pdfTeX und teTeX unter Linux (S.u.S.E. 5.3) — Ein Installations– und Testbericht*

Jürgen Prenzel

19. Oktober 1998

Zusammenfassung

pdfTeX ist eine Variante von TeX, die statt dvi pdf erzeugen kann. Es wird eine Möglichkeit vorgetellt pdfTeX unter Linux neben teTeX zu installieren ohne die vorhande Installation zu gefährden. Der vorliegende Bericht wird gleichzeitig als Testmaterial genommen. Die Darstellung über PDF sieht am Ende genau so aus wie die über dvi und dvips. Eine vermutete größere Kompaktheit der pdf-Darstellung, betrug nur ca. 20%.

1 Einleitung

Das Portable Document Format (PDF) der Firma Adobe scheint sich gegen das ältere PostScript, das auch von Adobe stammt, immer mehr durchzusetzen. Ein Vorteil von PDF ist angeblich seine Kompaktheit im Vergleich zu PostScript. Ein Autor, der Dokumente in LaTeX verfaßt, die einem größeren Leserkreis online zugänglich sein sollen, muß damit rechnen, daß seine Leser auf ihren Windows Systemen vielleicht ein Programm zum Lesen von PDF aber nicht von PostScript installiert haben. Es gibt zwar gsview als freie Software mit komfortabler Bedienung unter Windows, aber nicht jede(r) ist bereit oder in der Lage ein zusätzliches Programm zu installieren, um bestimmte Dokumente ansehen zu können. Sogar das reine Weiterreichen von PostScript an einen Drucker, der das verarbeiten kann, wird manchmal zum Problem. Es scheint also wünschenswert, aus LaTeX PDF machen zu können.

Die Lösung, die Han The Thanh unter dem Namen pdfTeX entwickelt hat, arbeitet mit Ergänzungen zu TeX, die in dem eigentlichen TeX–Programm, also nicht nur in den Makros, vorgenommen wurden. Entsprechend muß das Programm virtex ersetzt werden, um diese Funktionalität nutzen zu können. pdfTeX ist noch im beta–Stadium. Es scheint dehalb angebracht, es nur als zusätzliche Variante zu installieren, ohne die vorhandene teTeX–Installation zu stören.

Dieser Bericht beschreibt eine Installation von pdfTeX in einem separaten Benutzer–Verzeichnis ohne root–Privilegien, also ohne Eingriff in teTeX. Basis ist ein Linux–System nach SuSe (5.3) mit teTeX.

Ein Ziel war, daß die gleichen LaTeX–Files sowohl mit dem unveränderten `latex` zu Erzeugung von `.dvi` wie mit dem neuen System zur

*Diese Version wurde mit LaTeX und dvips erstellt.

Erzeugung von `.pdf` benutzt werden können. Dies ist eine Vorsichtsmaßnahme, die wahrscheinlich überflüssig wird, wenn man sich auf pdfTeX voll verlassen kann. Denn es erzeugt wahlweise auch `.dvi` und soll also wohl eine Obermenge der Funktionalität des Standard–TeX–Programms realisieren.

Dieser Bericht dient gleichzeitig als Testbeispiel. Er wird deshalb um einiges Material wie Grafiken und Gleichungen angereichert und wird dann sowohl mit klassischem LaTeX über dvips als PostScript erzeugt wie mit pdfTeX.

2 Quellen

Die notwendigen Files befinden sich auf dem Göttinger ftp–Server `ftp.gwdg.de` in `pub/dante/systems/pdftex` und `pub/dante/fonts`. Hilfreich war ein Installationsbericht unter `http://linux.anderson.ucla.edu/pdftex/README-pdftex.html`. Aktuelle Informationen finden sich bei `http://www.tug.org/applications/pdftex`. Dort gibt es auch ein Manual für pdfTeX vom Juni 1998, das sich aber auf eine etwas neuere Version (0.12+) bezieht, als die, die hier getestet wurde (0.10)

3 Installation

Es wird ein neues Verzeichnis angelegt, in meinem Fall `/home/jprenze/pub/pdftex`, das im folgenden als PDFTEXDIR bezeichnet wird. Die folgenden Vorgänge laufen in PDFTEXDIR ab:

Die sources (`pdftex.tar.gz`) und das binary für Linux (`pdftex-bin.tar.gz`) werden ausgepackt.

Das File `texmf.cnf` wird in `texmf.cnf.dist` umbenannt. Eine Kopie von `/usr/lib/teTeX/texmf.cnf` wird erzeugt und an diese der hintere Teil von `texmf.cnf.dist` (ab "Part 3 ...") angehängt.

Jetzt wird das File `pdflatex.fmt` durch das folgende Skript (`ini.sh`) erzeugt:

```
#!/bin/sh
export TEXMF=/usr/lib/teTeX/texmf
export TEXMFDBS=$TEXMF
ln -sf $TEXMF/tex/latex/base/latex.ltx pdflatex.ltx
pdftex -ini pdflatex.ltx
```

Für die benötigten Fonts wird ein Unterverzeichnis `$PDFTEXDIR/fonts` eingerichtet. Dort werden in weiteren Unterverzeichnissen die `.pfb`– und `.tfm` –Files der gewünschten PostScript–Type1–Schriften abgelegt. Für die LaTeX–Nutzer, die bei dem gewohnten Standard bleiben wollen, sind vor allem die Computer Modern (`cm`) Schriften wichtig. Für den vorliegenden Bericht wurden die bakuma–cm–Fonts installiert.

Als wichtig haben sich auch die Makros herausgestellt, die auf `ftp.gwdg.de in pub/dante/macros/pdftex` zu finden sind. Sie wurden in eine neues Unterverzeichnis `$PDFTEXDIR/macros` kopiert zusammen mit Kopien von `color.cfg` und `graphics.cfg`, die wie in `pdftex.def` vorgeschlagen manipuliert wurden. Tatsächlich gebraucht wurden diese Makros für die schrägen Beschriftungen in Abb. 2 (`rotating`).

Jetzt kann pdfTeX arbeiten, wenn man das folgende Skript (`pdflatex.sh`) anstelle von `latex` aufruft:

Abbildung 1: Ein farbiges Bild, das als PostScript Vektor–Grafik mit `graphicx` eingebunden wurde

```
#!/bin/sh
export PDFTEXDIR=/home/jprenze/pub/pdftex
export TETEXDIR=/usr/lib/teTeX
export TEXMF=$TETEXDIR/texmf
export TEXMFDBS=$TEXMF
export TEXMFCNF=$PDFTEXDIR:$TETEXDIR
export TEXMFL=$PDFTEXDIR//
export TEXMFS=$TEXMFL:!!$TEXMF
export TEXFORMATS=$PDFTEXDIR:!!$TEXMF
export T1FONTS=$PDFTEXDIR/fonts/type1//:$TEXMFS/fonts/type1//
export TFMFONTS=$PDFTEXDIR/fonts/type1//:$TEXMFS/fonts/tfm//
export TEXPSHEADERS=.:$PDFTEXDIR/web2c/libpdf/etc
$PDFTEXDIR/pdftex -fmt pdflatex $1
```

In einer anderen Umgebung muß jeweils die Definition von `PDFTEXDIR` angepaßt werden. Das Skript kommuniziert über die gesetzten Umgebungsvariablen mit kpathsea, dem File–Finder von TeX. Es ist sicherlich noch nicht ganz ausgereift. Man kann es dann, wenn kpathsea Probleme bekommt,d. h. benötigte Files nicht findet, über Versuch und Irrtum weiterentwickeln. Dazu ist das Studium der Files `texmf.cnf` und `texmf.cnf.dist` und der kpathsea–Dokumentation anzuraten.

Der Bequemlichkeit halber kann man in einem Verzeichnis, das in `$PATH` liegt, einen symbolischen link namens `pdflatex` auf `pdflatex.sh` einrichten.

Die Eingabefiles für `pdflatex` müssen besondere Befehle enthalten, sonst wird nur die gleiche dvi–Ausgabe erzeugt, wie mit normalem `latex`.

Dies sind vor allem:

```
\pdfoutput=1\relax
\pdfpagewidth=8.26in
\pdfpageheight=11.69in
\pdfcompresslevel=9
```

Der erste Befehl schaltet die PDF–Ausgabe ein. Die beiden nächsten stellen auf DIN A4 – Papier um und der letzte sorgt für maximale Kompression der Ausgabe.

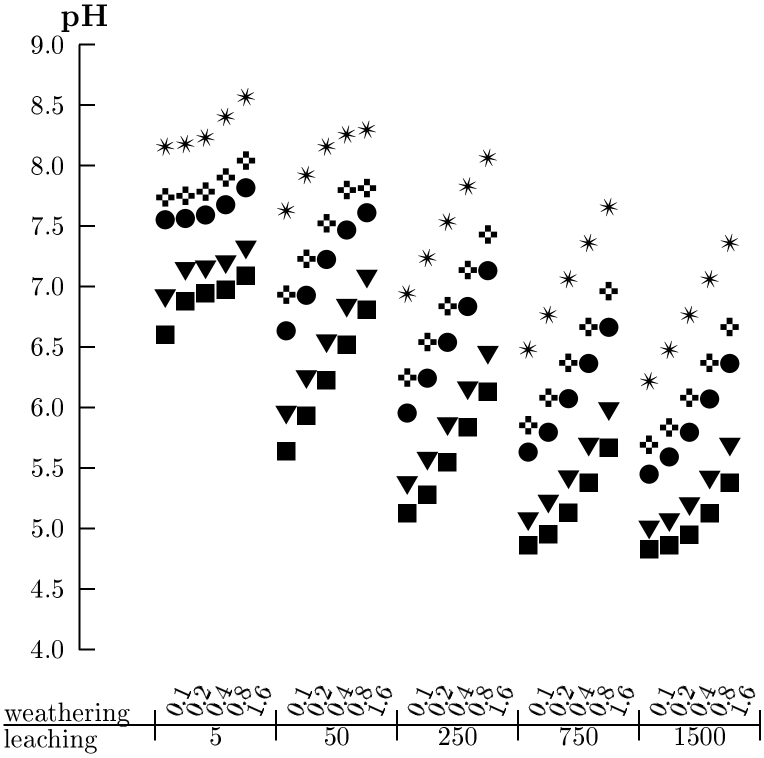

Abbildung 2: Ein Diagramm, das mit `epic` erzeugt wurde, unter Verwendung von `pifont` und `rotating`

Diese Befehle dürfen natürlich nicht enthalten sein, wenn das normale `latex` aufgerufen wird. Um das Umschalten zu erleichtern, verwende ich einen Schalter in der Präambel:

```
%\newcommand{\ifpdf}[2]{#1} % pdf on
\newcommand{\ifpdf}[2]{#2} % pdf off
```

der z.B. im Titel so verwendet wird:

```
Diese Version wurde mit \ifpdf{pdfTeX}{LaTeX und dvips} erstellt.
```

Weitere Abweichungen zwischen Eingabetext für `latex` und für `pdftex` entstehen bei der Einbindung von Grafiken.

4 Testmaterial

Die folgenden beiden Gleichungen und die beiden Abbildungen dienen dazu, die Fähigkeiten von pdfTeX und das Ausreichen der vorhandenen Fonts zu testen.

$$\mathsf{Al_2Si_2O_5(OH)_4 + H_2O \rightleftharpoons 2\ Al(OH)_3 + 2\ SiO_2} \qquad (1)$$

$$r = \left\{ \begin{array}{rcl} 1 & : & IAP/K_{\mathrm{sp}} < 0.1 \\ 0.1 & : & 0.1 < IAP/K_{\mathrm{sp}} \leq 1 \\ 0 & : & IAP/K_{\mathrm{sp}} \geq 1 \end{array} \right. \qquad (2)$$

Tabelle 1: Filegrößen verschiedener Darstellungsformen dieses Berichts

Behandlung	Form	kb
latex⇒dvips	ps	184
pdflatex	pdf	144
latex⇒dvips⇒Destiller	pdf	136
pdflatex⇒Acrobat (print)	ps	280
pdflatex⇒gzip	pdf.gz	104
latex⇒dvips⇒gzip	ps.gz	72

5 Testbericht

Schwierig war vor allem die Einbindung des Tigers (Abbildung 1). PDF stammt zwar von PostScript ab, kann aber keine frei geschriebenen Programme in PostScript interpretieren, bzw. die PDF–Interpreter wie Acrobat sind keine PostScript–Interpreter. Die Grafik muß jeweils vor Einbindung in PDF gerastert werden. Dies geschah hier über die Hilfsprogramme `pstopnm` und `pnmtopng`. Hinter dem ersten verbirgt sich das bekannte ghostscript und hinter dem zweiten eine Programmbibliothek zur Komprimierung und Manipulation von Rasterbildern (libpng, zlib), die auch von pdfTeX intern verwendet wird.

Die Grafik, die mit der LaTeX `picture` Umgebung hergestellt wurde, lies sich dagegen problemlos in PDF überführen. Aus die mathematischen Zeichen und die Zapfdingbats wurden gefunden.

Fraglich bleibt der Vorteil, den man von der Verwendung von PDF in diesem Kontext hat. Der Vorsprung von PDF bei der Kompression eines zu übertragenden Dokuments ist im Fall dieses Berichts nicht sehr groß .Die verschiedenen Darstellungsformen dieses Berichts hatten (kurz vor der Fertigstellung) die in Tab. 1 angegebenen Größen. Dabei wurden die Adobe Programme AcrobatReader und Destiller in Version 3.0 unter Windows NT benutzt. Die Benutzung von pdfTeX führt zu einer Einsparung auf 144 kbyte gegenüber 184 kbyte in dem von dvips erzeugten PostScript. Wahrscheinlich ist dieses PostScript schon relativ kompakt. Das sieht man im Vergleich zu der Ausgabe, die der AcrobatReader an den Drucker schickt (280 kbyte). Ein Nachteil für pdfTeX scheint zu sein, daß Rastergrafiken eingebunden werden müssen. Allerdings ist der Tiger hier im png–Format nur 20 kbyte groß gegenüber 80 im PostScript Vektor Format. PdfTeX bietet keinen wesentlichen Nachteil gegenüber dem Weg über dvips und den Adobe Destiller (144 statt 136 kbyte). Falls es aber auf Kompaktheit wirklich ankommt, bietet gzip in Verbindung mit der Ausgabe von dvips die besten Möglichkeiten.

Das Argument für PDF kann also kaum die Kompaktheit der internen Darstellung eines Dokuments wie des vorliegenden Berichts sein. Es mag andere Gründe geben, die jemanden veranlassen PDF gegenüber PostScript zu bevorzugen.

Die grafische Qualität der verschiedenen erzeugten Druckausgaben war nicht unterscheidbar. Für viele mit LaTeX erzeugte Dokumente bietet also wohl pdfTeX eine gute Möglichkeit, PDF zu erzeugen.

6 Installation bei der GWDG in Göttingen

Die Installation befindet sich bis auf weiteres im NFS Cluster der GWDG unter `/usr/users/jprenze/pub/pdftex`. PdfTeX ist als das dortige Skript `pdflatex.sh` auf den Linux Rechnern im Cluster (gwdu101 etc.) ausführbar. Dieser Bericht einschließlich der LaTeX Eingabe (`instrep.tex`) ist unter doc/testbericht zu finden. Die Installation umfaßt 8.5 Mbyte, wobei 3.7 Mbyte die Fonts und 1.4 Mbyte die Dokumentation einschließlich diverser Fassungen dieses Berichts ausmachen.