁹ G. Zeuner (1886), p. 11.

⁵⁰ W. Lexis (1903).

⁵¹ W. Lexis (1903), pp. 3-4.

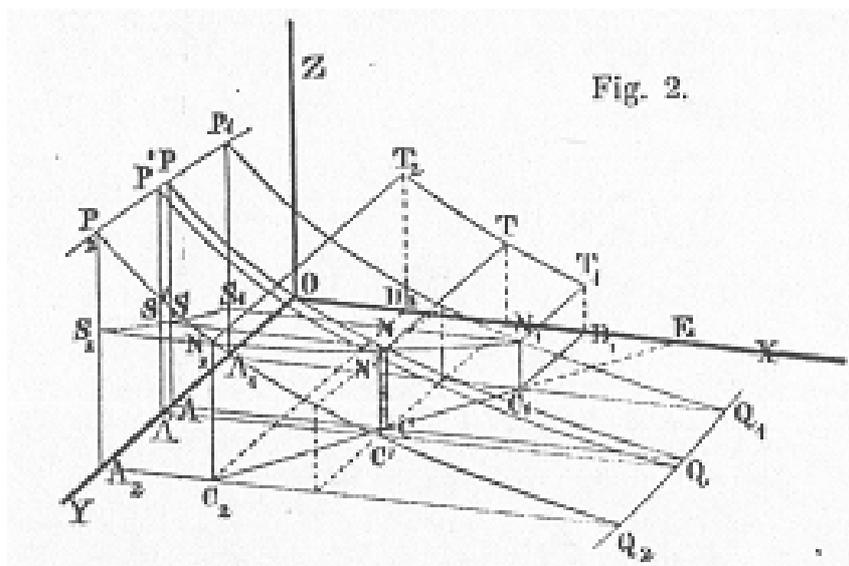
2.6.3.2. Les nominés par ordre chronologique

a. Zeuner

C'est en 1869 que Zeuner édite son diagramme (*cf.* figure 2.38). Cette figure mérite le qualificatif de « stéréométrique » dans le sens où elle représente non seulement les trois coordonnées, mais aussi en plus les effectifs qui prennent place sur l'axe vertical Z. Le plan horizontal formant la base du volume se délimite par les axes du moment de naissance (OY) et de l'âge (OX) ; il s'agit indubitablement d'un diagramme en moment de naissance et âge, à l'exacte image de ce que Lexis a proposé. Zeuner ne fait pas apparaître les lignes de vie dont il n'a pas besoin vu le caractère stéréométrique de son diagramme. Il parle par contre d'obliques décroissantes permettant de localiser un effectif au moment précis d'un recensement, par exemple.

Si Zeuner a bien utilisé comme base de sa figure une des versions du diagramme, par contre, il n'a pas ajouté de façon systématique les réseaux de droites pour en faciliter l'utilisation, ce qui s'explique une fois de plus par sa volonté de proposer une représentation stéréométrique.

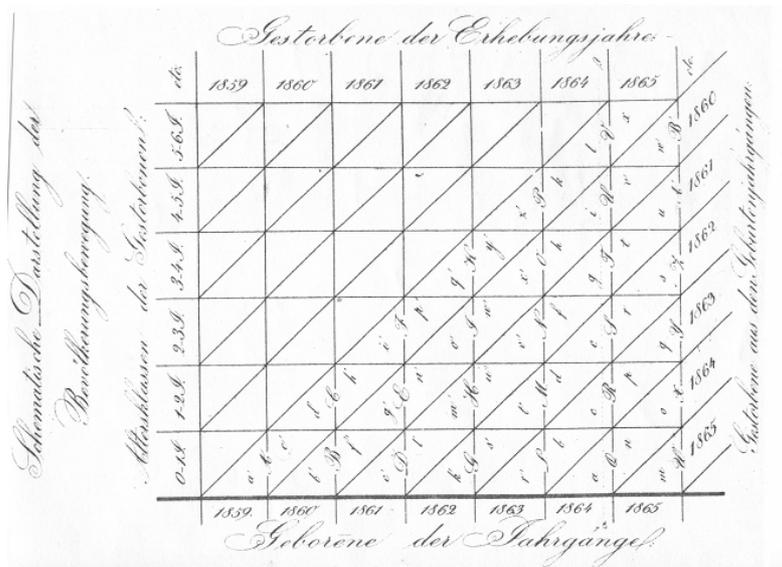
Figure 2.37. Le diagramme de Zeuner



Source : G. Zeuner (1869), p. 13.

b. Brasche

Brasche publie en 1870, soit un an après Zeuner, un essai sur le calcul de la table de mortalité. Pour sa démonstration, il se sert d'un diagramme en temps et âge (*cf.* figure 39). Cette figure comporte bien les réseaux de droites et suppose la présence sous-jacente des lignes de vie pour son emploi. Il l'utilise pour répartir les décès en fonction de l'année de survenance, de l'âge des décédés ainsi que de l'année de naissance. Il désigne les triangles élémentaires par des lettres. C'est à l'aide des effectifs de décédés ainsi localisés qu'il reconstitue le mouvement de la population ou qu'il établit ses formules pour le calcul des probabilités de mourir.

Figure 2.38. Le diagramme de Brasche

Source : O. Brasche (1870), hors texte.

Dès 1870, la version la plus utilisée actuellement du diagramme de Lexis (après sa remise au goût du jour par Pressat dans les années '50) est donc au point et son concepteur l'utilise de façon tout à fait moderne. On ne peut que regretter son laconisme à propos de l'élaboration du diagramme :

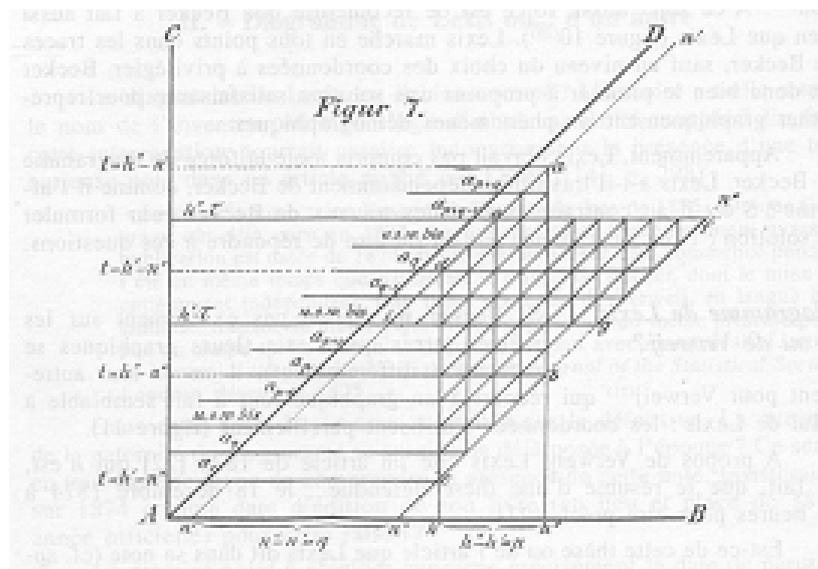
« Pour une démonstration plus précise de cette proposition de même que des explications qui vont suivre, on a esquissé sur une table lithographique la présentation schématisée du mouvement de la population. Si on se représente le schéma étendu de manière quelconque dans la direction du temps (vers le haut et vers le bas) et le nombre des colonnes des classes d'âge des personnes décédées continuées jusqu'à la plus haute classe d'âge possible, si d'autre part, on tient de plus compte du fait que la grandeur relative des différents champs est sans importance, le contenu de ce qui suit devrait rendre toute explication supplémentaire du schéma inutile »⁵².

Il est bien dommage que Brasche n'ait pas été plus explicite à propos de sa réflexion graphique. Par ailleurs, il ne signale aucune filiation pour sa construction. Il laisse le lecteur complètement sur sa faim à ce sujet, alors qu'il est le premier, à notre connaissance, à proposer un diagramme en plan ayant toutes les qualités voulues ! Ce défaut d'explication a sans doute pesé lourd dans le manque de reconnaissance dont sa réflexion en matière de représentation graphique a souffert, notamment de la part de Lexis. Il est aussi par ailleurs possible que l'essai de Brasche ait été desservi par certaines méthodes qu'il a utilisées, notamment pour corriger les problèmes au niveau des données.

c. Becker

En 1874, dans un ouvrage une fois de plus consacré à la mortalité, Becker propose aussi une figure pour représenter les données. Tant les lignes de vie que les réseaux de droites se retrouvent dans les figures de Becker. Au contraire de Brasche, Becker agrmente le graphique de commentaires assez complets permettant d'en comprendre le fonctionnement, et notamment la logique suivie pour établir les lignes de vie.

⁵² O. Brasche (1870), p. 20 (note de bas de page numéro 3, se référant à la page 4 du texte).

Figure 2.39. Le diagramme de Becker

Source : K. Becker (1874), hors texte.

Becker précisera entre autres que :

« On peut appeler "ensembles"⁵³ de vivants, de décédés ou de migrants des groupes limités ainsi, désignés par certaines caractéristiques concernant la date de naissance, la date de l'observation ou l'âge et issus de la masse observée. On peut ainsi observer toutes sortes d'ensembles, et aussi déduire de la manière la plus simple tous les rapports entre eux. Dans la mesure où seuls les rapports entre vivants, décédés ou migrants sont en question, on peut faciliter l'observation en négligeant toutes les lignes de vie et en considérant strictement les lignes ou surfaces délimitées comme les ensembles en question »⁵⁴.

Comme le montre cet extrait, Becker a reconnu l'importance des lignes de vie pour comprendre le fonctionnement de sa figure, mais aussi leur aspect superfétatoire quand il s'agit d'explicitier la logique d'un calcul : seule compte la forme contenant les événements concernés.

d. Verwey

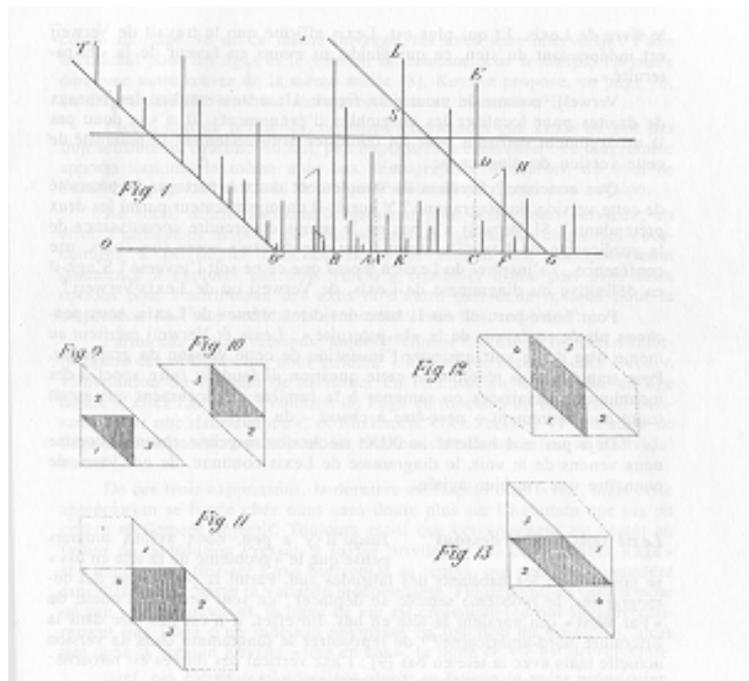
En décembre 1874, Verwey défend une thèse où il abordera également la question de la table de mortalité. Lui aussi, il introduit une figure pour expliciter les méthodes qu'il décrit ; il s'agit d'un diagramme en plan avec les axes dévolus au moment de naissance et à l'âge (cf. figure 2.41). Il utilise les lignes de vie ainsi que les réseaux de droites pour délimiter les ensembles de vivants ou de décédés intervenant dans ses calculs.

Verwey reconnaît avoir pris connaissance du livre de Becker, mais à un moment où sa thèse était déjà achevée, selon ses dires. S'il cite Knapp, Zeuner et Becker, il ne mentionne pas Lexis, même en décembre 1875 quand il publie un article en anglais sur le sujet de sa thèse.

⁵³ D'après la terminologie de Knapp, qui est acceptée ici.

⁵⁴ K. Becker (1874), p. 15.

Figure 2.40. Le diagramme de Verwey

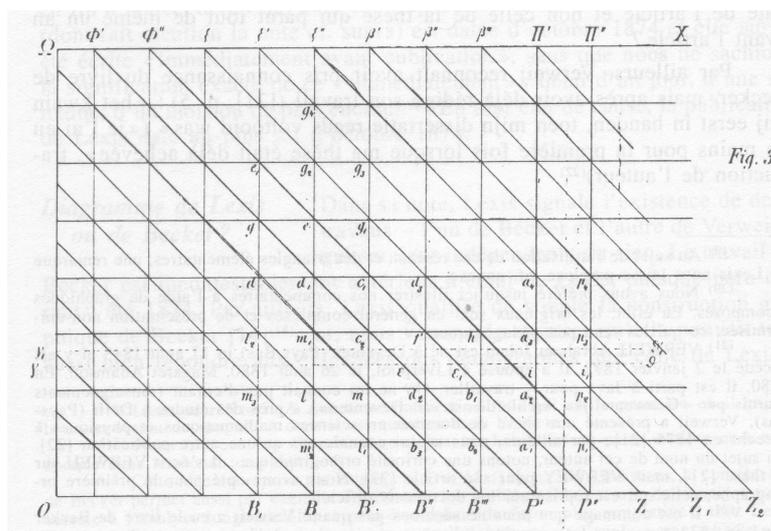


Source : A. Verweij(1874), hors texte.

e. Lexis

C'est en 1875, selon la date d'édition figurant sur le livre, que Lexis décrit pour la première fois le graphique qui portera son nom⁵⁵. Diagramme en moment de naissance et âge, c'est l'exacte réplique de celui de Verwey avec notamment les réseaux de droites pour faciliter la localisation d'observations (cf. figure 2.42). C'est à propos de ce livre que Lexis débat de la véritable date de publication dans son article de 1880 en insistant sur le fait que l'ouvrage était déjà paru en 1874 (cf. supra).

Figure 2.41. Le diagramme de Lexis par Lexis...



Source : W. Lexis (1875), hors texte.

⁵⁵ W. Lexis (1875).

2.6.3.3. Rendons à Zeuner et à Brasche ce qui n'appartient pas à... Lexis

Un diagramme de Lexis ne compte que deux axes, mais permet de recourir à trois coordonnées⁵⁶. À notre connaissance, Zeuner fut le premier en 1869 à avoir proposé une solution satisfaisante à ce sujet avec son diagramme en moment de naissance et âge. Brasche le suit de près dans cette voie en publiant en 1870 un diagramme avec des axes dévolus cette fois au temps et à l'âge. Becker, enfin, propose en 1874, la troisième version du diagramme, en privilégiant le temps et le moment de naissance.

En 1874/75, Verwey et Lexis reviennent sur la solution de Zeuner, mais en supprimant l'axe des effectifs ; ils sont donc les premiers à proposer un graphique purement en plan avec le moment de naissance et l'âge comme variables privilégiées. Il est bien difficile de départager ces deux auteurs en termes d'antériorité. La thèse de Verwey porte le millésime de 1874 et le livre de Lexis, celui 1875. Lexis a insisté sur le fait que son livre avait déjà paru en 1874 ; vu la date de la préface, on peut préciser que la parution ne peut remonter avant octobre 1874⁵⁷.

La thèse de Verwey parut, selon les dires de Lexis, peu de temps après son ouvrage. À ce sujet, on peut se demander pourquoi Lexis cite à chaque fois les références de l'article en anglais paru en décembre 1875 quand il évoque cette thèse en néerlandais défendue en décembre 1874 ! Quoi qu'il en soit, Lexis signale lui-même que la thèse n'a pas de rapport avec son travail, ce qui en toute logique devrait l'empêcher de tout revendiquer pour lui seul.

Que conclure à propos de la version moment de naissance et âge ? Malgré l'avantage de Verwey en matière d'antériorité officielle, on peut penser que ces deux auteurs ont à partager la paternité de l'une des trois versions du diagramme dans sa configuration plane. Revenons-en maintenant à l'invention du diagramme en tant que tel, et plus seulement à celle de la version en moment de naissance et âge.

En prenant le critère des deux axes pour trois coordonnées, au pire Verwey et Lexis disparaissent du palmarès, si Zeuner est maintenu dans ses droits malgré le troisième axe et au mieux, si Zeuner est éliminé, se retrouvent à la troisième place derrière Brasche et Becker. Si le critère des réseaux de droites se surimpose au premier, Zeuner disparaît ; Brasche se classe en tête, suivi de Becker et finalement du tandem Verwey et Lexis.

Il est assez remarquable de constater que la version utilisée le plus souvent actuellement, du moins dans le monde francophone, a été mise au point la première et est restée quasi ignorée pendant plus de 75 ans avant d'être « ressuscitée » par Pressat dans les années '50.

Il est piquant de constater que, en définitive, celui qui a laissé son nom au diagramme ne peut nullement prétendre à sa paternité ; s'il a inventé quelque chose, c'est simplement la version qui est apparue la dernière. De plus, il doit partager cette invention avec un concurrent, Verwey. Ce n'est pas tout : la version de Lexis s'inscrit dans le droit chemin du travail de Zeuner, paru en 1869, ce que Lexis a toujours refusé d'admettre, malgré l'argumentation avancée par Zeuner. D'une manière générale, on peut avoir l'impression que Lexis n'a pas toujours bien saisi tous les tenants et aboutissants du graphique :

- de la version de Brasche, il prétend qu'il s'agit d'une ébauche⁵⁸ alors qu'elle constitue une version complète et, qui plus est, utilisée très adroitement en respectant des principes d'utilisation qu'il énonce notamment en 1880 : « Il n'est d'ailleurs pas absolument

⁵⁶ Pour cette partie, nous nous inspirons largement d'une publication antérieure : C. Vandeschrick (2001)

⁵⁷ À la fin de la préface, se trouve la mention « Dorpat, im October 1874 » (cf. W. Lexis (1875), p. iv).

⁵⁸ Cf. W. Lexis (1903), p. 3.

nécessaire que ces points soient dessinés ou que leur densité plus ou moins grande soit représentée graphiquement ; la chose principale est toujours le réseau de lignes entre lesquelles on se figure mentalement les points mortuaires »⁵⁹ ; c'est exactement la façon dont Brasche utilise son diagramme, et ce, cinq bonnes années avant Lexis !

- de la version de Becker, il ne comprend pas qu'elle présente les mêmes qualités que la sienne (cf. point 2.6.2) ;
- au départ de sa propre version, il lui est arrivé de produire des commentaires étonnants⁶⁰.

S'il fallait rebaptiser le diagramme, deux choix s'offrent à nous :

- soit le diagramme de Zeuner, vu sa primauté en matière de représentation de trois coordonnées sur un système à deux axes ;
- soit le diagramme de Brasche, vu sa primauté en ce qui concerne l'ajout de réseaux de droites et sa façon d'utiliser le diagramme.

Pour notre part, nous inclinierions pour la deuxième possibilité, notamment en fonction d'un argument présenté par Lexis pour se défendre des attaques de Zeuner :

« Ma construction donne donc de suite une vision d'ensemble des décès dont les ensembles sont tous délimités par des droites sur la surface de base. Chez Zeuner, par contre, les ensembles de personnes décédées n'apparaissent pas dans la surface de base, mais dans une surface de coordonnées verticales sur le plan de base et sur le côté, ce qui a pour conséquence que les ensembles élémentaires importants apparaissent avec des délimitations en partie tordues, dans un dessin à la perspective inconfortable »⁶¹.

Un diagramme, pour être utile, doit nous aider à localiser aisément les données ; il faut reconnaître que la construction de Zeuner ne peut se prévaloir de simplicité vu le recours à la troisième dimension. De ce point de vue de la simplicité, Brasche devance Lexis de 4 ou 5 ans, ce que ce dernier, pour rappel, n'a jamais véritablement compris ou admis !

Malgré cela, c'est bien l'appellation « diagramme de Lexis » qui s'est imposée et de manière solide, même si, selon les sources, l'accord sur le rôle prépondérant de Lexis n'est pas total. Ainsi, par exemple, Seligman écrit à propos de Lexis⁶² : « Special mention must be made of... a new method of graphic treatment of mortality, first broached in his *Einleitung in die Theorie des Bevölkerungsstatistik* (Strasbourg 1875) and developed in later publications ». De même Petersen et Petersen considèrent Lexis comme l'initiateur⁶³ : « Lexis diagram, first developed by the German statistician Wilhem Lexis, has become a widely used tool in demography ». J. et M. Dupâquier accordent aussi une primauté à Lexis. Après avoir parlé de la construction de Becker, ils écrivent⁶⁴ : « A la fin de la même année (bien que le titre porte la date de 1875), le mathématicien allemand W. Lexis propose une solution beaucoup plus complète, qui va permettre de résoudre tous les problèmes de passage de l'analyse transversale à l'analyse longitudinale et réciproquement ».

⁵⁹ Cf. W. Lexis (1880), p. 301.

⁶⁰ Cf. C. Vandeschrick (1992), pp. 1256-1260.

⁶¹ Cf. W. Lexis (1903), p. 4.

⁶² E. Seligman (Editor in chief) (1963), pp. 426-427.

⁶³ W. et R. Petersen (1986), p. 523. Le suite du commentaire décrit la version *temps – âge*, sans signaler qu'il ne s'agit pas de celle utilisée par Lexis.

⁶⁴ J. et M. Dupâquier (1985), p. 386. Pour rappel, notons que la construction de Becker est aussi performante que celle de Lexis ; elle permet tout autant de résoudre le passage de l'analyse transversale à l'analyse longitudinale et de façon peut-être encore plus suggestive : à l'horizontalité du longitudinal s'oppose la verticalité du transversal !

Au contraire, Westergaard voit en Lexis un continuateur de Zeuner et Knapp ; après avoir expliqué les apports de ces derniers, il écrit⁶⁵ : « As a worthy supplement a contribution by W. Lexis can be mentioned. He too uses geometric methods ». Westergaard mentionne également Verwey, mais en citant son article en anglais et non sa thèse en néerlandais. Caselli et Lombardo signalent que Becker et Lexis ont résolu le problème de la représentation sur un plan des 3 coordonnées démographiques⁶⁶. Vu l'antériorité de sa publication, l'avantage irait à Becker, même si ces auteurs n'abordent pas cette question. Cette idée est reprise dans un dictionnaire encyclopédique russe où les apports de Zeuner, Knapp et Becker sont mis en évidence⁶⁷ ; par ailleurs, la version utilisée actuellement y est attribuée à Pressat.

L'apport de certains des auteurs cités dans l'historique de l'élaboration du diagramme est plus ou moins régulièrement reconnu (Knapp, Zeuner, Becker et bien évidemment Lexis) ; par contre d'autres restent largement inconnus (Verwey ou surtout Brasche). Par ailleurs, le rôle de Lexis dans cette saga diffère nettement selon les sources : seul initiateur selon les uns, continuateur selon les autres. Mais dès lors, pourquoi donc l'appellation « diagramme de Lexis » s'est-elle imposée ? Indépendamment du fait que sa version a été régulièrement utilisée jusqu'à ce que Pressat repropose celle de Brasche et du fait que Pressat n'a nullement cherché à changer d'appellation, plusieurs hypothèses peuvent être, à notre avis, avancées :

- Lexis a longtemps travaillé sur ce diagramme ; sur une période de 30 ans, il a publié différents écrits où le diagramme occupe une place plus ou moins importante. Par ailleurs, Lexis ne s'est pas contenté d'établir un graphique ; il en a aussi proposé différents modes d'utilisation, comme par exemple la figuration des nombres d'évènements différents par des cercles de différentes couleurs⁶⁸. A ce sujet, la question est de savoir si des développements de ce genre étaient utiles. A notre avis, non. L'utilisateur moderne s'en est d'ailleurs éloigné !
- Lexis semble avoir été un scientifique considérable en son temps, sans doute plus considérable que certains des autres chercheurs cités dans l'historique. Par ailleurs, son activité a largement dépassé le cadre de la dynamique des populations ;
- Lexis a aussi publié en français, langue importante en matière de statistique au dix-neuvième siècle, ce qui lui a donné une envergure internationale reconnue.

Le diagramme est mal nommé, mais son appellation perdurera sans doute. Dommage pour Brasche et Zeuner et tant mieux pour Lexis !

2.7. Conclusions du chapitre 2

L'apport principal du deuxième chapitre aura été l'identification des conditions mathématiques à respecter pour procéder à la cohortisation d'un phénomène. Plutôt que d'aborder directement cette question par sa face purement mathématique (point 2.4), nous avons préféré opérer d'abord un détour en empruntant le chemin graphique (point 2.3). En effet, à l'instar de ce qui se passe pour l'exposé des méthodes d'analyse démographique, le diagramme de Lexis – sans être indispensable à strictement parler – rend beaucoup plus palpables les éléments en jeu. Comme souligné au début du chapitre (point 2.2), le diagramme de Lexis revêt deux caractéristiques principales :

⁶⁵ H. Westergaard (1932), pp. 222-223.

⁶⁶ G. Caselli et E. Lombardo (1990), pp. 402-403.

⁶⁷ Grande Encyclopédie Soviétique (éditeur) (1994), pp. 21, 93, 439-441.

⁶⁸ À ce sujet, cf. I. Esenwein-Rothe (1992).

- bien que ne comportant que deux axes, il permet d'utiliser trois variables pour y localiser des observations ;
- l'utilisation du diagramme de Lexis suppose la présence de trajectoires forcées, dans le sens où une unité d'observation ne peut quitter celle qui lui est dévolue.

Pour que ces deux caractéristiques émergent d'un diagramme, deux conditions doivent être remplies :

- la relation dans une combinaison linéaire des trois variables en jeu, condition nécessaire pour se contenter de deux axes ;
- pour chaque individu, l'invariance de l'une des trois variables, ce qui entraîne l'apparition des trajectoires forcées.

Les trois variables *âge*, *temps* et *moment de naissance* satisfont bien les deux conditions :

- elles se combinent dans la relation suivante :
$$\text{âge} = \text{temps} - \text{moment de naissance} ;$$
- pour un individu, le moment de naissance est invariant.

Ces trois variables peuvent donc servir à confectionner un diagramme de Lexis ; par extension, si des observations se classent en les utilisant comme système de référence, le phénomène en cause pourra s'étudier en suivant les méthodes de la cohortisation.

L'explication – tant graphique que mathématique – du principe de la réduction du volume en plan nous a aussi permis de mettre en évidence l'équivalence de trois solutions envisageables ; celles-ci se différencient selon le plan du volume initial où les lignes de vie se projettent. Par ailleurs, dans un premier effort de généralisation, nous avons montré que non seulement l'âge mais aussi le temps pouvaient très bien céder leur place respective à d'autres durées ; en conséquence, nous avons proposé une définition généralisée pour la notion de cohorte en cas de variables temporelles.

Sur un plan plutôt technique, il nous a semblé opportun d'envisager à la fois des diagrammes de Lexis à plus de deux axes et les « Lexis pencils ». À nos yeux, ces deux voies d'utilisation ne présentent que peu d'intérêt ; en effet, elles supposent l'abandon de l'une des qualités remarquables du diagramme de Lexis : sa capacité à représenter visuellement de façon simple les données. L'intrusion d'une troisième dimension est en rupture avec ce principe de simplicité.

Finalement, dans un point historique, l'appellation du diagramme de Lexis a été remise en cause. En effet, sur base d'une reconstitution aussi complète que possible des étapes de la mise au point de cet outil, il appert clairement que Lexis ne peut en aucun cas être considéré comme l'inventeur du diagramme qui porte son nom ; bien plus, on a tout lieu de penser qu'il a été plutôt dans les derniers « à sauter sur la balle ». Brasche, qui a proposé quatre à cinq ans avant lui une figure ayant toutes les qualités voulues, mériterait bien plus que Lexis d'avoir laissé son nom au diagramme.

Dans le chapitre suivant, nous allons étendre la généralisation de l'utilisation du diagramme de Lexis en envisageant notamment le cas de variables non temporelles. Au-delà de la simple présentation de données sur un diagramme de Lexis, nous essayerons ensuite, dans le chapitre 4, de montrer que l'analyse en cohorte pourrait aussi s'appliquer dans ces circonstances et déboucher sur des considérations intéressantes en matière d'analyse de l'observé.