

Comment faire du PDF avec L^AT_EX

David MONNIAUX

8 juin 2001

1 Pourquoi du PDF ?

Certains ne voient pas très bien l'intérêt de produire des fichiers au format PDF, sachant qu'il est très facile de produire des fichiers DVI ou PS à l'aide de T_EX et dvips ; ces fichiers sont ensuite lisibles par un certain nombre de visualisateurs, dont certains sont gratuits, voire libres. Cependant, cette approche souffre de certains inconvénients.

1. Il est parfois difficile de visualiser des fichiers DVI ou PostScript.
 - Les visualisateurs DVI sont intégrés dans les systèmes T_EX, notamment parce qu'ils ont besoin des polices de caractères standard de ces systèmes. Peu de gens hors de certains milieux scientifiques possèdent ce type d'installation, et peu de gens voudront prendre la peine d'installer une distribution T_EX juste pour visualiser quelques documents.
 - Les visualisateurs PostScript gratuits basés sur Ghostscript, sont assez lents et d'une installation peu facile pour un néophyte.

En comparaison, le logiciel Acrobat Reader, qui permet de naviguer dans les fichiers PDF, est disponible gratuitement auprès d'Adobe (<http://www.adobe.com>), s'installe facilement sur les systèmes les plus répandus (Windows et Mac) et d'ailleurs est souvent déjà installé par défaut sur ces systèmes (beaucoup d'éditeurs de logiciels et de fabricants de matériel le font installer automatiquement pour l'affichage des manuels en ligne). Acrobat Reader est également disponible pour les systèmes Sun Solaris et PC Linux. Par ailleurs, les visualisateurs PostScript gratuits basés sur Ghostscript (GSview, gv...) permettent d'afficher le PDF.

2. Les fichiers DVI ne peuvent pas contenir directement de figures PostScript, et l'utilisation de polices de caractères non standard est rendue difficile par la nécessité pour le lecteur d'installer ces polices. Les formats PostScript et PDF permettent l'inclusion des polices non standard et des figures.
3. Le format PDF a des fonctions hypertextes :
 - table des matières escamotable placée à côté du texte et permettant d'aller sur un chapitre ou une section d'un simple clic ;
 - liens et renvois hypertextes.

2 Les outils

Il y a essentiellement deux grands types de méthodes : celles qui passent par l'intermédiaire d'un fichier PostScript, reprenant en cela la méthode originale des produits d'Adobe, et celles qui s'en passent.

2.1 Passage par Postscript

La première catégorie de méthodes produit du PostScript en 3 étapes : compilation vers DVI, production de PostScript avec `dvips`, obtention du PDF à partir du PostScript.

2.1.1 Acrobat Distiller

La société Adobe, créatrice du format PDF, propose un logiciel appelé Acrobat Distiller, produisant du PDF à partir de PostScript. Malheureusement, ce logiciel n'est disponible que pour Microsoft Windows et Macintosh et est de plus quelque peu coûteux. Je ne l'ai jamais utilisé.

2.1.2 `ps2pdf`

`ps2pdf` est un logiciel inclus dans Ghostscript (<http://www.ghostscript.com>). En fait, c'est un simple script, qui se contente d'appeler Ghostscript avec les bonnes options. L'utilisation en est simple : on produit un fichier PostScript avec `dvips`, puis on fait `ps2pdf fichier.ps fichier.pdf`.

Les avis sur les PDF produits avec `ps2pdf` sont très divergents. Certains se plaignent que le PDF obtenu est « illisible » ; d'autres prétendent que l'on est alors forcé de n'utiliser que la police *Times*. Comme souvent dans ce genre de situations, le problème est qu'on parle de différentes versions du même logiciels ainsi que de différentes façons de l'utiliser.

Voilà ce qu'il faut retenir : *n'utiliser `ps2pdf` qu'avec une version de Ghostscript supérieure ou égale à 6.0*. Maintenant, pour ceux qui sont intéressés, voici l'explication. Les versions antérieures de Ghostscript ne savent pas inclure des polices de caractères Type 1 dans les fichiers PDF ; lorsque le document utilise des polices hors les 14 polices Acrobat de base (Times, Helvetica, Courier, Symbol et variantes), celles-ci sont incluses dans le fichier PDF sous forme de petites images *bitmap* — or Acrobat Reader est très mauvais pour l'affichage de polices de caractères de ce type. L'impression ne pose par contre aucun problème.

2.2 `pdfLATEX`

`pdfTEX` est une version modifiée de T_EX qui produit du PDF au lieu de DVI.

2.3 `dvipdfm`

`dvipdfm` produit du PDF à partir de DVI.

Package	De base :		
	Sérief	Sans sérief	Monoespacée
défaut	Computer Modern Roman	C.M. Sans Serif	C.M. Typewriter
times	Times	Helvetica	Courier
palatino	Palatino	Helvetica	Courier
newcent	New Century Schoolbook	AvantGarde	Courier
bookman	Bookman	AvantGarde	Courier

FIG. 1 – Les polices de caractères

3 Polices de caractères

Tout d’abord, commençons par abattre un mythe très répandu à propos de \LaTeX , à savoir que celui-ci ne permet pas de changer de police de caractères. Cela est faux, voir d’ailleurs la figure 1. Ce qui est vrai, par contre, c’est que la sélection des polices est quelque peu complexe ; toutefois, des *packages* spécialisés permettent de changer facilement les polices par défaut du document (Fig.1, colonne de gauche). Nous verrons comment tirer profit de ces *packages* afin de faciliter la production de PDF.

Maintenant, passons à une petite explication sur la nature des polices de caractères. Traditionnellement, les polices de caractères utilisées avec \TeX étaient décrites par des programmes pour le logiciel METAFONT. Celui-ci produit des images des caractères à partir d’une part du programme décrivant la police, d’autre part de paramètres dépendant de l’imprimante. Les images de caractères produites sont dites *bitmap*, c’est-à-dire que chaque caractère est représenté par une mosaïque de points. Lorsque `dvips` produit un fichier PostScript, par défaut, il utilise METAFONT pour produire des polices *bitmap* adaptées au modèle d’imprimante réglé (typiquement une imprimante laser à 600 points par pouce), qu’il inclut dans le fichier PostScript (format dit Type 3).

Voyons maintenant le cas du PDF. Tout d’abord, tout visualisateur PDF est censé savoir afficher les 14 polices de base (Times, Helvetica, Courier, Symbol et variantes). Par contre, toutes les autres polices utilisées doivent être fournies dans le fichier PDF. Le format PDF accepte des polices Type 1, Type 3 ou TrueType ; toutefois, il est recommandé de ne pas utiliser de polices Type 3, dont le rendu à l’écran par Acrobat Reader est souvent désastreux (caractères flous et illisibles).

Nous voyons donc un problème évident : par défaut, `dvips` produit des polices dont l’affichage par Acrobat Reader est de très mauvaise qualité ! La solution : *il faut n’utiliser que des polices PostScript Type 1 en plus des polices de base du PDF.*

Nous ne sommes pas au bout de nos peines. \LaTeX , par défaut, utilise une famille de polices de caractères appelées *Computer Modern*, dessinée par Donald E. KNUTH (l’auteur de \TeX et de METAFONT). Lorsqu’on utilise `\usepackage[T1]{fontenc}`, par défaut, on utilise des polices *European Computer Modern*, une extension des précédentes mieux adaptée aux langues européennes écrites en signes latins. Le problème est que toutes ces polices sont a priori au format METAFONT, donc produisent des polices de Type 3 dans le fichier PDF, d’où l’affichage illisible dans Acrobat Reader. Que faire ?

La première solution est de ne pas utiliser les polices METAFONT. On peut par

TAB. 1 – Quelques combinaisons utiles

Police	Type	Préambule \LaTeX	Commande dvips
Times	Sans maths	<code>\usepackage[T1]{fontenc}</code> <code>\usepackage{times}</code>	<code>dvips fichier.dvi -o</code>
Times	Avec maths	<code>\usepackage[T1]{fontenc}</code> <code>\usepackage{times}</code> <code>\usepackage{mathptm}</code>	<code>dvips -Ppdf fichier.dvi -o</code>
CM	Sans maths	<code>\usepackage{ae,aecompl}</code>	<code>dvips -Ppdf fichier.dvi -o</code>
CM	Avec maths	<code>\usepackage{ae,aecompl}</code> <code>\usepackage{mathptm}</code>	<code>dvips -Ppdf fichier.dvi -o</code>

exemple faire `\usepackage{times}`, qui va utiliser les polices Times, Helvetica, Courier et Symbol, polices PDF de base. Pour les symboles mathématiques, on peut soit utiliser la police commerciale MathTime, conçue pour s’harmoniser avec Times, soit utiliser `\usepackage{mathptm}` ou `\usepackage{mathptmx}`. Malheureusement, dans ce dernier cas, certains symboles mathématiques sont pris dans les polices mathématiques Computer Modern, il faudra donc appliquer les méthodes du paragraphe suivant.

La seconde solution est de forcer l’utilisation de polices Type 1 à la place des polices METAFONT. Heureusement pour nous, des polices Computer Modern Type 1 ont été réalisées et sont livrées avec les distributions \TeX modernes. Pour les utiliser, on lancera `dvips -Ppdf fichier.dvi -o` ou `dvips -Pwww fichier.dvi -o`. Il y a cependant un hic : il n’existe pas de polices European Computer Modern Type 1 gratuites ! On utilisera donc `\usepackage{ae,aecompl}`, qui tente de les émuler tant bien que mal (`aecompl` rajoute les guillemets français, hélas non disponibles en Type 1!).¹ En ce qui concerne les polices mathématiques, on utilisera `\usepackage[psamsfonts]{amsfonts}` ; en effet, certaines tailles de polices AMS ne sont pas disponibles gratuitement et devront donc être émulées.

4 Liens hypertextes et autres fonctionnalités

Les liens hypertexte en \LaTeX sont gérés par HyperRef. Celui-ci doit être paramétré suivant la méthode utilisée pour produire le PDF :

`\usepackage[dvips]{hyperref}` pour dvips + Acrobat Distiller ou ps2pdf

`\usepackage[dvipdfm]{hyperref}` pour dvipdfm

`\usepackage[pdflatex]{hyperref}` pour pdflatex

HyperRef fournit quatre types de services :

- Les renvois (numéros de notes de bas de page, de chapitres, de sections, d’équations, références bibliographiques...) deviennent automatiquement des liens hypertextes.
- Les URL (`\usepackage{url}`, commande `\url{http://www.exemple.org}`) sont actives (cliquer dessus lance un navigateur WWW).

¹Utilisez la version 1.3 du package `ae`.

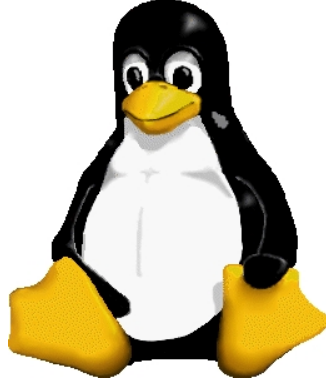


FIG. 2 – Un exemple d’image JPG incluse facilement

- La table des matières PDF est automatiquement générée suivant celle du document.
- On peut facilement régler les commentaires PDF (auteur, titre...).

Notons aussi l’utilisation de `dvipdft` au lieu de `dvipdfm`, qui va produire en plus les petites vignettes représentant les pages en réduction (un peu lent, à réserver pour la version finale du document).

5 Inclusions d’images

Un des principaux problèmes lors de la production de fichiers PDF est l’inclusion d’images et autres dessins par le document. En effet, la façon standard d’inclure une image lorsque l’on utilise `dvips` est de l’inclure au format PostScript ; c’est bien naturel puisque le fichier résultant sera lui aussi au format PostScript. Cela fonctionne avec `dvips` suivi de `Distiller` ou `ps2pdf`, mais pas avec `pdfLATEX`.

Tout d’abord, il faut savoir que PDF accepte directement les images JPEG, PNG et TIFF. Lorsque vous utilisez `dvipdfm` ou `pdfLATEX`, vous pouvez directement inclure les images :

```
\usepackage[dvipdfm]{graphics} % ou [pdfLATEX]  
\includegraphics{fichier.png}
```

Ensuite, `dvipdfm` accepte l’inclusion d’images PostScript, qu’il va convertir au vol en PDF. Il suffit donc de faire ²:

```
\usepackage[dvipdfm]{graphics} % ou [pdfLATEX]  
\includegraphics{fichier.ps}
```

Le seul problème avec `dvipdfm` est que celui-ci est très limité en ce qui concerne le PostScript inséré directement dans le DVI (ordres d’effets spéciaux, graphiques). Il faudra donc éviter ce genre de manipulations autant que possible. Ainsi, si on utilise la sortie `psLATEX` de `gnuplot`, il faudra mettre une option forçant l’utilisation d’un fichier externe pour le PostScript :

```
set terminal postscript auxfile
```

²Le fichier `dvipdfm.def` livré avec `LATEX` est souvent obsolète, car il ne prend pas en compte cette possibilité des versions récentes de `dvipdfm`. Le remplacer par celui livré avec `dvipdfm`.

6 En bref

```
\usepackage[dvipdfm]{graphics} % ou graphicx
\usepackage{ae,aecompl}
\usepackage[psamsfonts]{amsfonts}

latex fichier puis dvipdfm fichier.
```