Shrinkage Methods and PDEs

Workshop “Bildverarbeitung und Wavelets”
Dortmund, 24. – 25. April

Kristian Bredies, Dirk Lorenz, Peter Maaß

Overview

There are lots of different concepts in signal processing. For example:

- Linear/convolution filters:
$$u = f * g$$

- Diffusion or other partial differential equations like
$$u_t = \Delta u \quad \text{or} \quad u_t = \mathrm{div}(g(|\nabla u|)\nabla u)$$

- Nonlinear wavelet methods like soft or hard threshholding ("shrinkage").

Goal of this talk: Establish a connection between PDEs and shrinkage methods.

Outline of shrinkage methods

Take an invertible linear isometry

$$\mathcal{T} : L^2 \to L^2$$

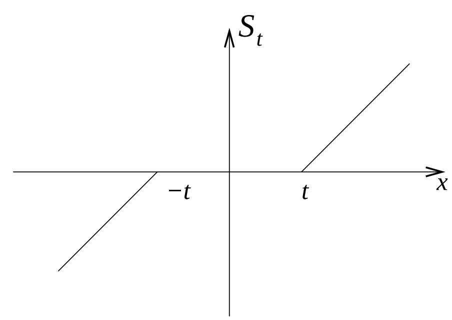
$$\text{resp. } \mathcal{T} : L^2 \to l^2$$

and define the $\mathcal{T}$-shrinkage of a function $f \in L^2$ by

$$u(t) = \mathcal{T}^{-1}(S_t(\mathcal{T}f))$$

with the shrinkage function

$$S_t(x) = \mathrm{sign}(x)(|x| - t)_+$$

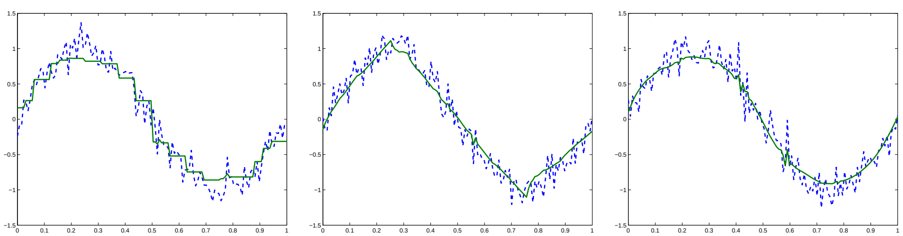


Example: Discrete wavelet shrinkage

For an orthogonal wavelet ψ the discrete wavelet-shrinkage looks like

$$u(t) = \sum_{j,k} S_t(\langle f | \psi_{j,k} \rangle) \psi_{j,k}.$$

Pictures:



$\rightsquigarrow$ Haar-Wavelet $\rightsquigarrow$ db2-Wavelet $\rightsquigarrow$ db4-Wavelet

History of shrinkage methods

- Development of discrete wavelet-shrinkage for data smoothing (Johnstone and Donoho, 1994, 1995)
- Translation invariant denoising by redundant discrete wavelet-shrinkage (Coifman and Donoho, 1995)
- Discrete wavelet-shrinkage as an evolution equation resp. smoothing Scale-Space (Chambolle, DeVore, Lee, Lucier, 1998)
- Translation invariant discrete wavelet-shrinkage as Scale-Space (Chambolle, Lucier 2001)

Discrete Shrinkage and PDEs

Chambolle and Lucier showed that the discrete wavelet-shrinkage is the solution of a descent problem:

$$u(t) = \sum_{j,k} S_t(\langle f | \psi_{j,k} \rangle) \psi_{j,k} \quad \Leftrightarrow \quad u_t + \partial |u|_{B^1_{1,1}} \ni 0, \quad u(0) = f.$$

Short explanation:

To solve the descent problem, use backward differences in time:

$$\frac{u^n - u^{n-1}}{\Delta t} + \partial |u^n|_{B^1_{1,1}} \ni 0 \quad \Leftrightarrow \quad \text{minimize } \frac{1}{2\Delta t} ||v - u^{n-1}||^2_{L^2} + |v|_{B^1_{1,1}}.$$

To minimize the functional $\frac{1}{2\Delta t}||v-u^{n-1}||^2_{L^2}+|v|_{B^1_{1,1}}$ we use the wavelet expansion of u^{n-1} and v and the norm equivalence $|u|_{B^1_{1,1}} \asymp \sum_{j,k} |\langle u|\psi_{j,k}\rangle|$. So the minimization problem is:

$$\text{minimize } \sum_{j,k} \frac{1}{2\Delta t}\Big(\langle v|\psi_{j,k}\rangle - \langle u^{n-1}|\psi_{j,k}\rangle\Big)^2 - |\langle v|\psi_{j,k}\rangle|.$$

So we can compute the minimum for each wavelet coefficient separately.

We have to find the minimum of the convex function

$$E(x) = \frac{1}{2\Delta t}(x - x_0)^2 + |x|$$

and this is $x = S_{\Delta t}(x_0)$.

So we have

$$u^n = \sum_{j,k} S_{\Delta t}(\langle u^{n-1} | \psi_{j,k} \rangle) \psi_{j,k} = \sum_{j,k} S_{n\Delta t}(\langle f | \psi_{j,k} \rangle) \psi_{j,k}.$$

Going to the limit $n \to \infty$ while $n\Delta t = t$ yields

$$u(t) = \sum_{j,k} S_t(\langle f | \psi_{j,k} \rangle) \psi_{j,k}.$$

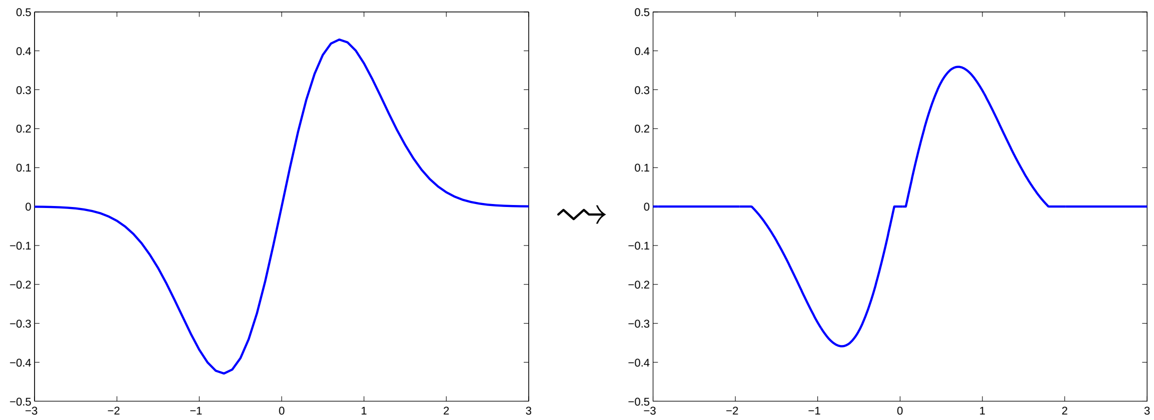
Is a result like this possible for other transformations $\mathcal{T}$ as the discrete wavelet transform?

Computation in the image domain of $\mathcal{T}$

Theorem: *For a given $f \in L^2$ the shrinkage $u(t) = S_t(f),\ t \geq 0$ is a subgradient descent:*

$$u_t + \partial ||u||_{L^1} \ni 0, \quad u(0) = f.$$

Where $||u||_{L^1} = \int |u|$ has to be understood as a functional on L^2.



How can this be seen?

Same arguments as above:

Use backward differences:

$$\frac{u^n - u^{n-1}}{\Delta t} + \partial ||u||_{L^1} \ni 0 \quad \Leftrightarrow \quad \text{minimize } \frac{1}{2\Delta t}||v - u^{n-1}||_{L^2}^2 + ||v||_{L^1}.$$

The minimizer is given by

$$u^n = S_{\Delta t}(u^{n-1})$$

and we have

$$u(t) = S_t(f).$$

Pulling back in the domain of $\mathcal{T}$

Now we see, that in the image domain of our transformation $\mathcal{T}$, the shrinkage is a descent along the L^1-Norm.

But how can we understand this in the domain of $\mathcal{T}$?

Lemma: *Let V, H be Hilbert spaces, $\mathcal{T} : V \to H$ be an invertible linear isometry, $\Phi : H \to]-\infty, \infty]$ be proper, convex and lower semi-continuous and $\Psi(u) = \Phi(\mathcal{T}u)$. Then solving*

$$u_t + \partial\Psi(u) \ni 0, \quad u(0) = f \text{ in } V$$

is equivalent solving

$$v_t + \partial\Phi(v) \ni 0, \quad v(0) = \mathcal{T}f \text{ in } H.$$

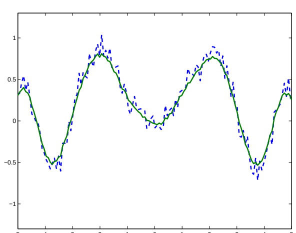
Fourier-shrinkage as a descent

The continuous Fourier-shrinkage is

$$u(t,x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} S_t(\mathcal{F}f(\omega)) e^{i\omega x} d\omega.$$

Applying the preceding lemma, we have, that the Fourier-Shrinkage is a descent along the subgradient of $\Psi(u) = \Phi(\mathcal{F}u) = ||\mathcal{F}u||_{L^1}$, i. e.

$$u_t + \partial ||\mathcal{F}u||_{L^1} \ni 0, \quad u(0) = f.$$

Picture: 

$\rightsquigarrow$Continuous Fourier shrinkage

Wavelet-shrinkage as a descent?

Unfortunately this is not working for the wavelet-shrinkage.

Problem: The wavelet-transform L_ψ is not onto and we can not use the lemma.

More precisely: For a signal u, $S_t(L_\psi u)$ is not in the image domain of the wavelet transform.

So in every time step we have to make a projection on the domain of the wavelet transform.

A more sophisticated argument:

In the Fourier case we had, that the shrinkage was a descent:

$$u_t + \mathcal{F}^{-1}\partial\Phi_F(\mathcal{F}u) \ni 0,\ u(0) = f$$

with $\Phi_F(v) = \int |v| dx$, moreover $\mathcal{F}^{-1}\partial\Phi_F(\mathcal{F}u) = \partial||\mathcal{F}u||_{L^1}$.

In the wavelet case it is not true, that ${L_\psi}^*\partial\Phi_W(L_\psi u) = \partial||L_\psi u||_{L^1(\frac{dadb}{a^2})}$.

With $\Phi_W(v) = \int |v|\frac{dadb}{a^2}$ we have

$${L_\psi}^*\partial\Phi_W(L_\psi u) \neq \emptyset \ \Leftrightarrow\ u = 0.$$

But the theory of continuous semi-groups say, that the descent problem $u_t + \partial\Phi(u) \ni 0$ has a unique continuous solution for every initial value $f \in \operatorname{dom}\Phi$, if Φ is convex, proper and lower semi-continuous.

Conclusion

Theorem: *Let $f \in L^2(\mathbb{R})$. Then a solution of the descent problem in the spatial domain*

$$u_t + \partial||\mathcal{F}u||_{L^1(\mathbb{R})} \ni 0, \quad u(0) = f$$

is given by the Fourier-shrinkage of f:

$$u(t,x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} S_t(\mathcal{F}f(\omega))e^{i\omega x} d\omega.$$

The descent problem in the wavelet domain

$$v_t + \partial||v||_{L^1(\mathbb{R}^2, \frac{dadb}{a^2})} \ni 0, \quad v(0) = L_\psi f$$

has a solution v for wich, with $u(t) = L_\psi^* v(t)$, the relation

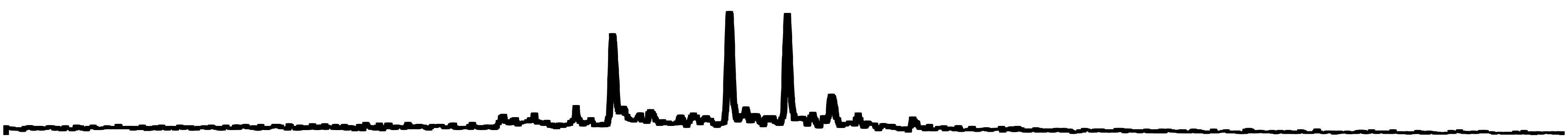
$$u(t,x) = \frac{1}{\sqrt{c_\psi}} \int_{\mathbb{R}} \int_{\mathbb{R}} S_t(L_\psi f(a,b))|a|^{-1/2} \psi\left(\frac{x-b}{a}\right) \frac{dadb}{a^2}$$

holds.

Application: Denoising of spectroscopy data

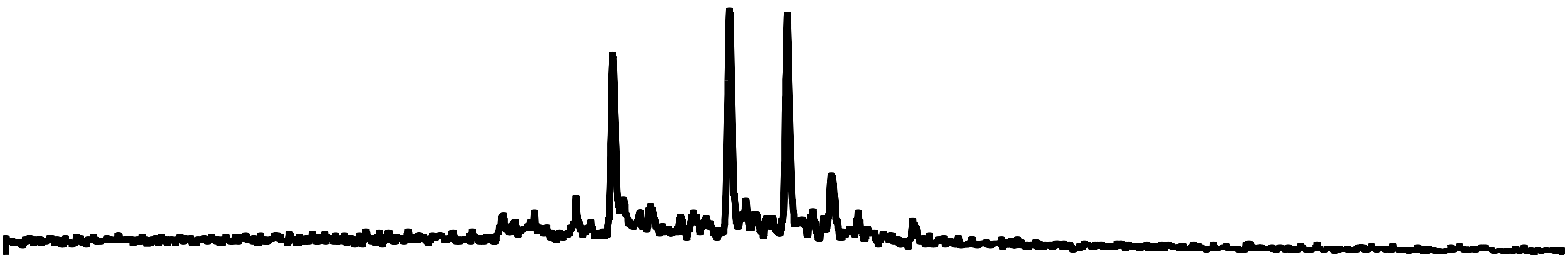
Bruker Daltonics is a manufacturer of mass spectrometry instruments and accessories for pharmaceutical, biochemical and chemical research.

Data of of spectometry of blood.

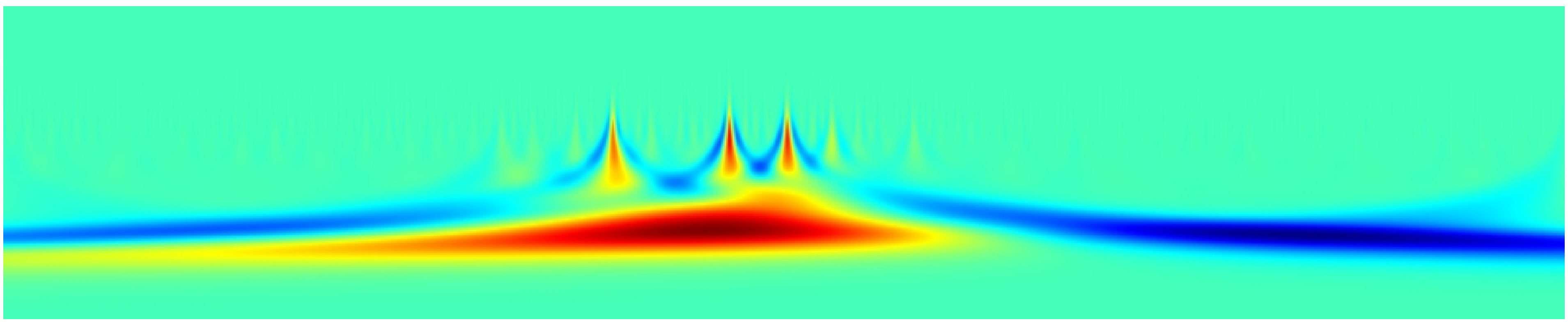


The peaks indicate different proteins. The concetration of the proteins shall be used for cancer diagnosis. Assumption: Cancer can be detected by “proteonics” before it can be seen be computer tomography.
Goal of preprocessing: Remove the noise and keep the “structure”.

The data and the denoising

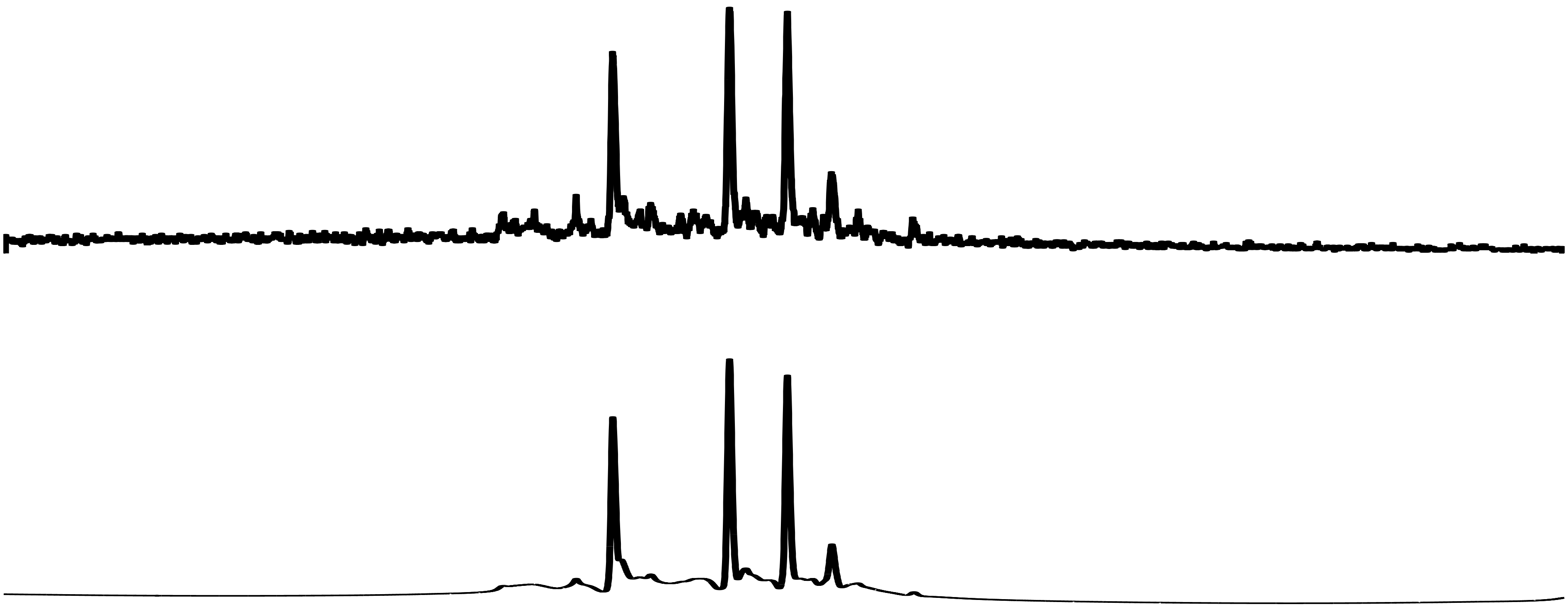


Continuous wavelet transform (mexican hat)

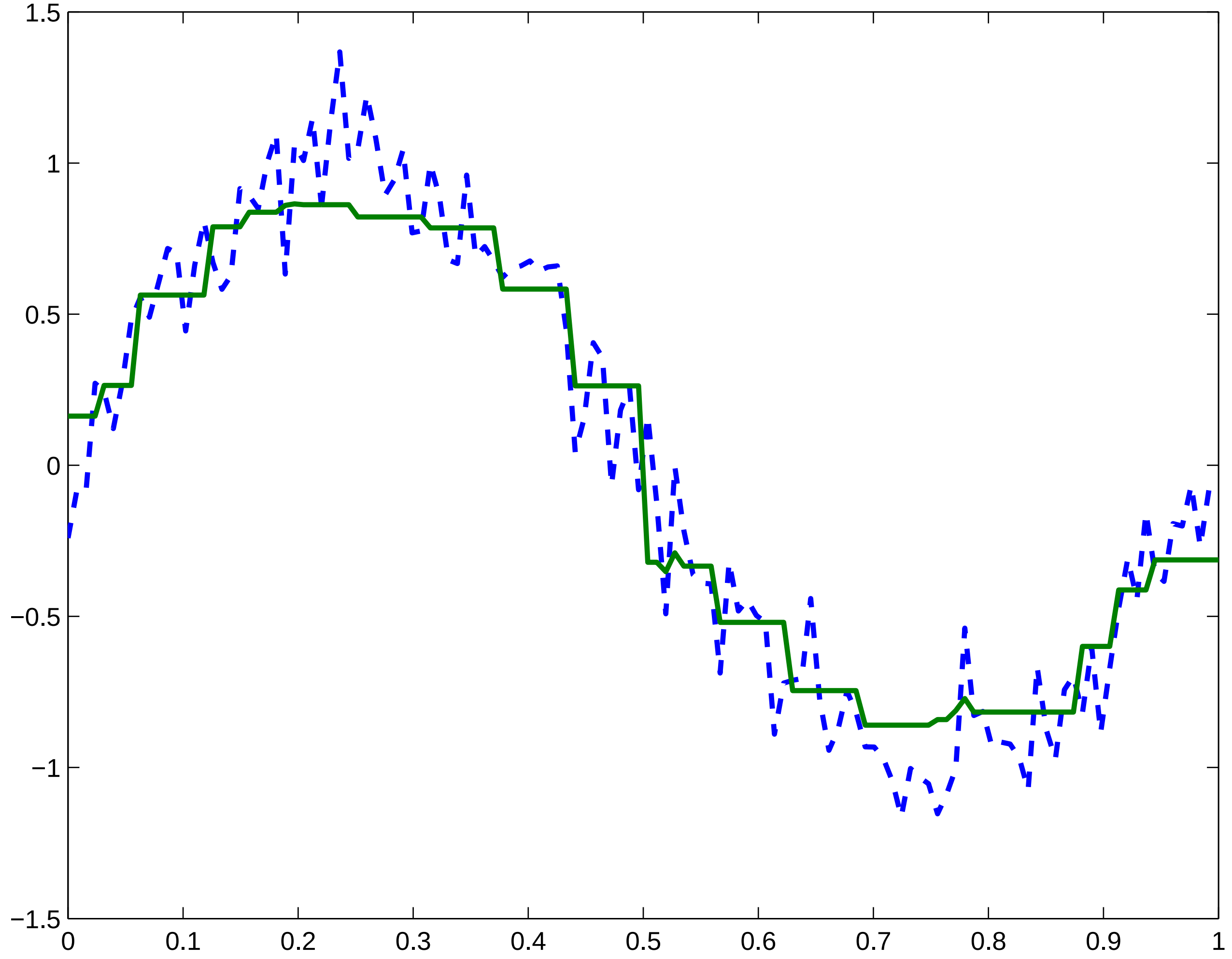


$\rightsquigarrow$The denoising process

The denoised data:

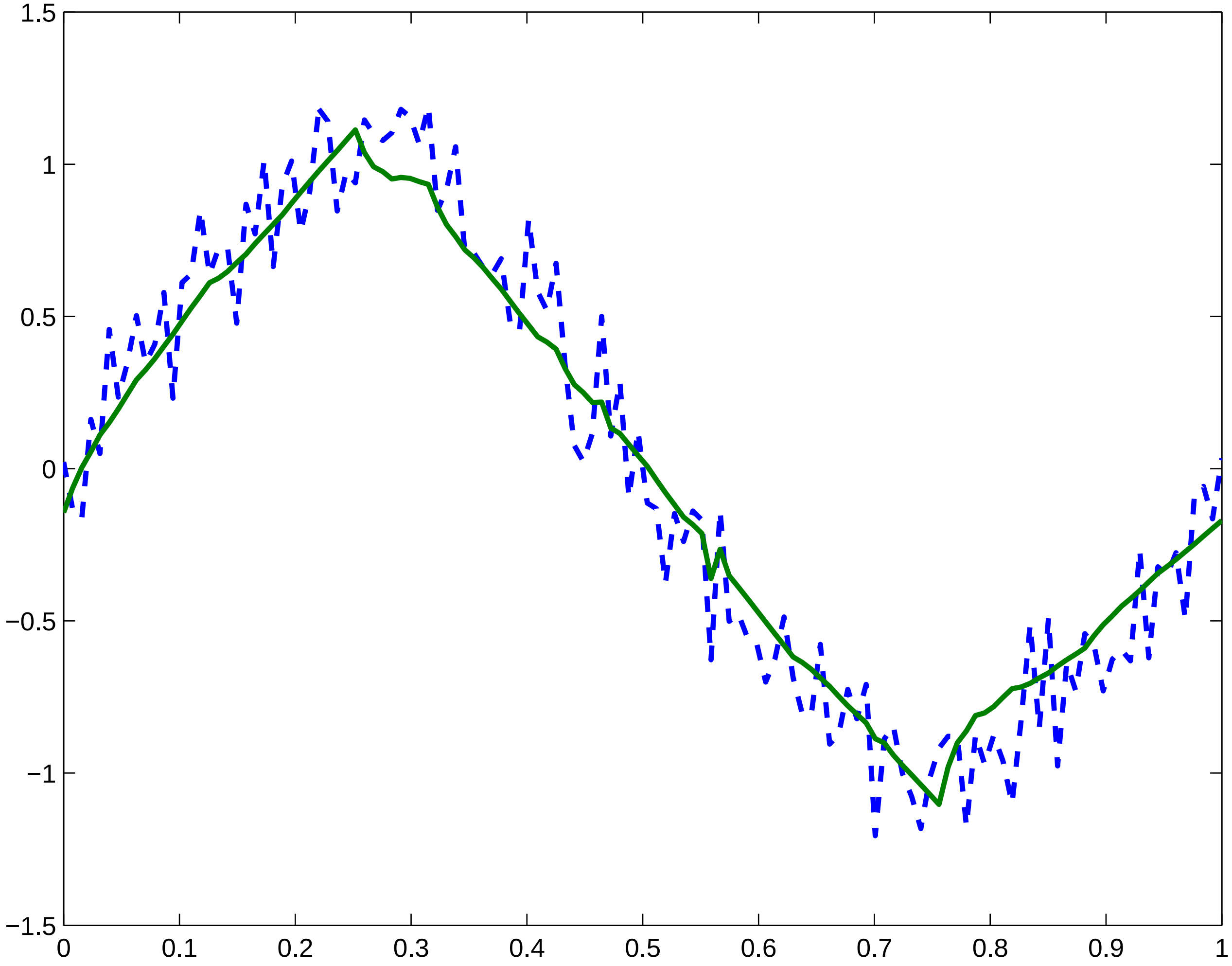


Wavelet-Shrinkage with the Haar-Wavelet



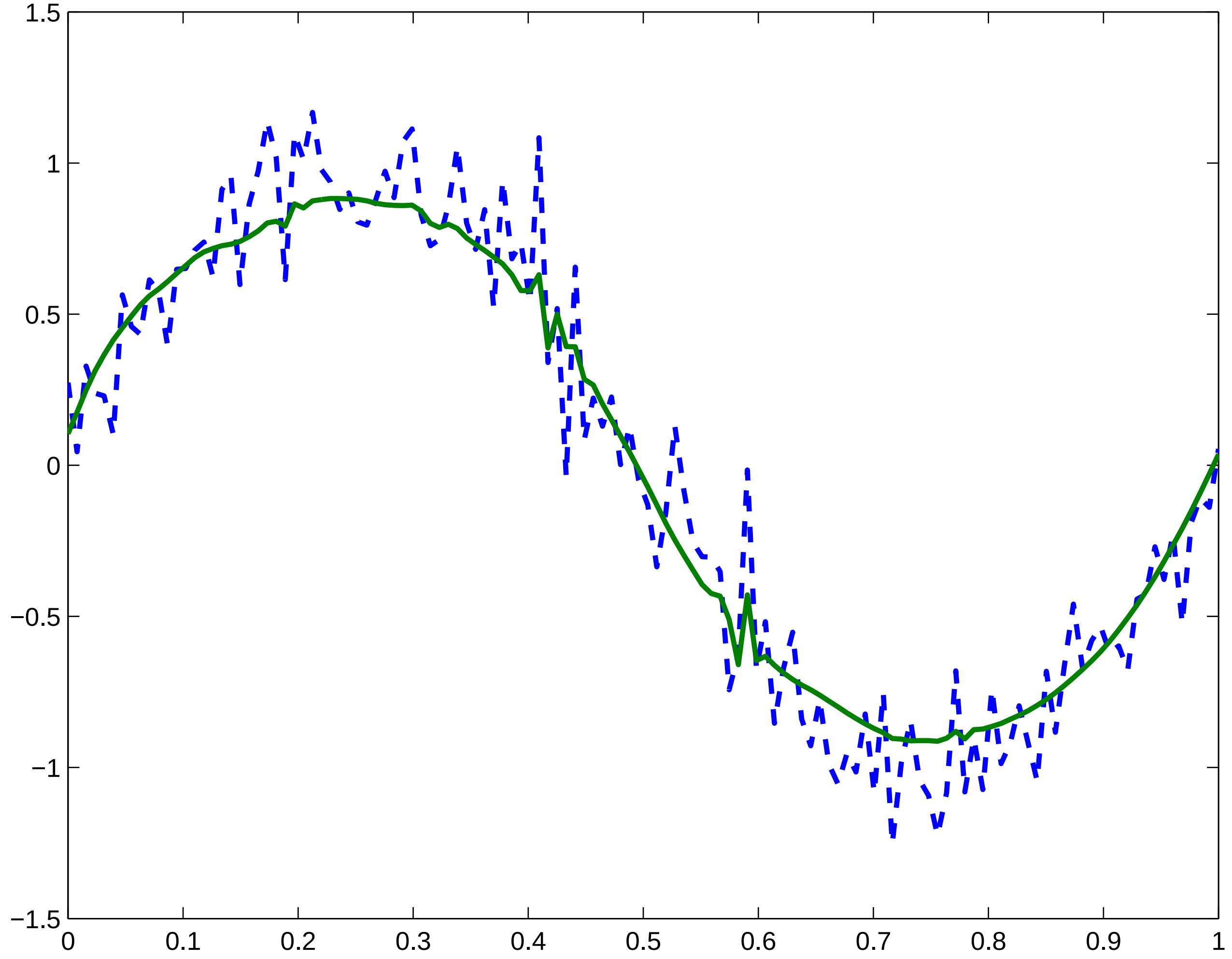
$\rightsquigarrow$Back

Wavelet-Shrinkage with the db2-Wavelet



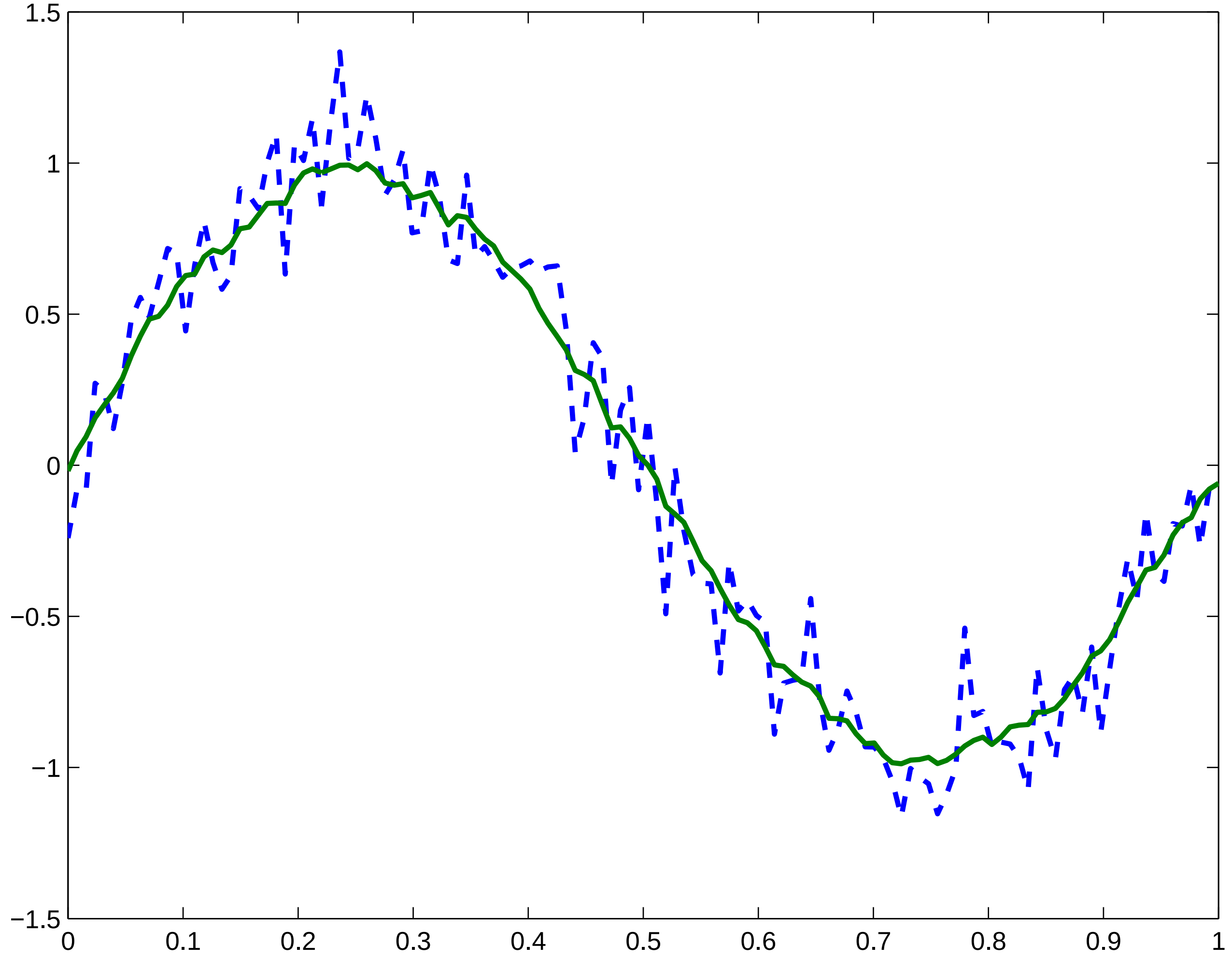
$\rightsquigarrow$Back

Wavelet-Shrinkage with the db4-Wavelet



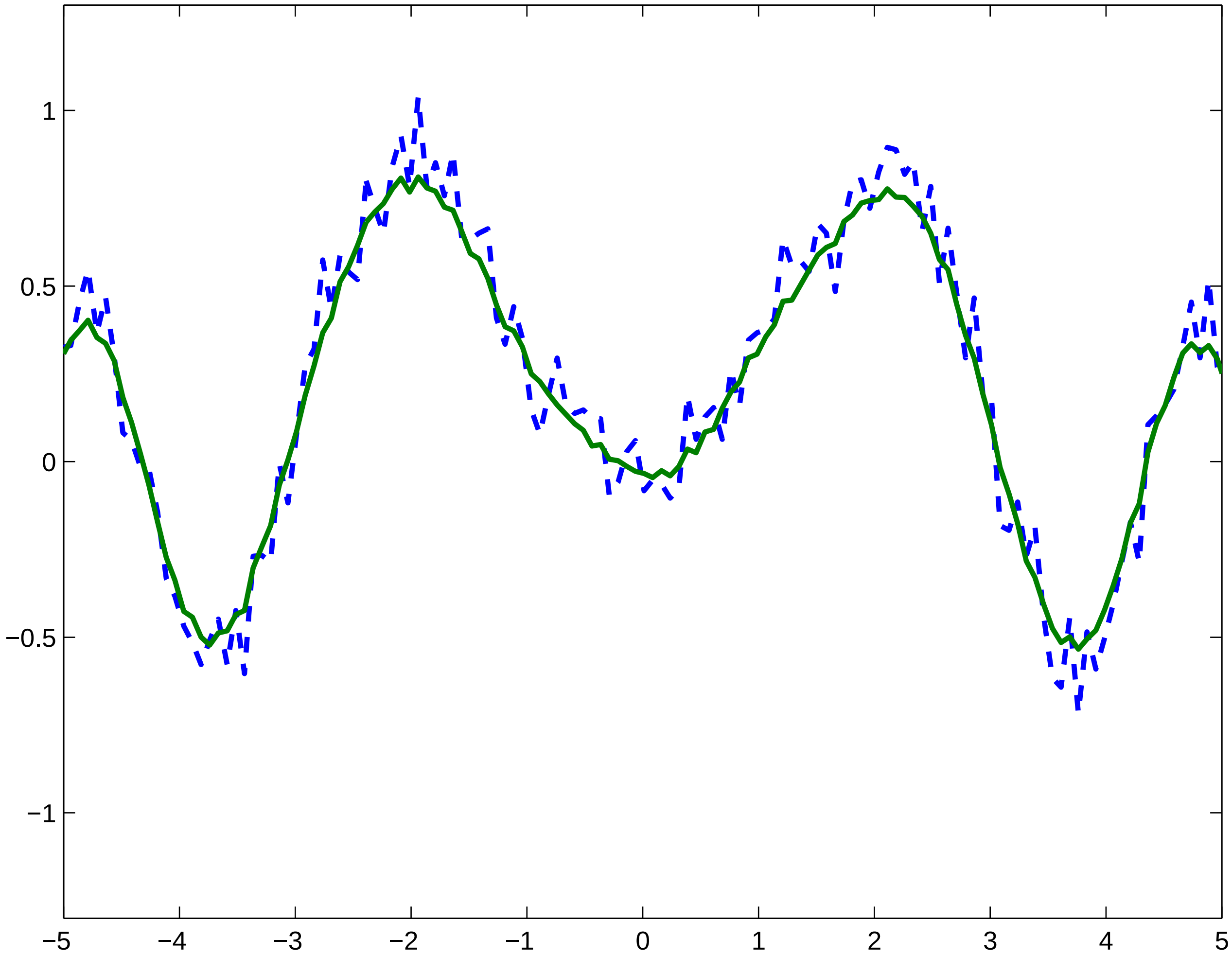
$\rightsquigarrow$Back

Fourier-Shrinkage

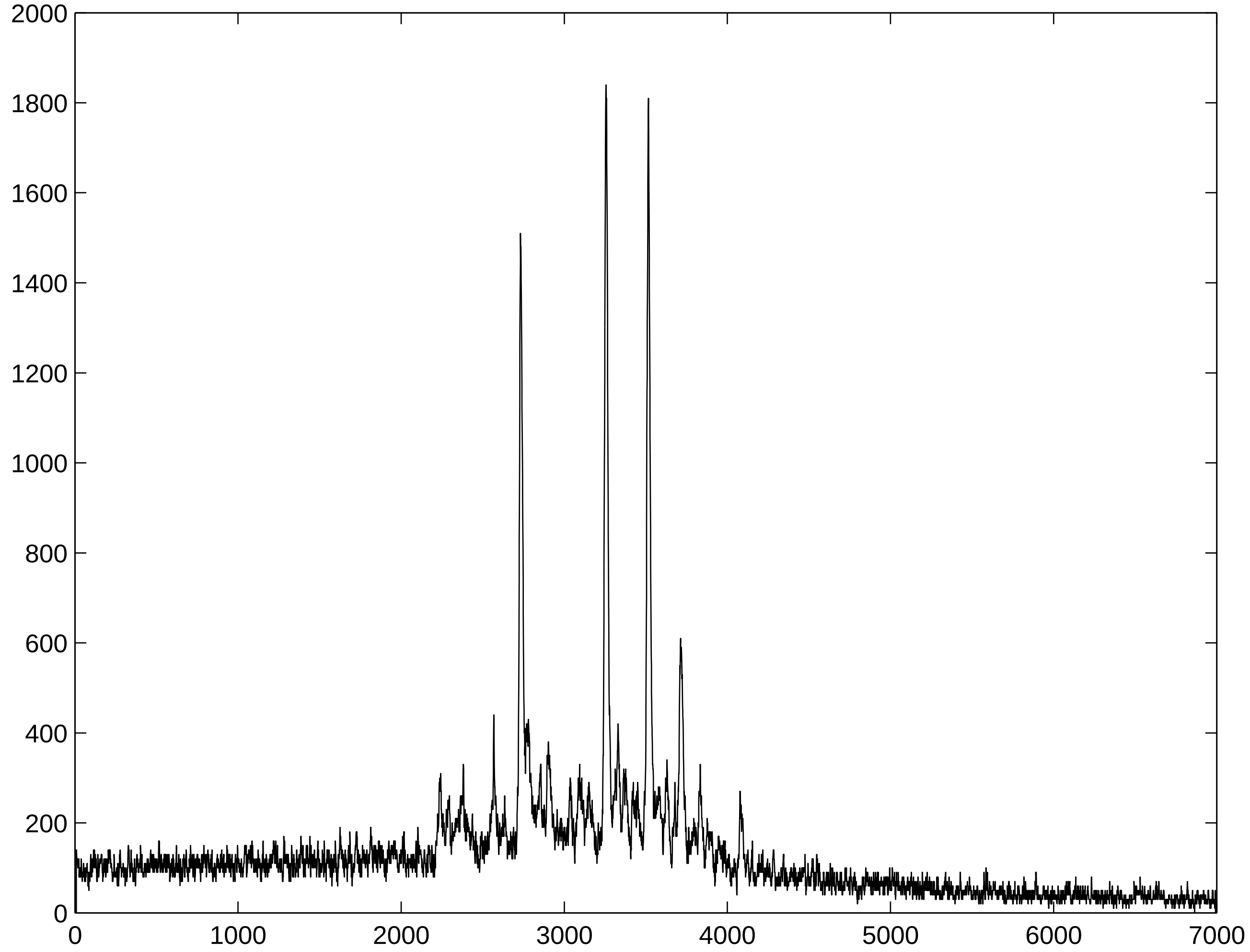


⇝Back

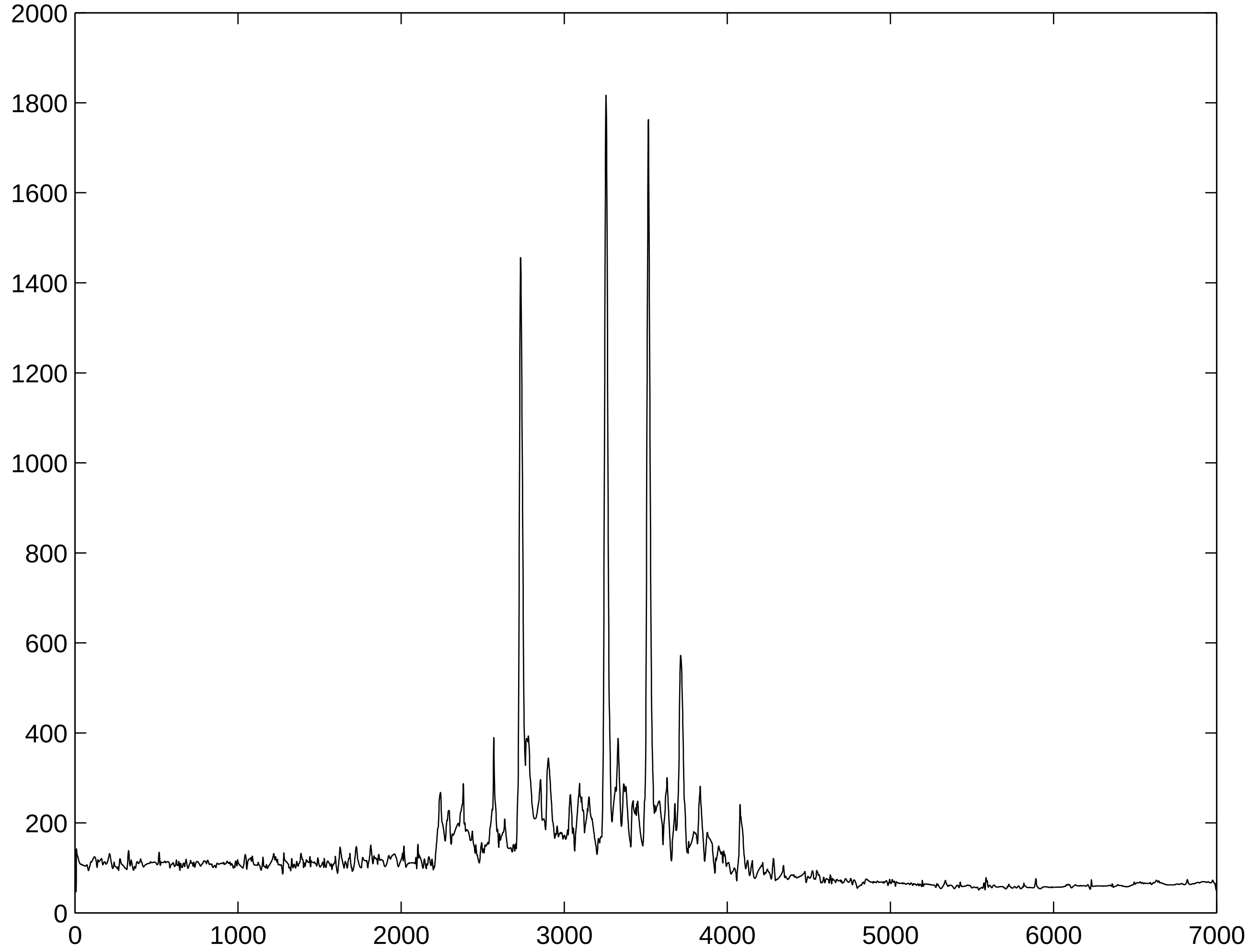
Continuous Fourier-Shrinkage



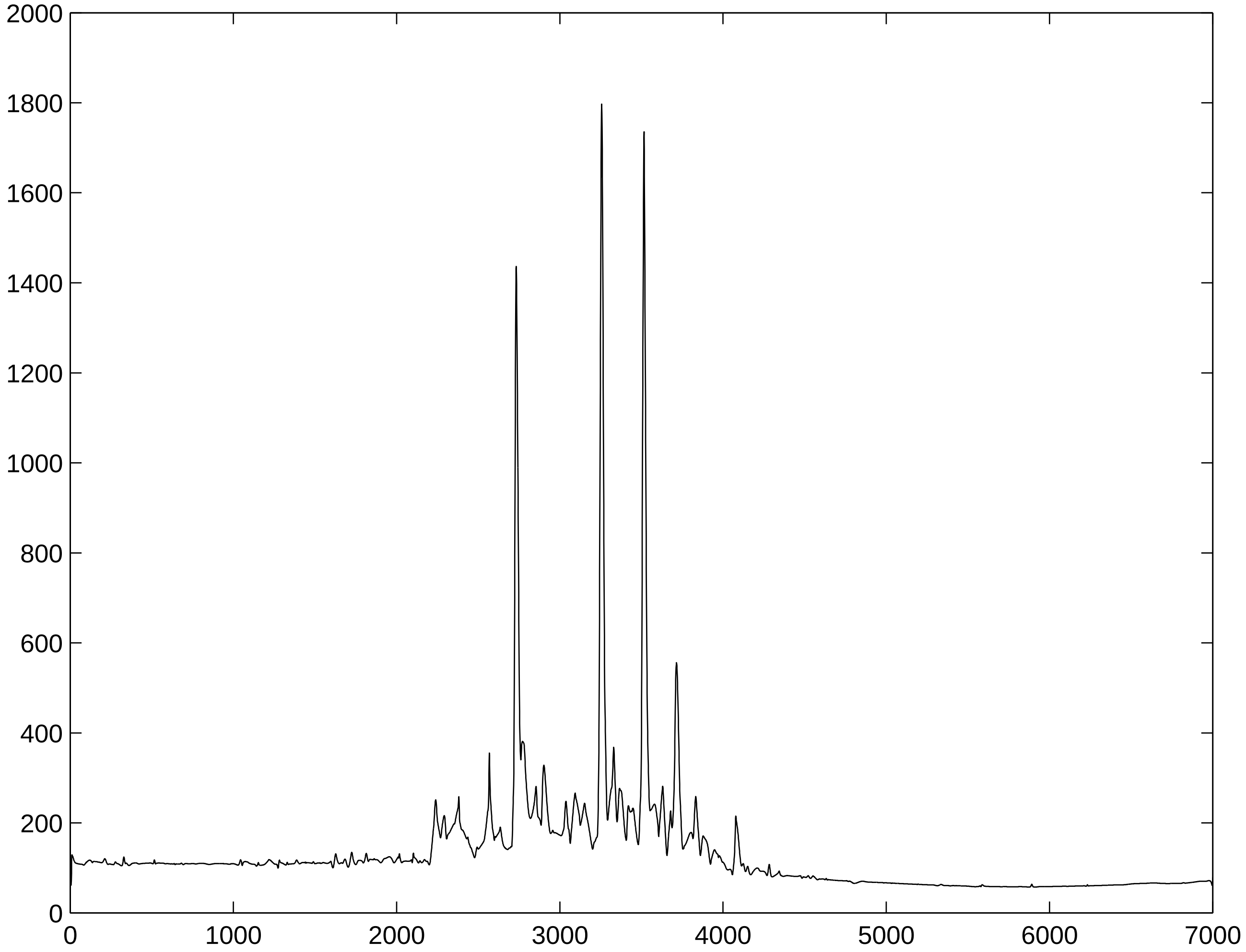
$\rightsquigarrow$Back


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

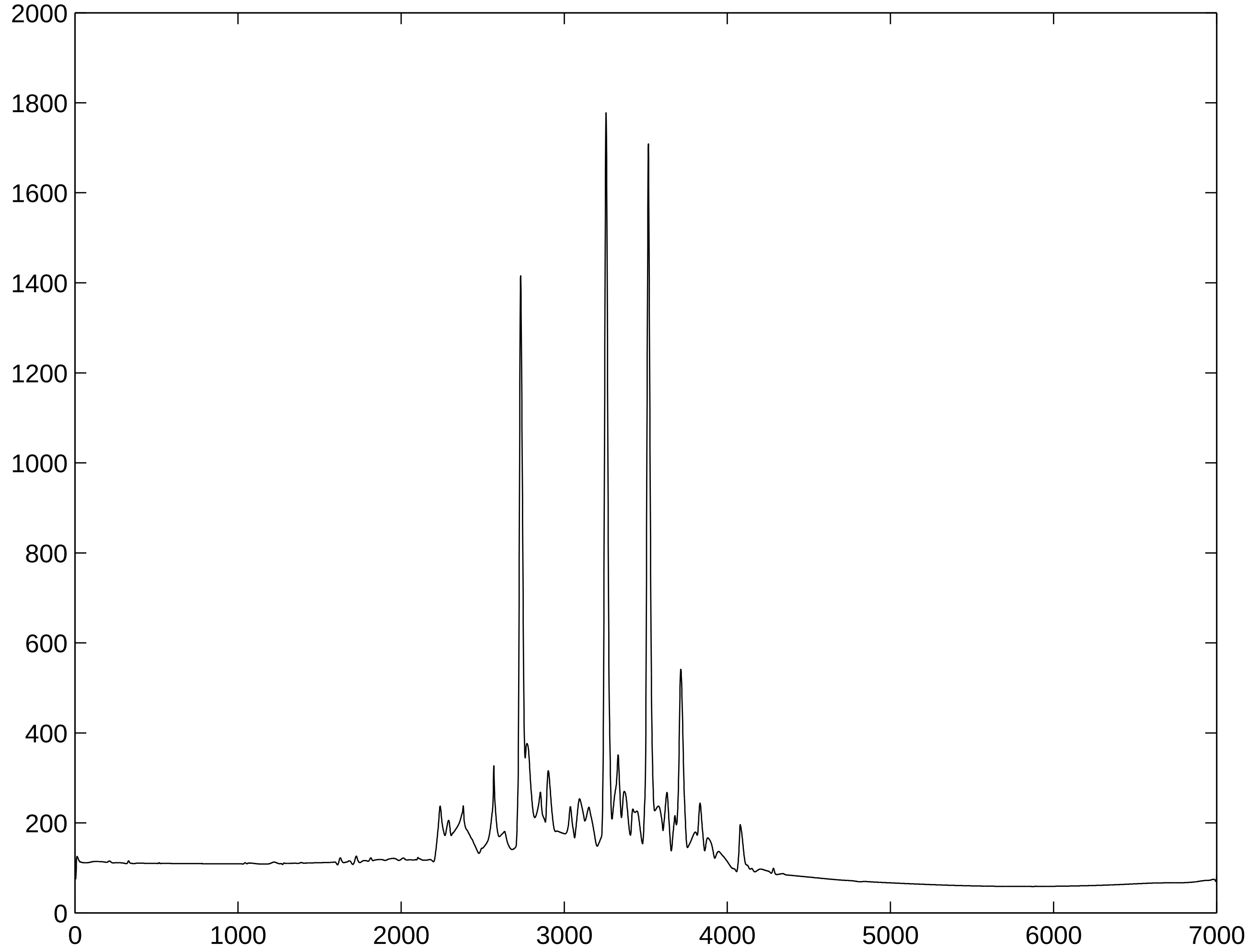
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

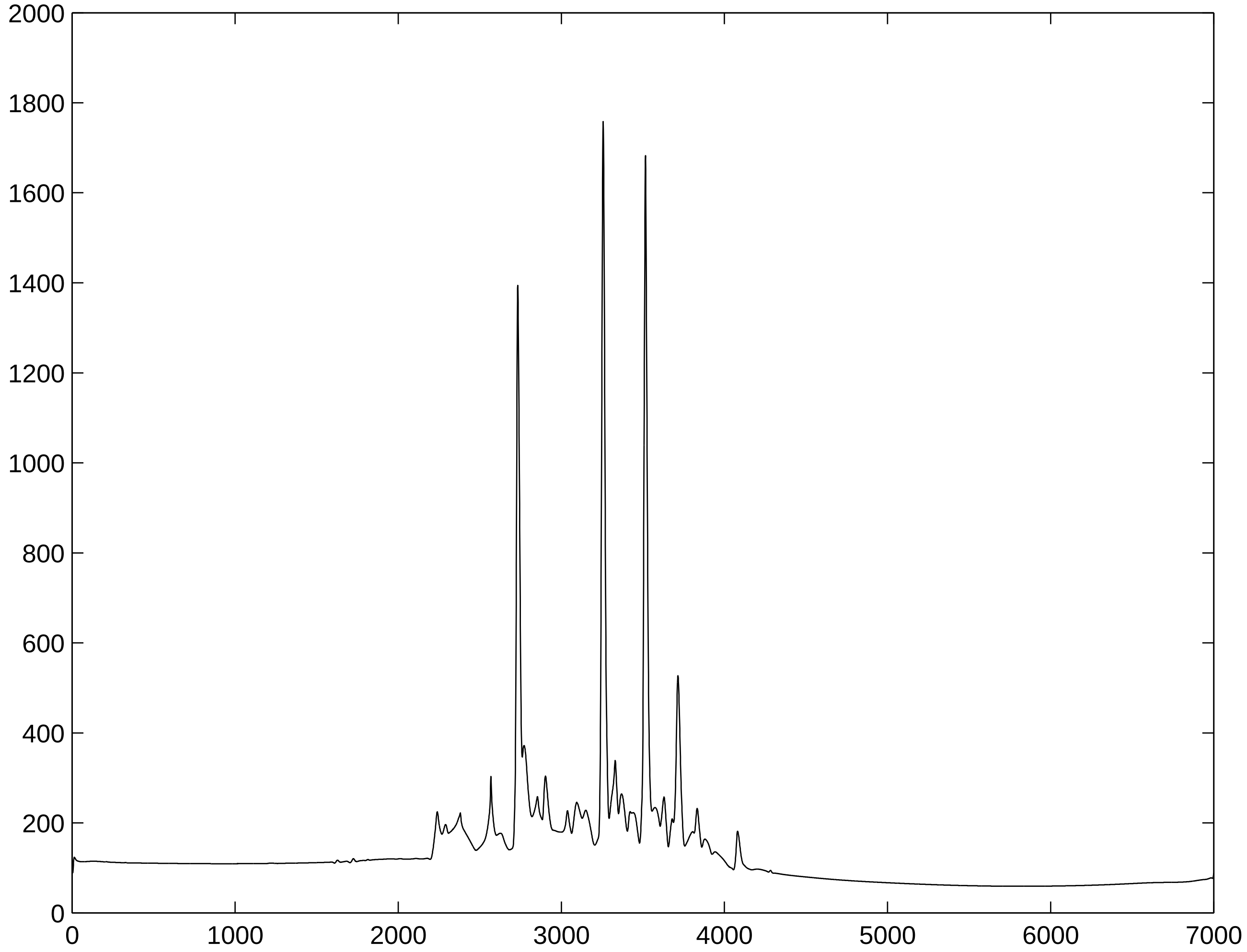
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

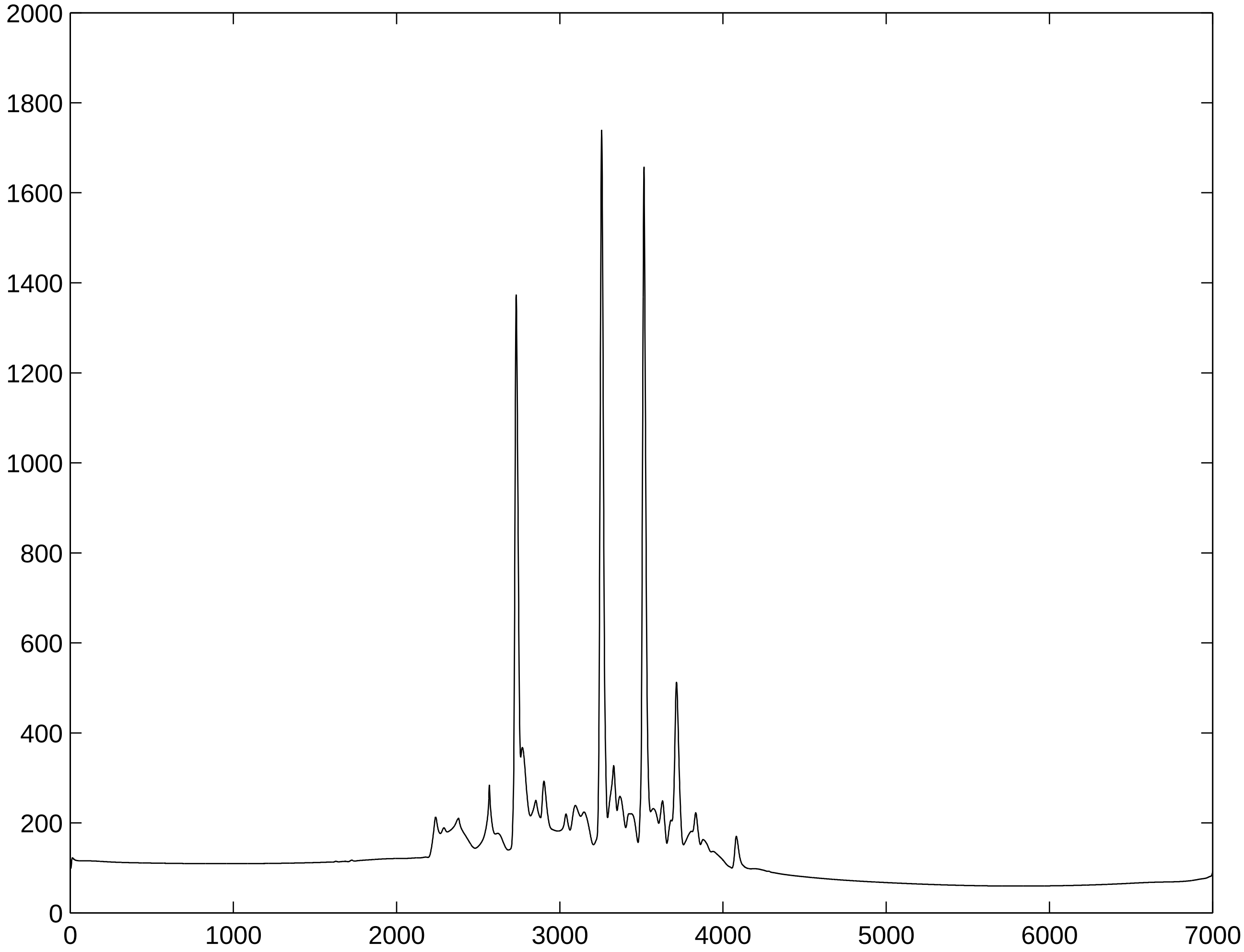
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

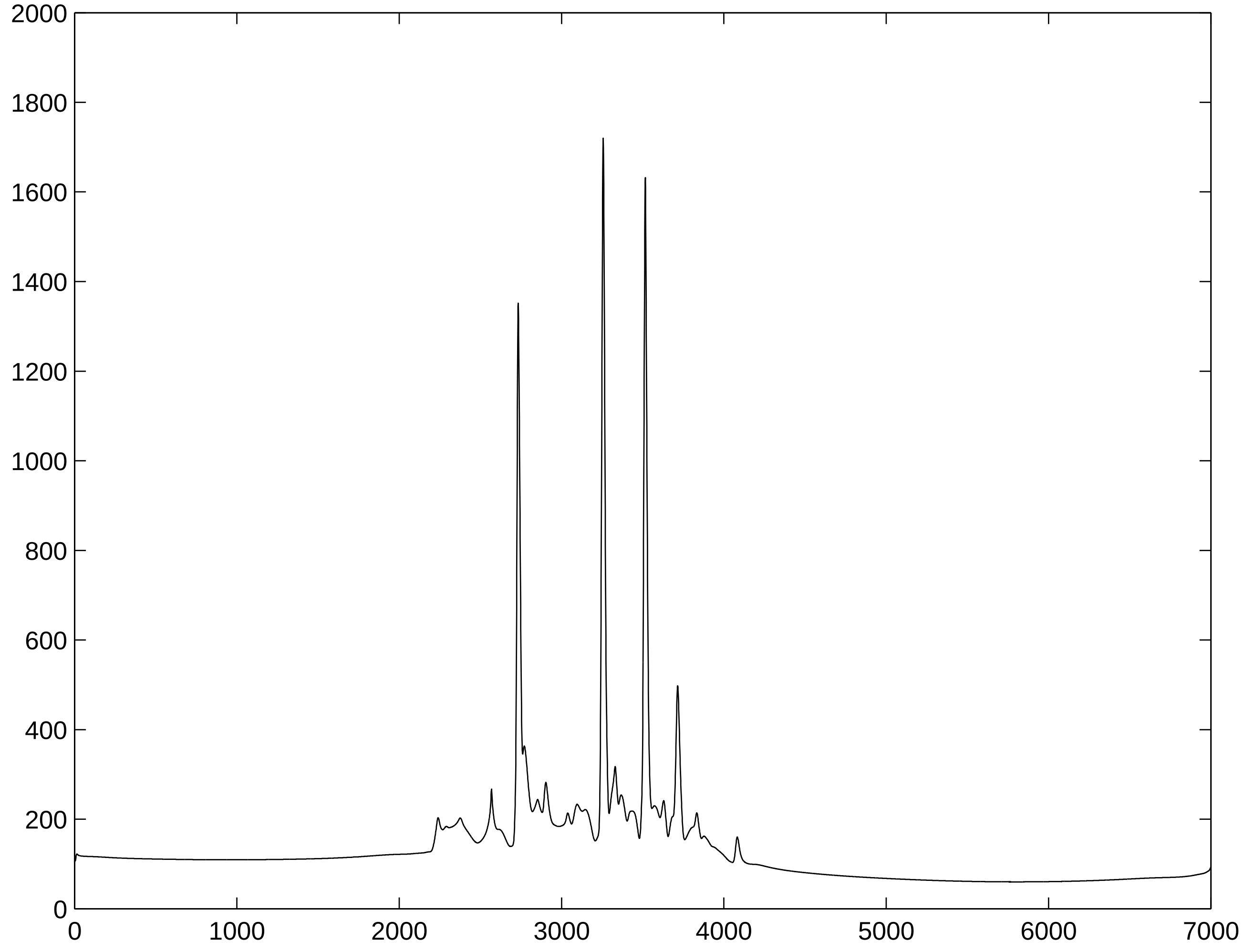
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

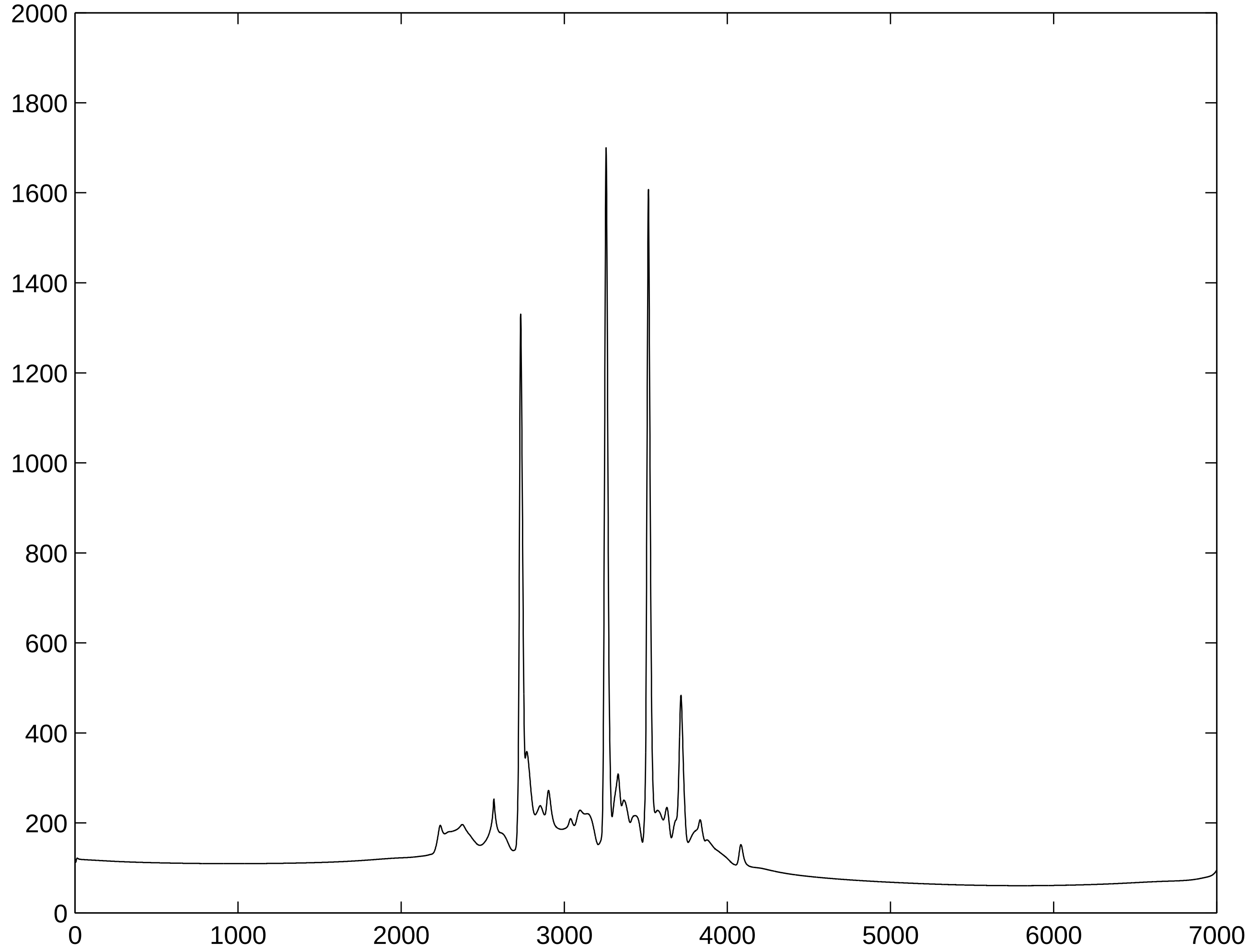
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

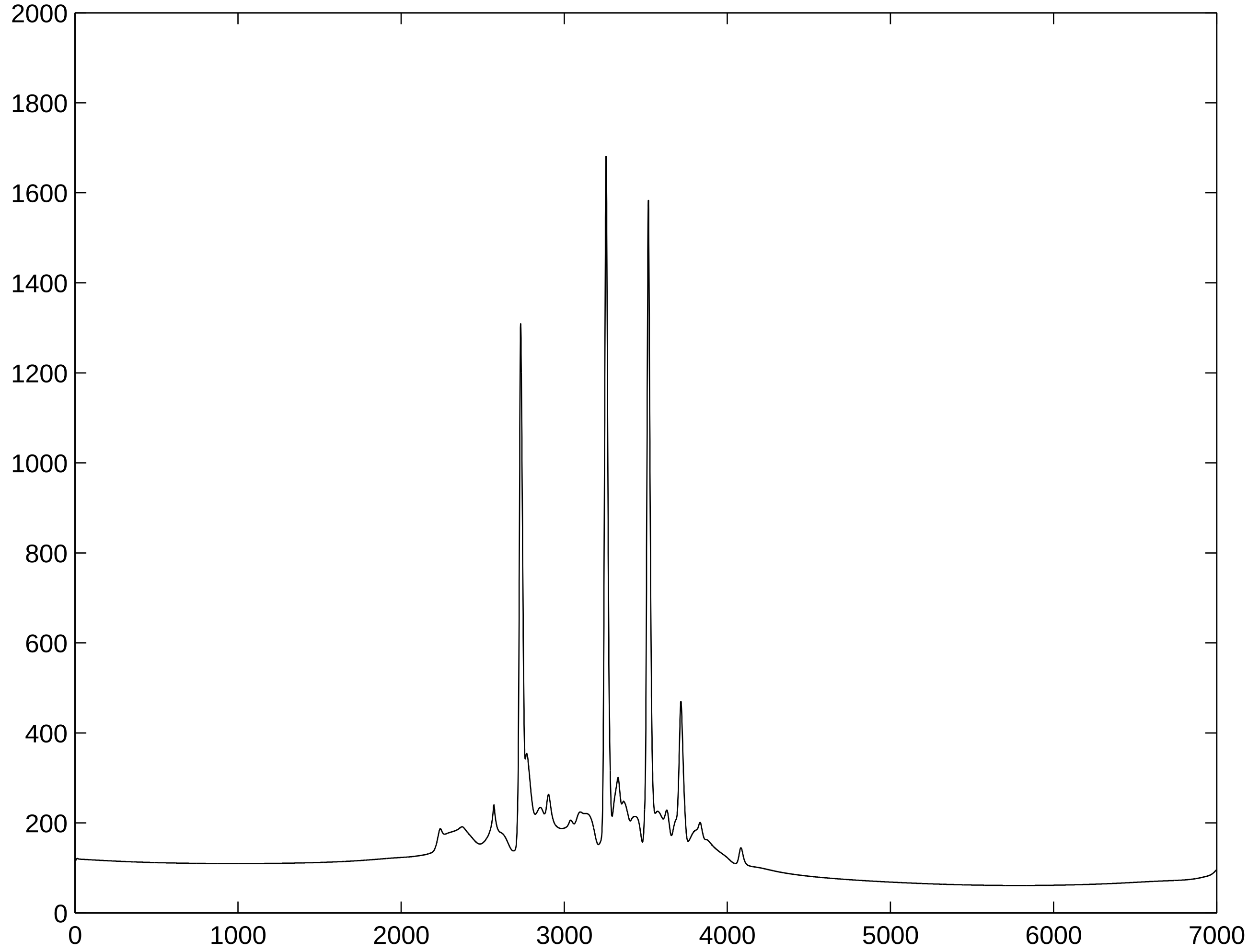
⇝Next



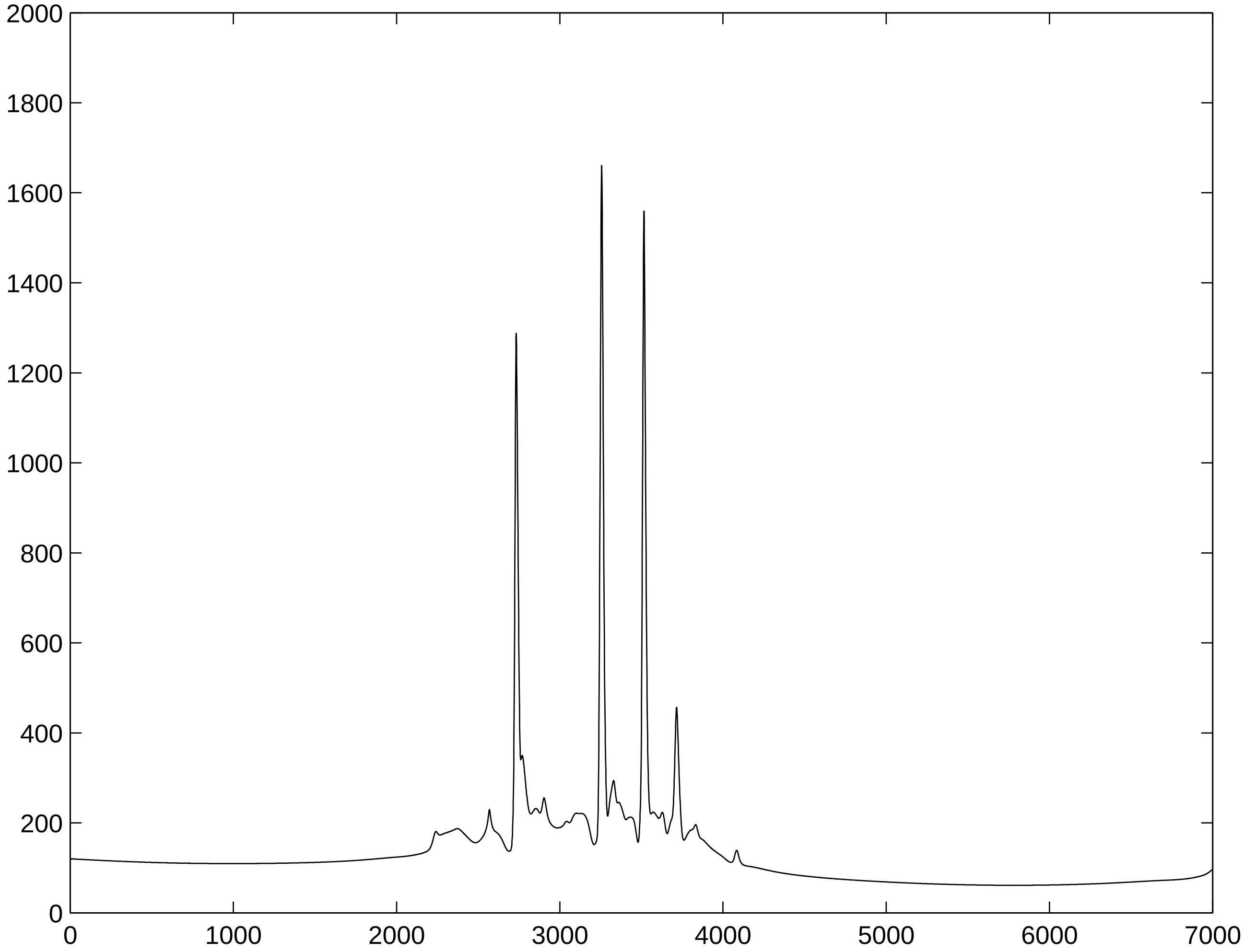
$\rightsquigarrow$Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

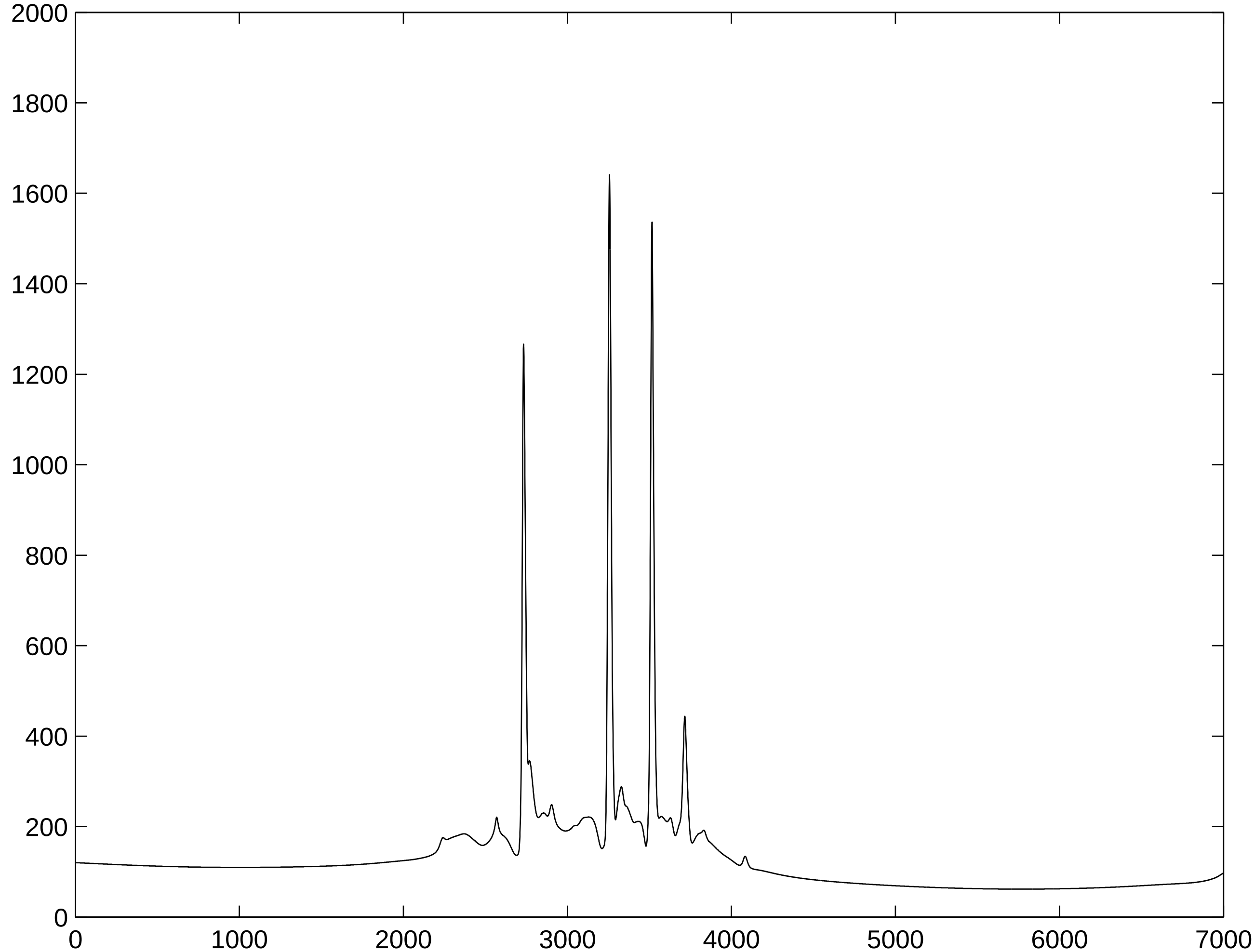
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

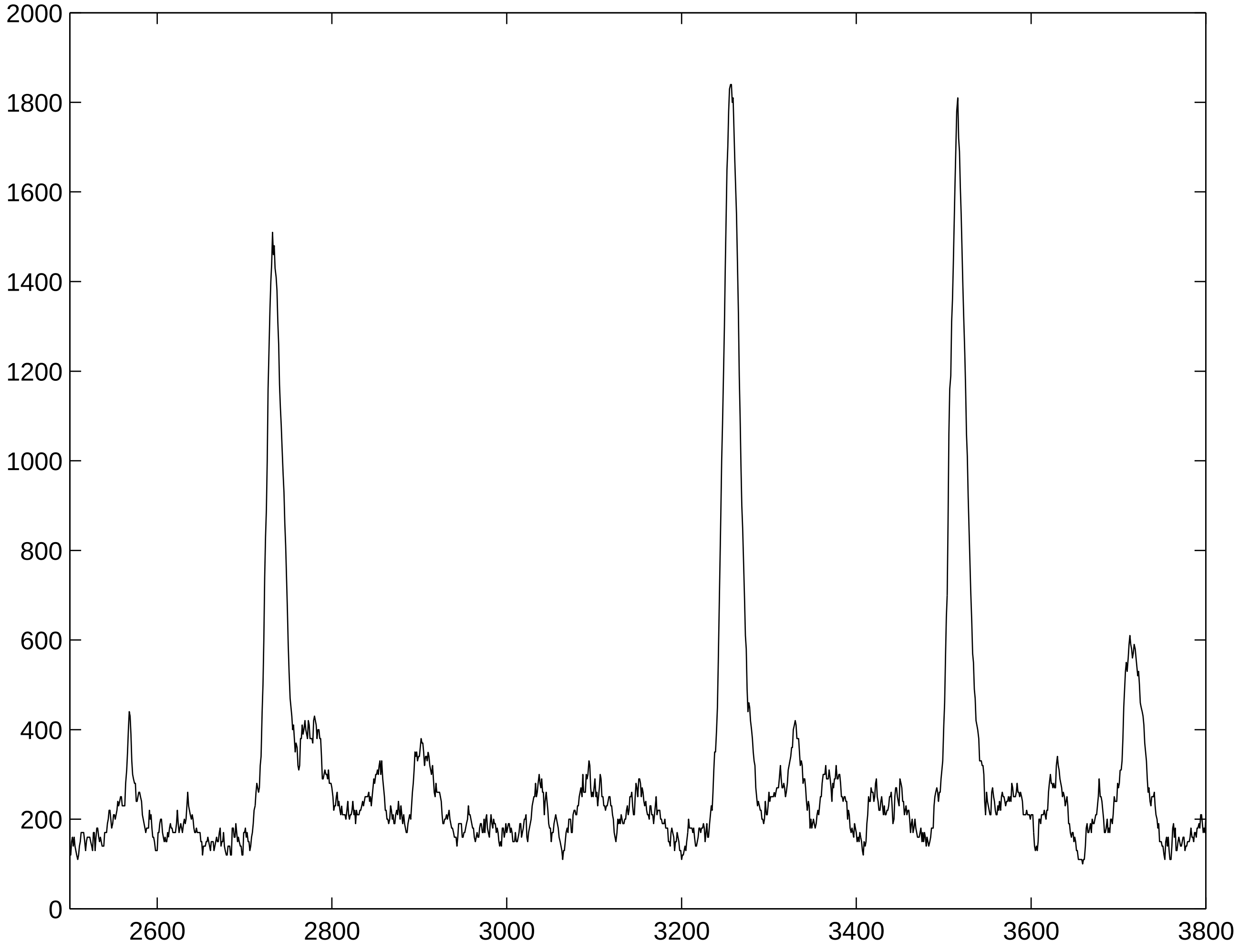
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
0
1000
2000
3000
4000
5000
6000
7000

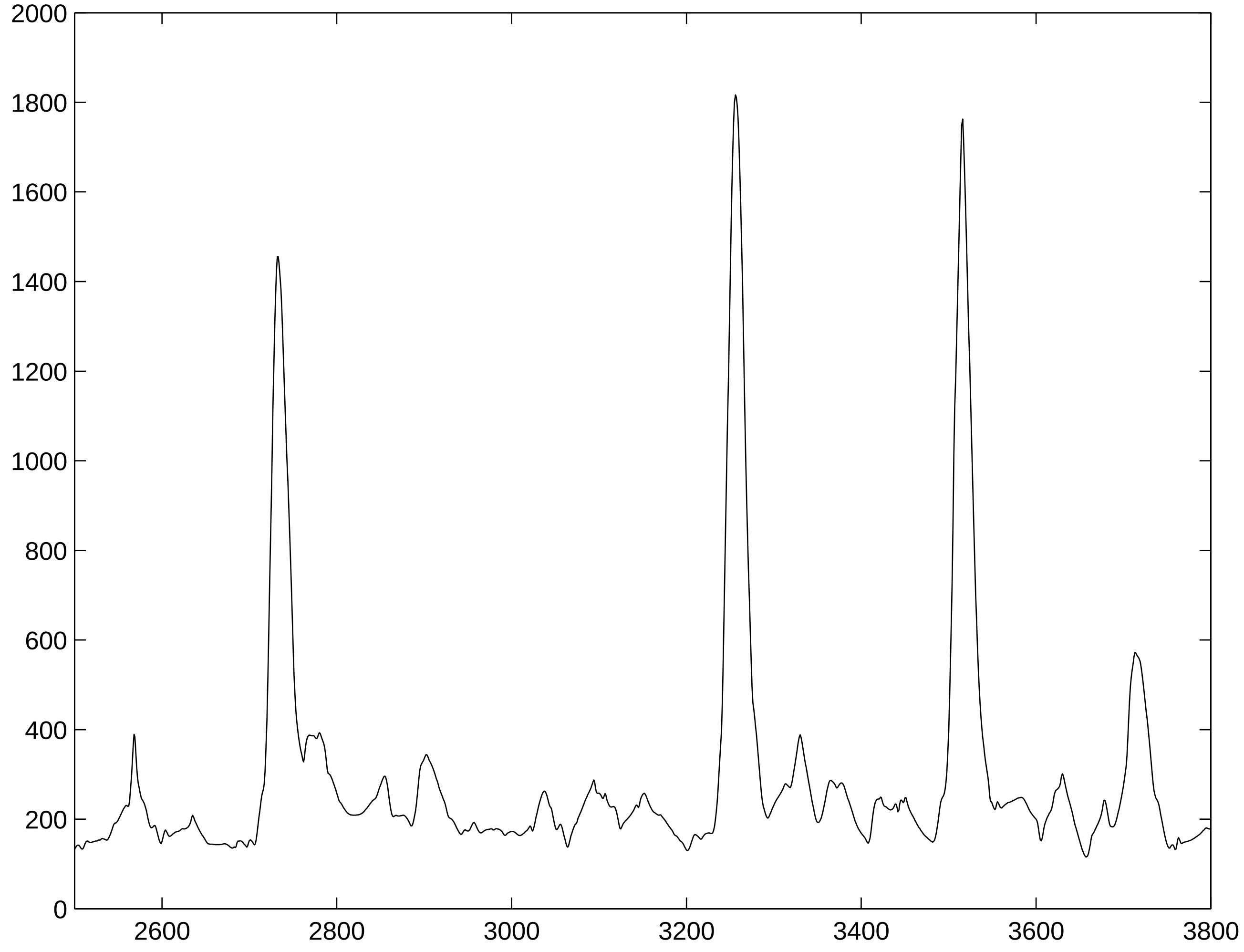
⇝Next



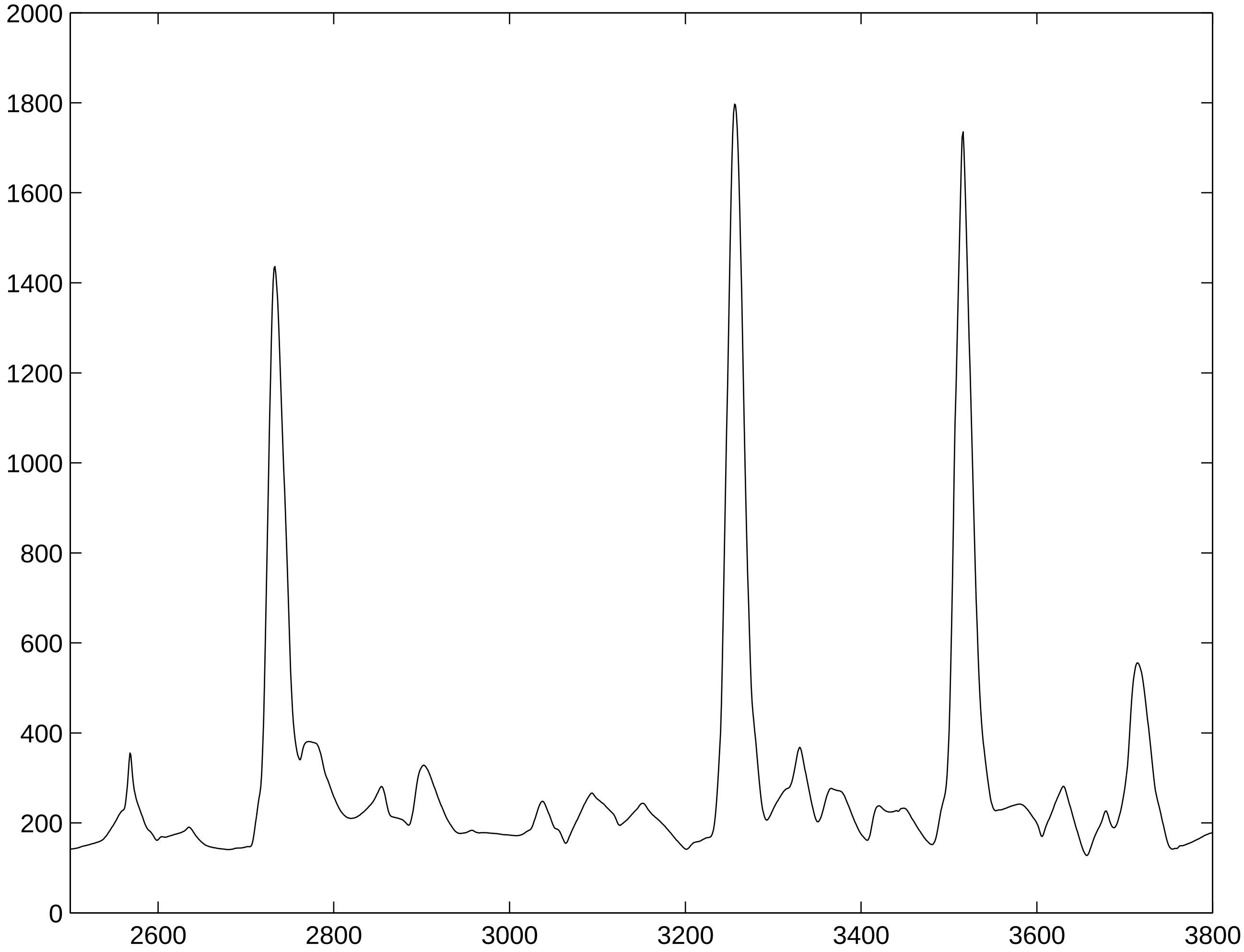
⇝Back
⇝Zoom on the denoising


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

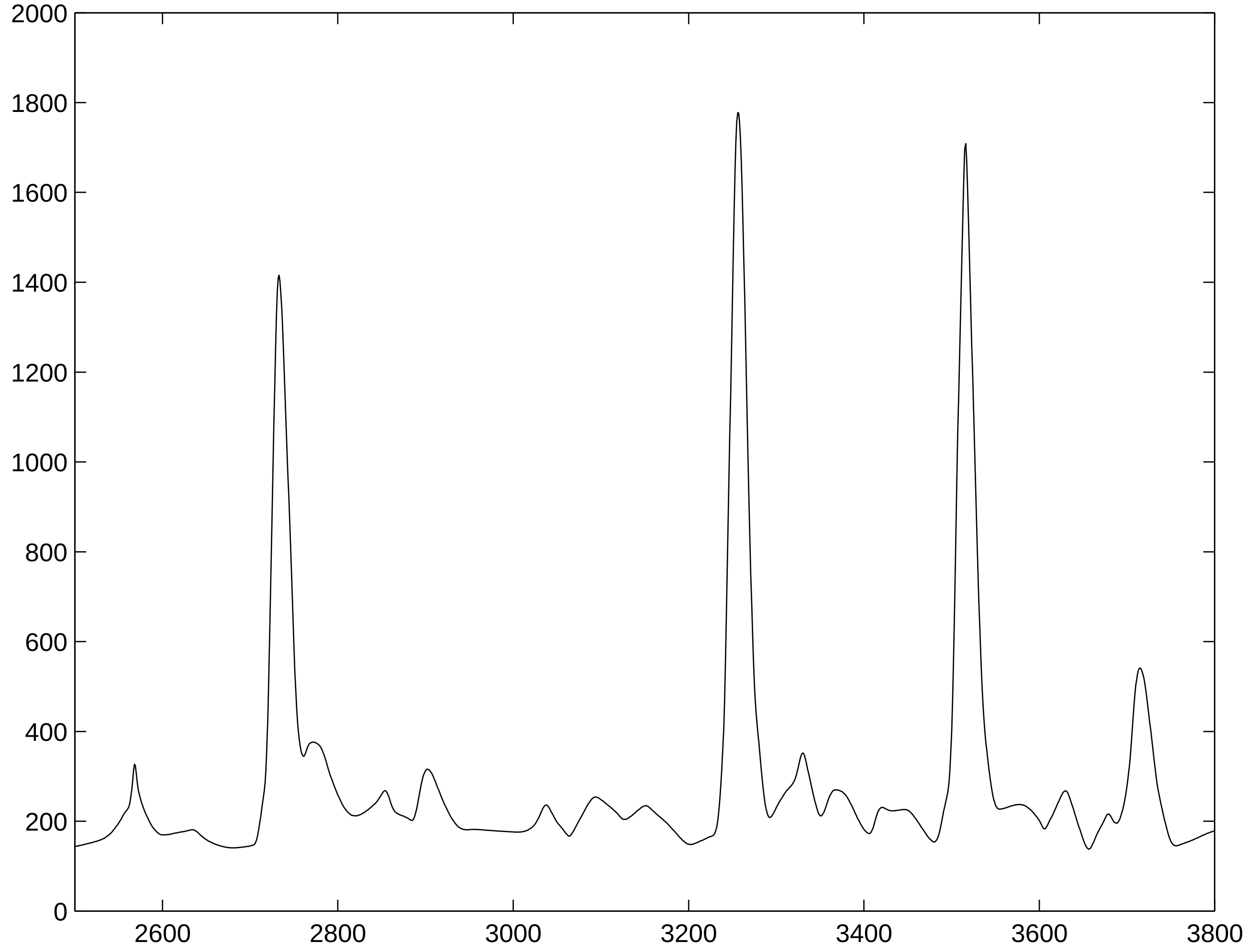
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

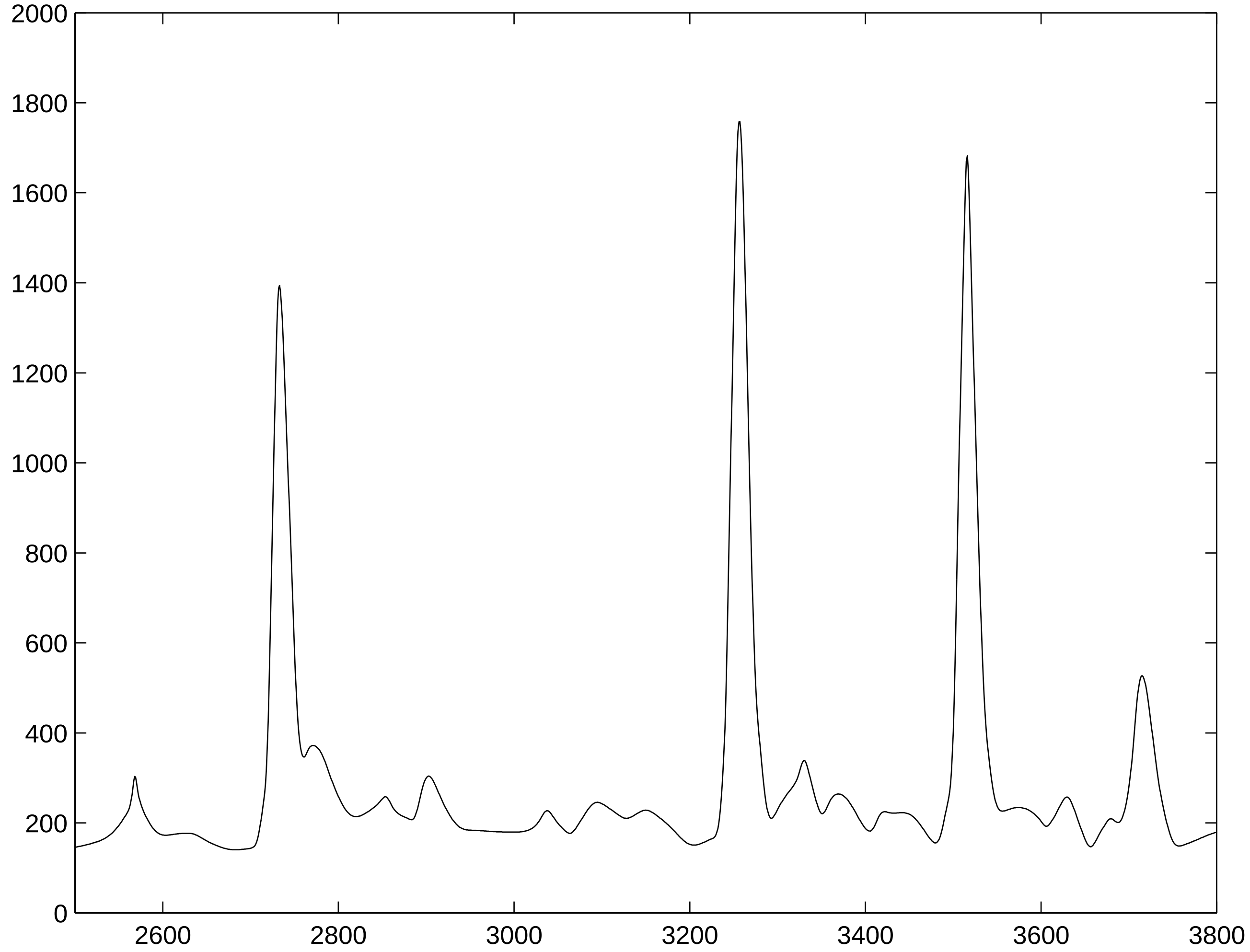
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

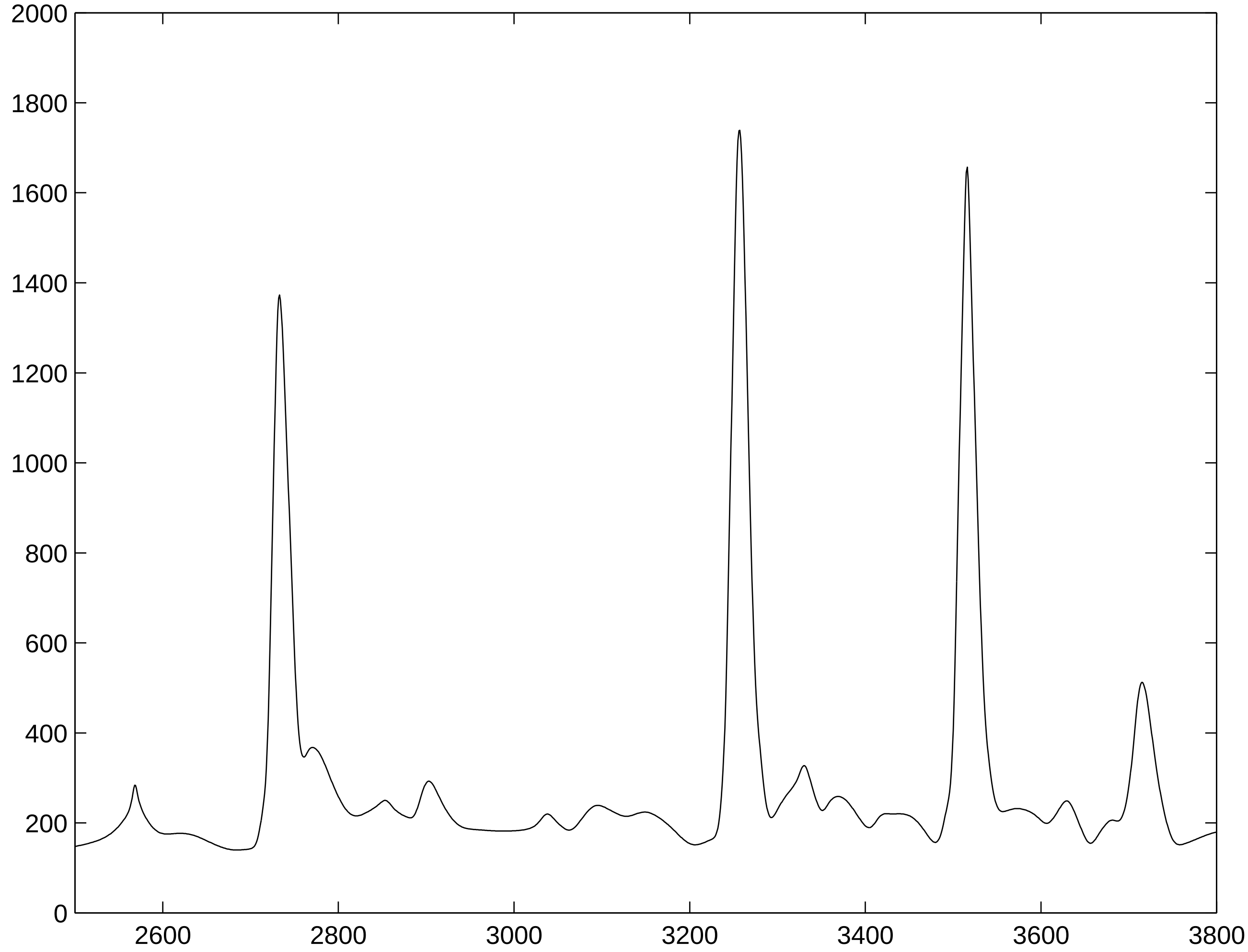
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

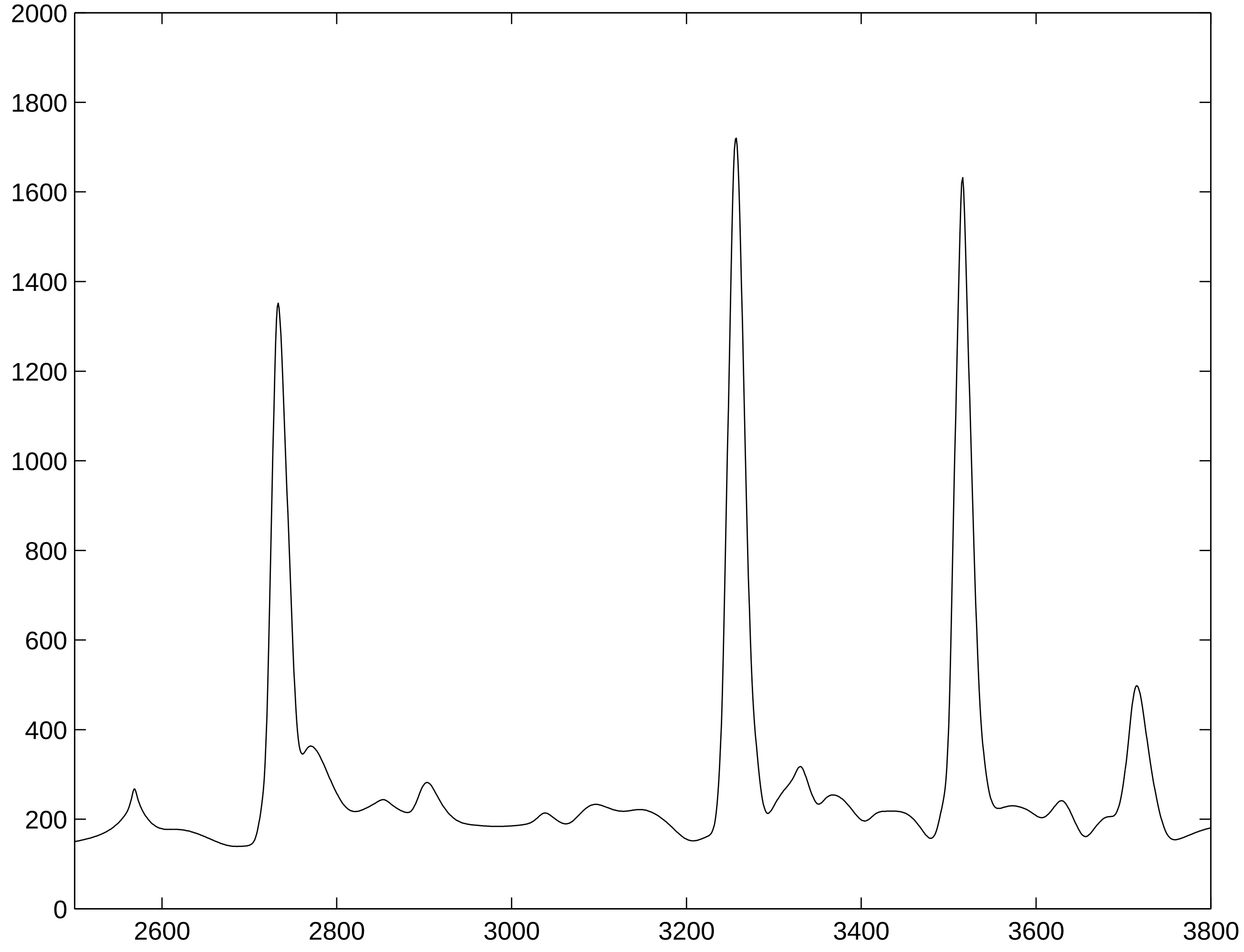
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

⇝Next

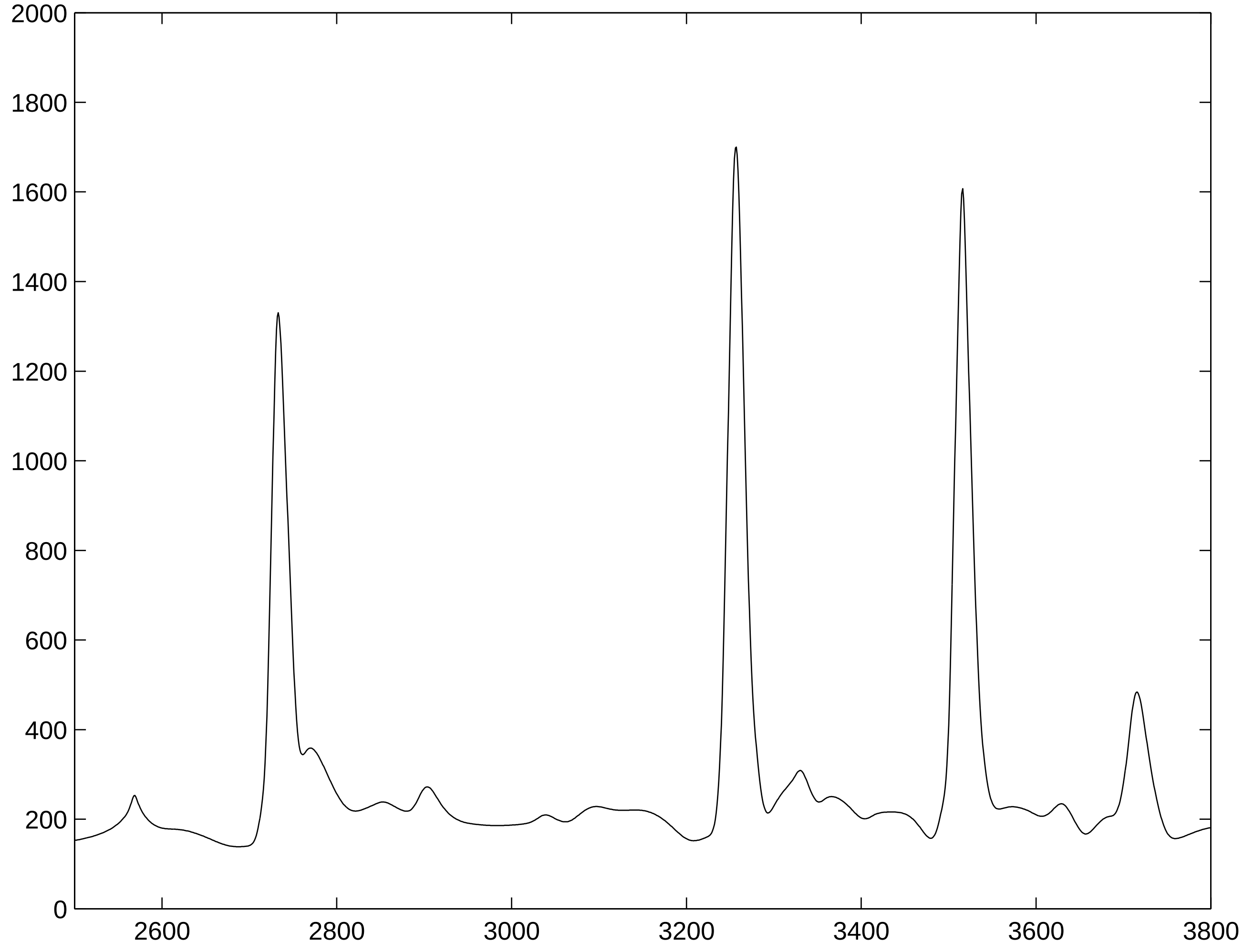

2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

⇝Next

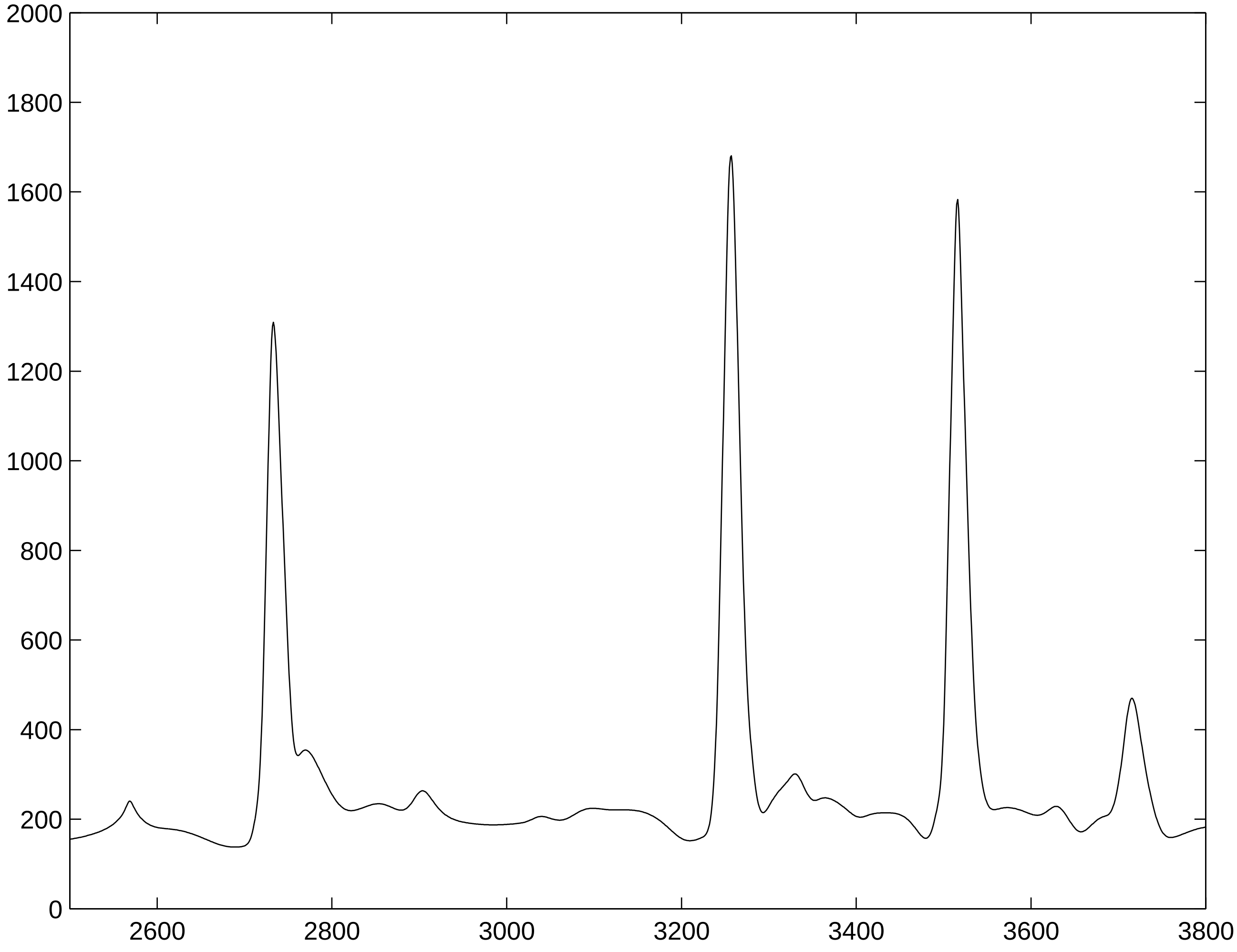


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

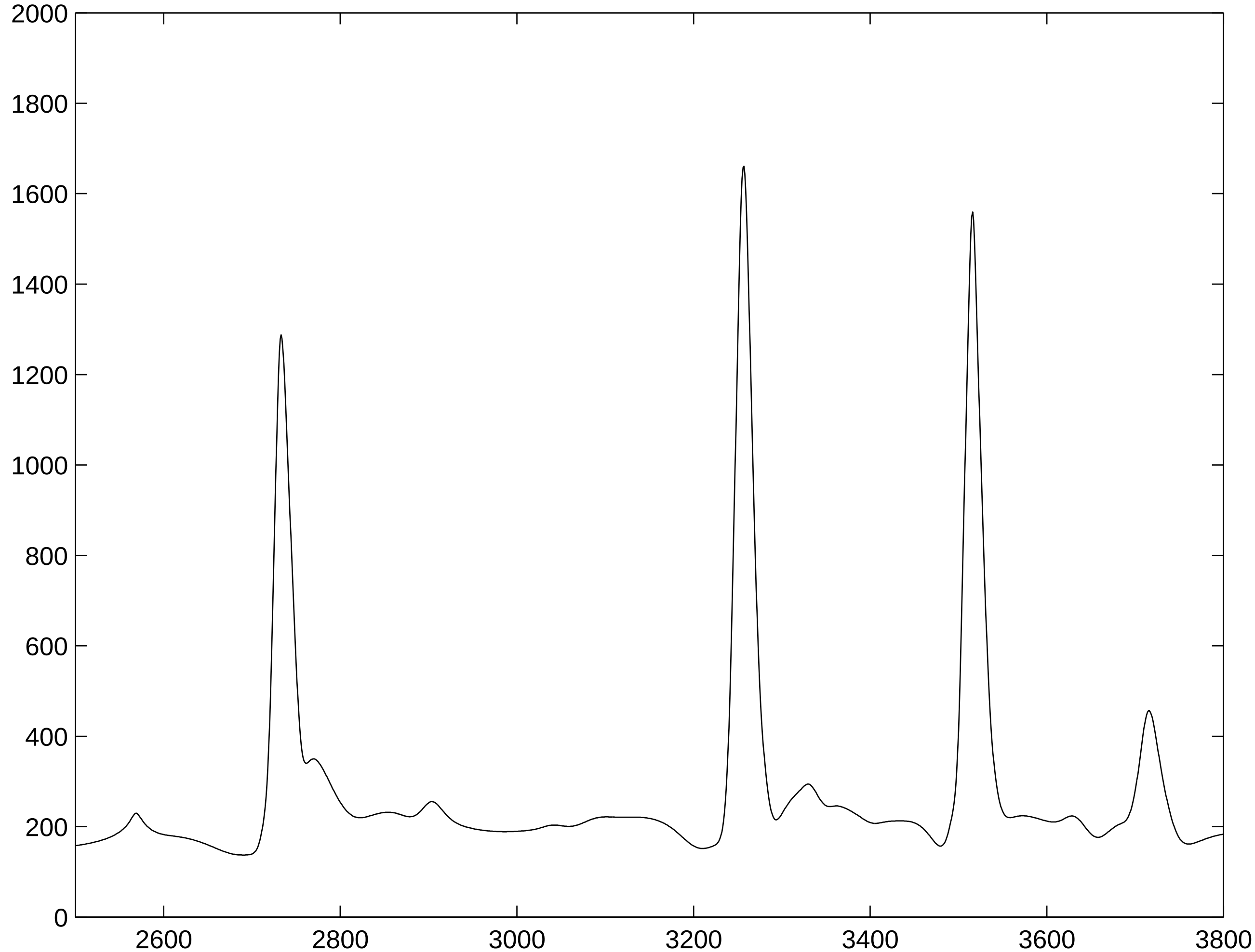
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

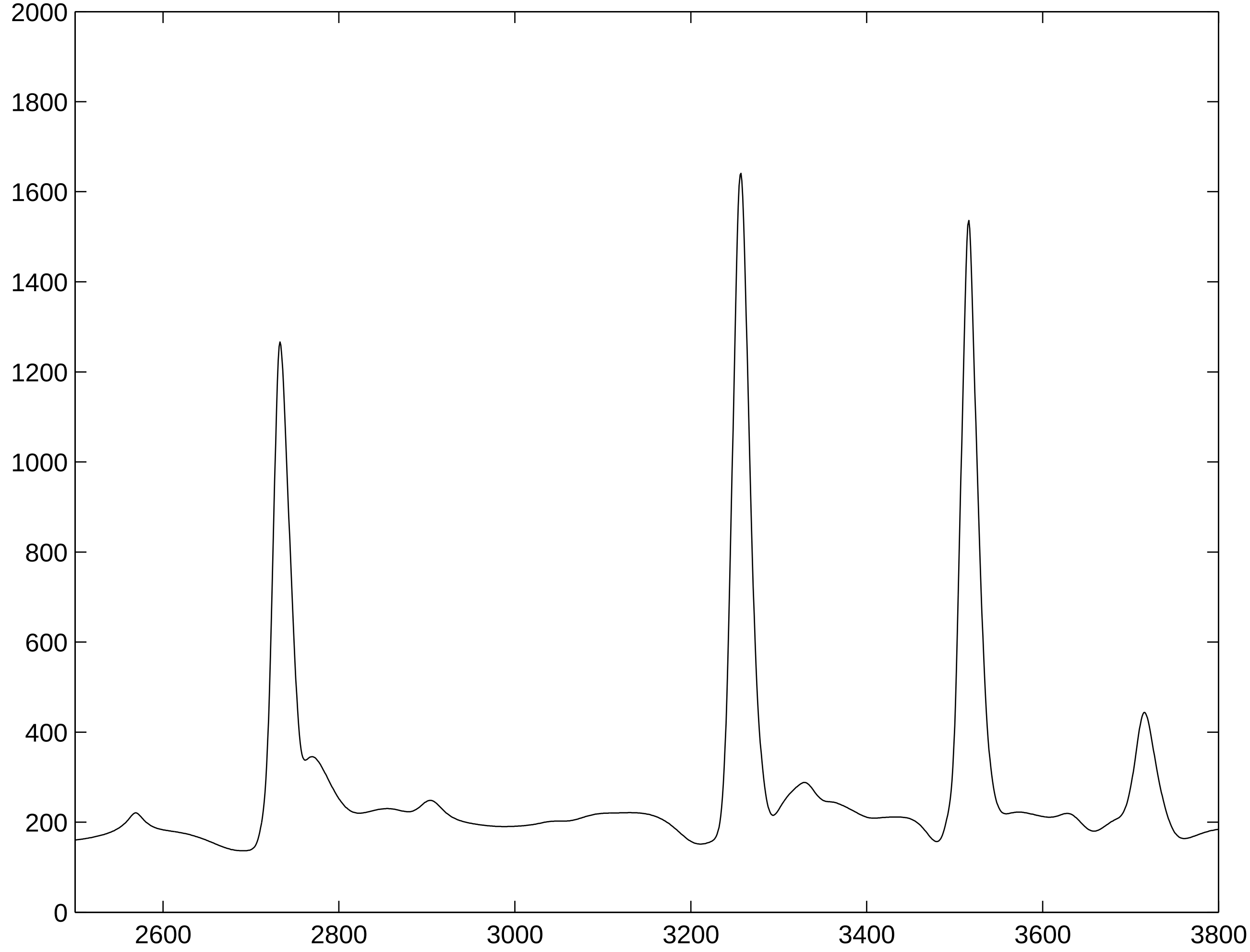
⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

⇝Next


2000
1800
1600
1400
1200
1000
800
600
400
200
0
2600
2800
3000
3200
3400
3600
3800

⇝Back