PUBLICATIONS OF THE ASTRONOMICAL SOCIETY OF THE PACIFIC
98:609–617, June 1986

AN OPTIMAL EXTRACTION ALGORITHM FOR CCD SPECTROSCOPY

KEITH HORNE

Space Telescope Science Institute, Homewood Campus, Baltimore, Maryland 21218

Received 1985 October 25, revised 1986 March 5

ABSTRACT

An optimal spectrum extraction procedure is described and examples of its performance with CCD data are presented. The algorithm delivers the maximum possible signal-to-noise ratio while preserving spectrophotometric accuracy. The effects of moderate geometric distortion and of cosmic-ray hits on the spectrum are automatically accounted for. In tests with background-noise limited CCD spectra, optimal extraction offers a 70% gain in effective exposure time in comparison with conventional extraction procedures.

Key words: data-handling techniques–spectroscopy

I. Introduction

Spectroscopists have witnessed greatly expanding capabilities in recent years with the development of CCD detectors and techniques for processing digital data. With the two-dimensional format, high quantum efficiency, linearity, and large dynamic range of the CCD, large sky backgrounds can be accurately subtracted from spectra of faint objects that were beyond the reach of previous detectors. The algorithms that are used to reduce and analyze CCD data have a direct impact on the quality of the resulting spectra. The development of optimal algorithms is necessary to realize the full potential of CCD spectroscopy.

This paper examines an optimal algorithm for extracting one-dimensional object spectra from a two-dimensional spectrum image. The standard spectrum extraction procedure is to sum the sky-subtracted image data over a range of pixels enclosing the object spectrum in the spatial dimension. Spectrophotometric accuracy can be achieved by using a sufficiently wide band of pixels, but information is then wasted because noisy pixels containing only a small fraction of the light must be included in the sum.

The optimal extraction algorithm resolves this dilemma by applying nonuniform pixel weights in the extraction sum in order to reduce statistical noise in the extracted spectrum to a minimum while simultaneously preserving its photometric accuracy. The algorithm can also eliminate cosmic ray hits on the object spectrum by detecting the distortion they produce in the spatial profile (strong emission lines are not mistaken for cosmic-ray events). The procedure depends on an assumption that the spatial profile varies slowly with wavelength. It therefore is ideal for unresolved point sources, but cannot be recommended for resolved sources whose continua and emission-line regions have different spatial distributions.

Two areas that benefit most significantly from the use of the optimal extraction procedure are faint-object spectroscopy and time-resolved spectroscopy of rapid variables. In these background-noise limited cases, the typical gain over conventional extraction procedures is a 70% increase in effective exposure time.

II. The Spectrum Extraction Algorithm

The formulae that comprise the optimal spectrum extraction algorithm are collected in Table I. The reader may wish to refer to this overview occasionally as we discuss the standard extraction procedure (Steps 1–4) and

TABLE I

Optimal Spectrum Extraction Algorithm

1	initial image processing	$\mathbf{D} = (\mathbf{C} - \mathbf{B})/\mathbf{F}$
2	initial variance estimates	$\mathbf{V} = V_0 + \lvert\mathbf{D}\rvert/Q$
3	fit sky background	$\mathbf{S} = \mathrm{FIT}_x[\mathbf{D}; \mathbf{V}]$
4	extract standard spectrum	$f = \sum_x (\mathbf{D} - \mathbf{S})$
	variance of standard spectrum	$\mathrm{var}[f] = \sum_x \mathbf{V}$
5	construct spatial profile	$\mathbf{P} = \mathrm{FIT}_\lambda[(\mathbf{D} - \mathbf{S})/f; \mathbf{V}/f^2]$
	enforce positivity	$\mathbf{P} = \mathrm{MAX}[\mathbf{P}; 0]$
	enforce normalization	$\mathbf{P} = \mathbf{P}/\sum_x \mathbf{P}$
6	revise variance estimates	$\mathbf{V} = V_0 + \lvert f\mathbf{P} + \mathbf{S}\rvert/Q$
7	mask cosmic ray hits	$\mathbf{M} = \begin{cases} 0 & \text{if } (\mathbf{D} - \mathbf{S} - f\mathbf{P})^2 > \sigma_{clip}^2 \mathbf{V} \\ 1 & \text{otherwise} \end{cases}$
8	extract optimal spectrum	$f = \dfrac{\sum_x \mathbf{MP}(\mathbf{D} - \mathbf{S})/\mathbf{V}}{\sum_x \mathbf{MP}^2/\mathbf{V}}$
	variance of optimal spectrum	$\mathrm{var}[f] = \dfrac{\sum_x \mathbf{MP}}{\sum_x \mathbf{MP}^2/\mathbf{V}}$
9	iterate Steps 5 thru 8	

its extension (Steps 5–9) to an optimal procedure. Section III presents a few examples to compare results obtained with the standard and optimal spectrum extraction procedures. Parts of the discussion address issues specific to CCD data, and a CCD provides the data used in the examples, but the principles employed in optimizing the extraction algorithm, and the optimal extraction algorithm itself, are applicable to any two-dimensional spectroscopic data.

A. *Image Processing and Standard Spectrum Extraction*

We begin with a brief description of the image processing that normally precedes the extraction of object spectra, since these steps too have an impact on the quality of the final spectrum. The raw spectrum image $C_{x\lambda}$ is filled with data numbers that are proportional to the number of photons detected during the exposure by each CCD pixel. Here the discrete indices x and λ refer to pixel coordinates in the spatial and spectral directions of the CCD. The true wavelengths and spatial positions are, of course, only approximately constant along pixel rows and columns, due to imperfect spectrograph optics and differential refraction in the Earth's atmosphere. Such distortions may be removed by resampling the image, but this practice consumes computer time, degrades the resolution of the image, and introduces correlations between the errors in adjacent image pixels. Resampling of CCD spectra is not usually necessary, however, since the optimum extraction procedure copes automatically with small distortions of up to several pixels across the image.

The initial image processing (Step 1) involves the pixel-by-pixel subtraction of a bias image B and division by a balance factor image F to produce the reduced image

$$D_{x\lambda} = \frac{C_{x\lambda} - B_{x\lambda}}{F_{x\lambda}} \quad . \tag{1}$$

These simple corrections make no attempt to fix wild pixel values caused by cosmic-ray hits, "hot" pixels, deferred charge, and other defects known to affect CCD data. The philosophy underlying our cavalier treatment of such blemishes is that bad data should be rejected, not repaired.

The purpose of subtracting a bias image is to eliminate the signal that is registered by the detector in the absence of exposure to light. For CCD data this signal includes a fixed pattern, which may sometimes be assumed to be constant along one dimension of the CCD, and an additive bias offset that varies from one exposure to the next but is uniform across the image. The required bias image B is constructed accordingly, by averaging a series of zero-exposure images, possibly averaging along one dimension of the CCD, and then adding a constant to eliminate the bias offset. The constant is found by comparing the average value of $(C - B)$ in a region of the image that is not exposed to light. The statistical noise introduced by bias subtraction can be reduced to a negligible level if the bias image is formed by averaging a sufficiently large number of zero-exposure images. If many exposures are to be combined to form the final spectrum of any object, the required number of zero-exposure images is correspondingly increased. Wild pixel values produced by cosmic ray hits must also be eliminated from the bias image. This may be accomplished conveniently by removing the bias offsets of the individual zero-exposure images and then computing a pixel-by-pixel median of the series.

Division by the balance factor image F serves both to compensate for pixel-to-pixel variations in detector sensitivity, and to remove long-scale response variations in the spatial direction, such as might be caused by vignetting of the spectrograph slit or nonparallel slit jaws. The balance factor image F is constructed from flat-field images obtained by exposing the CCD to a spatially-uniform continuum source such as the interior of the telescope dome under illumination by an incandescent lamp. The flat-field images are individually de-biased, following the procedure described above, and added together. The resulting image is then divided by a low-order polynomial in λ fitted to flat-field data summed in the spatial direction. This last step produces an image of balance factors with values of order unity. Division by the balance factors therefore approximately preserves the relationship between data numbers and detected photons, allowing direct calculation of Poisson variance estimates without reference to the flat-field frames. Since a large number of well-exposed flat-field images are combined to form the balance factors, statistical errors introduced by the balance factor division can be ignored.

The next operation is sky subtraction. The sky background image S is derived from the reduced image D by smoothly interpolating sky data on each side of the object spectrum. This operation is denoted by FIT_x in Step 3 of Table I. A simple average or median value may suffice to represent the spatial profile of the sky at many wavelengths, but quadratic or cubic polynomials in x are required near night sky lines that are appreciably curved or tilted with respect to the grid of image pixels. We therefore generally perform a weighted least-squares polynomial fit to the sky data at each λ. Individual sky pixels are given weights inversely proportional to their variances estimated as in Step 2 of Table I. (Section II.C further discusses the estimation of pixel variances.) A software mask is used to cover regions of the image that are contaminated by object spectra, and likewise any columns that are seriously affected by inefficient charge transfer. Cosmic ray hits and other blemishes are rejected from the sky fit by an iterative σ–clipping procedure that is described more fully in Section II.D. The interpolated sky background is assumed to be effectively noise-free, since generally a much larger number of pixels is devoted to the

sky than to the object spectrum.

With the image processing steps now complete, the object spectrum is extracted from the sky-subtracted image ($D-S$) by summing along the spatial dimension

$$f_\lambda^{\text{std}} = \sum_{x=x_1}^{x_2} (D_{x\lambda} - S_{x\lambda}) \ . \tag{2}$$

The variance of this standard spectrum estimator is

$$\text{var}\,[f_\lambda^{\text{std}}] = \sum_{x=x_1}^{x_2} V_{x\lambda} \ , \tag{3}$$

where $V_{x\lambda}$ are the statistical variances of the data values $D_{x\lambda}$.

We assume here that reliable estimates for the individual pixel variances $V_{x\lambda}$ are available; a discussion of their determination is provided in Section II.C. Equation (3) neglects statistical errors introduced by bias subtraction, balance factor division, and subtraction of the interpolated sky, which for reasons already noted contribute much less to the final spectrum variance than do the variances of the individual pixels in the extraction sum.

The 'object limits' x_1 and x_2 delimit the spatial range of pixels needed to enclose the object spectrum. If the spatial profile is Gaussian, object limits separated by three times the full width at half maximum enclose over 99% of the starlight. If narrow object limits are used, the photometric accuracy of the spectrum is sacrificed since an appreciable fraction of the star's light is not counted. With wide object limits the extracted spectrum becomes unnecessarily noisy as numerous pixels that contain very little starlight are included in the extraction sum. An optimum choice of object limits can be made by a compromise between the opposing requirements of low statistical noise and high photometric accuracy. In practice, however, object limits are often selected interactively on the basis of a grey-scale display of the spectrum image data, or a plot of its spatial profile summed over some range of pixels in the wavelength direction. This subjective procedure can introduce systematic errors if for example wider limits are selected for better-exposed spectra on which the wings of the spatial profile are more easily visible.

Because lower weights are assigned to the noisy pixels far from the center of the spatial profile, the optimal extraction algorithm produces a spectrum estimate that is not as sensitive to the exact choice of object limits. When the optimal extraction procedure is used, widely-spaced object limits should always be chosen.

B. *Optimization*

At this point, we expand the scope of our discussion to include the broad class of *linear* spectrum estimators, those of the form

$$d_\lambda^{\text{linear}} = \sum_x w_{x\lambda} (D_{x\lambda} - S_{x\lambda}) \ . \tag{4}$$

Thus any linear combination of the image data is considered a potential estimator of the object spectrum. The standard spectrum estimator discussed in Section II.A is one member of this class, with the particular extraction weights $w_{x\lambda}$ equal to unity between the extraction limits x_1 and x_2 and zero otherwise. The aim of this section is to discover how the extraction weights $w_{x\lambda}$ may be chosen in an optimal way. Our discussion generalizes that given by Robertson (1983).

Optimization of the extraction weights depends upon accurate information about the spatial distribution of the detected starlight at each wavelength. Let $P_{x\lambda}$ be the probability that a detected photon with wavelength λ is registered in pixel x rather than at some other position along the spectrograph slit. The image P is strictly nonnegative, and is normalized at each λ to a unit sum over x

$$\sum_x P_{x\lambda} = 1 \ . \tag{5}$$

We assume for the moment that this normalized spatial profile image is accurately known. An example is illustrated in Figure 1, and the procedure used in practice to generate P from the spectrum image is discussed in Section II.D.

A knowledge of the spatial distribution of the starlight allows us to identify and limit further attention to extraction weights that give unbiased estimates of the object spectrum. In the absence of noise, f_λ would be given exactly by $(D_{x\lambda} - S_{x\lambda})/P_{x\lambda}$. This result follows directly from the definition of $P_{x\lambda}$ as the fraction of the starlight at a given wavelength that falls on an individual pixel. When noise is present, the quantities $(D_{x\lambda} - S_{x\lambda})/P_{x\lambda}$ provide independent and unbiased estimates of the object spectrum for every pixel in which $P_{x\lambda}$ is positive. Thus any spectrum estimator that is both *linear* and *unbiased* is a linear combination of these unbiased individual pixel estimates and may be expressed in the general form

$$f_\lambda^{\text{unbiased}} = \frac{\Sigma_x W_{x\lambda}(D_{x\lambda} - S_{x\lambda})/P_{x\lambda}}{\Sigma_x W_{x\lambda}} \ . \tag{6}$$

We now determine how to choose the weights $W_{x\lambda}$ in order to minimize the variance of $f_\lambda^{\text{unbiased}}$. For each λ, $f_\lambda^{\text{unbiased}}$ is a simple weighted average of the independent random variables $(D_{x\lambda} - S_{x\lambda})/P_{x\lambda}$, which have identical mean values but different variances. The variance of the weighted mean $f_\lambda^{\text{unbiased}}$ can therefore be minimized by choosing weights that are inversely proportional to the variance of the individual random variables

$$\frac{1}{W_{x\lambda}} = \text{var}\left[\frac{(D_{x\lambda} - S_{x\lambda})}{P_{x\lambda}}\right] = \frac{V_{x\lambda}}{P_{x\lambda}^2} \ . \tag{7}$$

The optimal (minimum-variance unbiased) spectrum esti-

mate therefore is

$$f_\lambda^{\rm opt} = \frac{\Sigma_x P_{x\lambda}(D_{x\lambda} - S_{x\lambda})/V_{x\lambda}}{\Sigma_x P_{x\lambda}^2/V_{x\lambda}} \ , \tag{8}$$

and the variance of this optimal spectrum estimate is

$$\mathrm{var}[f_\lambda^{\rm opt}] = \frac{1}{\Sigma_x P_{x\lambda}^2/V_{x\lambda}} \ . \tag{9}$$

These formulae are equivalent to determining the spectrum $f_\lambda^{\rm opt}$ by scaling a known spatial profile $P_{x\lambda}$ to fit the sky-subtracted data $(D_{x\lambda} - S_{x\lambda})$ whose variances are $V_{x\lambda}$.

Optimal extraction achieves the highest possible signal-to-noise ratio by assigning lower weights to pixels far from the peak of the spatial profile that receive little light from the star. The expected reduction in variance may be evaluated quantitatively by comparing equations (3) and (9). We consider the specific case of a well-sampled Gaussian spatial profile with a standard deviation of σ pixels, and assume that the noise distribution is independent of x, as in background-limited spectroscopy. The variance of the standard spectrum estimate is then larger than that of the optimal spectrum estimate by the factor

$$\frac{\mathrm{var}[f^{\rm std}]}{\mathrm{var}[f^{\rm opt}]} = (x_2 - x_1)\int_{x_1}^{x_2} \frac{dx}{2\pi\sigma^2}\, e^{-x^2/\sigma^2} \tag{10}$$

$$\approx \frac{x_2 - x_1}{2\sqrt{\pi}\sigma}$$

for wide object limits. If the object limits span the range $\pm 3\sigma$, optimal extraction increases the effective exposure time by a factor of $3/\sqrt{\pi} \doteq 1.69$.

In the opposite limit of negligible background noise, the pixel variances approach the Poisson limit

$$V_{x\lambda} \propto P_{x\lambda} \ , \tag{11}$$

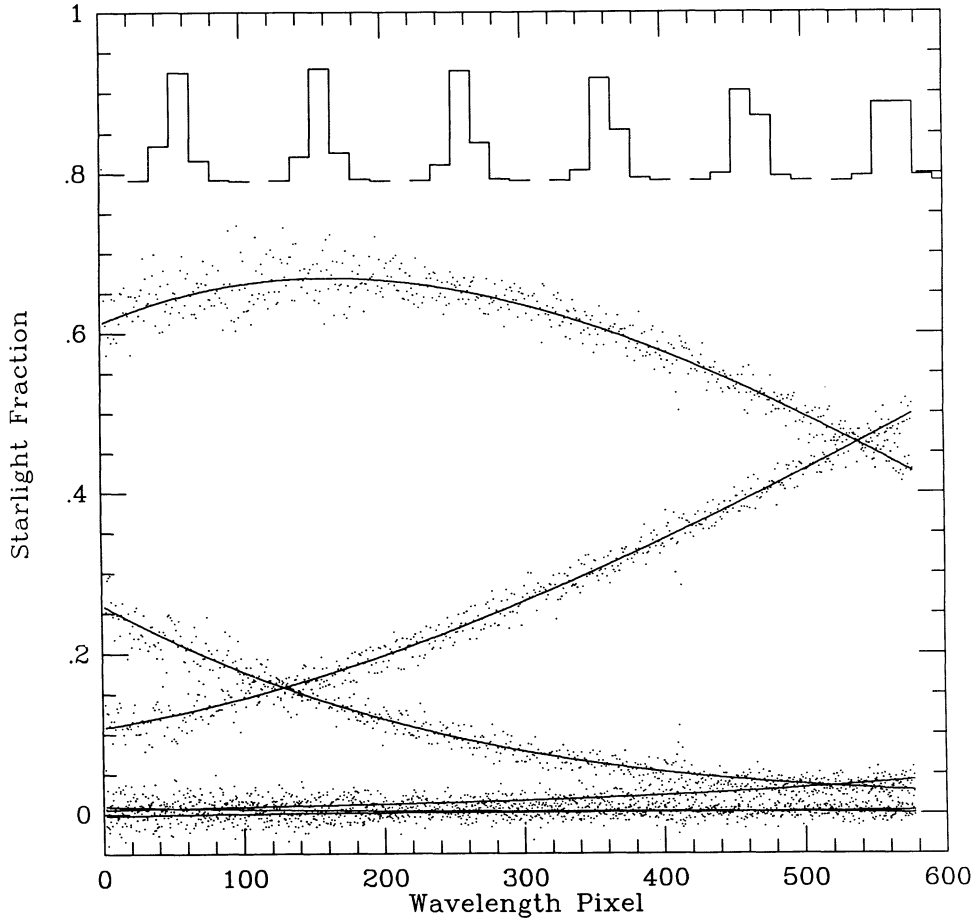

FIG. 1–The spatial profile of starlight detected by a CCD spectrograph is shown at the top for six different wavelengths, and the fraction of the starlight collected by individual CCD pixels is shown below. The smooth curves are weighted polynomial fits to the empirical estimates based on individual pixels. A conventional spectrum extraction would in this case require a sum over four pixels at each wavelength. The resulting spectrum is noisier than it needs to be because the two outer pixels, which contain only 10%–15% of the starlight, nearly double the statistical variance. A less-noisy spectrum is produced by the optimal extraction procedure, which applies smaller weights to pixels that contribute little light from the star. The polynomial model of the spatial distribution of the starlight is used to determine optimal pixel weights that minimize statistical noise while preserving the spectrophotometric accuracy of the spectrum.

and the optimal and direct spectrum estimates become formally identical.

C. *The Variance Image*

An important element of the optimal extraction procedure is the image V containing estimates of the variance of the spectrum image D. These pixel variances are needed for the weighted polynomial fits that are used to represent the sky and spatial profile images, as well as for estimates of the uncertainty in the final spectrum.

The pixel variances are initially taken to be

$$V_{x\lambda} = V_0 + |D_{x\lambda}|/Q \ , \tag{12}$$

where $\sqrt{V_0}$ is the root-mean-squared readout noise (in data numbers) and Q is the effective number of photons per data number. The parameters V_0 and Q, which characterize the noise properties of the CCD, are determined empirically by examining fluctuations as a function of data number in one or more reduced (de-biased and balanced) zero-exposure and flat-field images. I assume that constant values of V_0 and Q apply over the entire CCD frame, but independent values for each CCD pixel may of course be determined if a multitude of zero-exposure and flat-field frames is available. The absolute value of $D_{x\lambda}$ taken in equation (12) is important to prevent pixel variance estimates from falling below the level of readout noise as a result of statistical fluctuations.

After a spectrum estimate f, a normalized spatial profile image P, and a sky image S are available, the initial pixel variances are replaced by better estimates:

$$V_{x\lambda} = V_0 + |f_\lambda P_{x\lambda} + S_{x\lambda}|/Q \ . \tag{13}$$

This substitution of predicted data $fP + S$ for the observed data D guards against the application of excessive weight to pixel values that are low as a result of statistical fluctuations. If this precaution is not taken, the extracted spectra develop a skewed error distribution with a bias toward low values. The result can be disastrous for spectra with low signal-to-noise ratios.

D. *The Spatial Profile Image*

The image P of normalized spatial profiles is the most crucial component of the optimal extraction procedure. The spatial profiles must be accurately determined if spectrophotometric accuracy and the highest possible signal-to-noise ratio are to be achieved. The extracted spectrum will be effectively unbiased only if the relative error in $P_{x\lambda}$ is much smaller than the relative statistical uncertainty in $(D_{x\lambda} - S_{x\lambda})$.

In some applications, for example a stationary spectrograph fed by optical fibers mounted at fixed positions along the spectrograph slit, the spatial profiles may be assumed to be identical for all objects observed. In this case a suitable profile image may be obtained by summing all of the two-dimensional sky-subtracted spectrum images, setting any negative values to zero, and renormalizing at each λ to ensure a unit sum over x. We consider below the more difficult case; for example, a long-slit Cassegrain spectrograph, in which the spatial profiles are unique to each observation and so must be derived from the spectrum image itself.

The first step in the construction of P is to form an initial estimate by dividing the sky-subtracted spectrum image by its sum over x

$$P_{x\lambda} \approx \frac{D_{x\lambda} - S_{x\lambda}}{\Sigma_x (D_{x\lambda} - S_{x\lambda})} = \frac{D_{x\lambda} - S_{x\lambda}}{f_\lambda^{std}} \ . \tag{14}$$

This initial estimate satisfies the desired normalization condition, but is unacceptably noisy and may contain negative values.

These faults can be remedied by smoothing in the wavelength direction, assuming that the spatial profile is a slowly varying function of wavelength. If the spatial profiles were identical at all wavelengths, a simple average over λ would yield a profile image suitable for use in the optimal extraction. In practice, however, we must take into account the slow wavelength dependence that arises from optical and geometric distortion in the spectrograph, detector misalignment, differential refraction in the Earth's atmosphere, and wavelength-dependent seeing.

Under many circumstances the spatial profile may be assumed to have a simple analytic form, such as a Gaussian. The parameters of the chosen model profile may then be fitted to the sky-subtracted data at each wavelength, and continuity with wavelength can be imposed by smoothing the fitted parameter values as functions of wavelength. This approach was taken in several programs for the extraction of spectra from *IUE* data and has led to improved signal-to-noise ratios in comparison with the standard extraction program (de Boer and Snijders 1981, and references therein). There is, however, a concern about the effect of broad wings on the *IUE* point-spread function. Profile-fitting should work best for low-to-moderate signal-to-noise ratio, but may become unsatisfactory at higher signal-to-noise levels as differences between the true form of the spatial profile and the assumed model become increasingly important.

The more general approach taken here makes no assumption about the analytic form of the spatial profile, only that it vary continuously with wavelength. Figure 1 shows that the slow variation down each row of pixels can be well approximated by a low-order polynomial in λ. The fitted polynomials provide a smooth and unbiased model for the spatial profile image P. The order of the polynomials must of course be high enough to permit them to follow the wavelength variations, and should be the same for all objects observed. The empirical points to which the polynomials are fitted should be weighted in inverse proportion to their variances, as is indicated in Step 5 of

Table I.

Wild pixel values produced by cosmic ray hits near the spectrum must not be allowed to affect the fitted polynomials. The following iterative σ–clipping procedure has been found to be satisfactory. After performing each weighted polynomial fit, all pixels having a squared normalized residual $(D_{x\lambda} - S_{x\lambda} - f_\lambda P_{x\lambda})^2/V_{x\lambda}$ in excess of a fixed threshold $\sigma_{clip}^2 \sim 16$ are identified as outliers. The weighted polynomial fit is then repeated omitting the outliers, and a new set of outliers is identified. This procedure is repeated until the set of outliers is the same on two successive iterations; usually 1 to 3 iterations are required. Because the number of wavelengths is generally much larger than the number of polynomial coefficients, σ_{clip} may be set to a value as small as 2 without adversely affecting the result. Graphs like that of Figure 1 indicate that the σ–clipping procedure works well in practice if the variance estimates $V_{x\lambda}$ are realistic.

The profile image $P_{x\lambda}$ is generated by evaluating the fitted polynomials. Any negative values are set to zero, and the image is renormalized at each λ to ensure that its sum over x is unity.

The polynomial model described here can cope with the moderate spectrum distortions usually encountered in CCD spectroscopy. With other detectors the spatial centroid of the spectrum may shift by more than a few spatial widths across the length of the spectrum, and a polynomial may not be adequate to represent the wavelength variations. In this case a running median filter, suitably modified to treat end effects, might be substituted for the polynomial model. Because the spatial profile at each wavelength combines information from a large number of adjacent wavelengths, the profile is effectively noise-free.

E. *Elimination of Cosmic Rays*

Cosmic-ray hits are one of the most annoying problems facing the CCD spectroscopist. The impact of a cosmic ray deposits a spurious signal in a connected region that depends in detail on the specific trajectory of the cosmic ray through the detector, but typically spans a few pixels in each direction. The magnitude of the disturbance depends on the energy of the cosmic ray, and thus varies over a wide range. Weak events may be difficult to recognize by simple inspection of the CCD image. Hits on the object spectrum bear a resemblance to a narrow emission line.

Access to an accurate spatial profile at each wavelength makes it possible to reliably detect cosmic-ray hits that land on the object spectrum. A cosmic-ray hit near the object spectrum produces a distortion of the spatial profile that extends over a few adjacent wavelength pixels. Because the spatial profile of the object generally is different from the residue left by a cosmic ray, a hit can usually be recognized even if it is centered on the spectrum. There is no danger with this procedure of mistakenly removing strong narrow emission lines from the object spectrum, as such lines do not alter the spatial profile.

The effects of cosmic ray hits can be eliminated automatically by a rejection cycle similar to that employed in fitting polynomials to form the sky and profile images. At each wavelength λ, the image data $D_{x\lambda}$ and the corresponding variances $V_{x\lambda}$ are compared to the predicted data $f_\lambda P_{x\lambda} + S_{x\lambda}$. The pixel with the largest value of

$$(D_{x\lambda} - f_\lambda P_{x\lambda} - S_\lambda)^2/V_{x\lambda}$$

is rejected if that value exceeds a preset threshold of 25. We then recalculate f and V (eqs. (8) and (13)) omitting the rejected pixel. This procedure is then iterated, allowing one additional pixel to be rejected on each iteration, until the set of rejected pixels converges (typically 1–3 iterations is required).

The pixel rejection scheme used at this stage is more cautious than the one used during the polynomial fits to sky and spatial profile images because a cosmic-ray hit can wipe out a significant fraction of the pixels available at a given λ. We therefore permit only one additional pixel to be rejected on each iteration. Cosmic-ray hits always increase the data numbers, suggesting that only pixels with positive residuals should be rejected, but the effects of localized charge-transfer defects and other CCD blemishes can be reduced if pixels with large negative residuals are rejected as well. (The spectrum is usually placed on a part of the CCD free of such blemishes.)

F. *Iteration*

A brief comment on the iteration implied by Step 9 in Table I is in order. Iteration is a necessary part of the optimal extraction algorithm because the spatial profile image P, the variance image V, the bad-pixel mask M, and the object spectrum f are all determined by a self-consistent solution of the equations given in Steps 5–8.

Details of the iteration scheme should not have a significant effect on the extracted spectrum, provided enough iterations are performed to achieve a roughly self-consistent solution. However, the amount of computing time required by the algorithm will depend on the iteration scheme and the number of iterations performed. I have not explored this aspect of the problem in detail. In my implementation, which was designed for computational speed and minimal storage requirements, the extraction is broken into two stages each of which involves an iteration.

The first stage iterates Steps 5 and 6 to determine the polynomials $P_x(\lambda)$ that define the profile image $P_{x\lambda}$. For each value of x the weighted polynomial fit is repeated until the set of pixels rejected is the same on two successive iterations. Because this stage uses the initial spectrum estimate found in Step 4, a handful of good pixels—those with λ the same as but x different from a cosmic ray

hit—are incorrectly rejected. This small loss of information is acceptable, however, because of the large number of wavelengths; the polynomial fits are not significantly altered by the omission of a few pixels (see Fig. 1).

The second stage uses the spatial profiles determined in the first stage to extract the one-dimensional object spectrum. For each wavelength λ Steps 6, 7, and 8 are iterated until the set of rejected pixels converges. Step 8 essentially scales the fixed spatial profile to achieve a weighted fit to the sky-subtracted data.

III. Examples

The optimal spectrum extraction algorithm was developed to analyze time-resolved CCD spectra of dwarf novae. The main goal of these observations is to detect absorption lines from the cool secondary stars of these short-period binary systems in order to determine spectral types and to measure radial-velocity curves. A series of 24 spectra of Z Chamaeleontis were acquired on 1984 March 13 with the CCD spectrograph on the CTIO 4-m telescope. The individual spectra had short (7-min) exposure times in order to prevent smearing of spectral lines by the orbital motion of the binary. Consequently, the CCD readout noise (13 electrons rms) is significant. The spectra also suffer from numerous cosmic ray hits.

Figure 2 compares spectra of Z Cha extracted with the standard and optimal procedures. The increased signal-to-noise ratio obtained with the optimal extraction procedure is evident in the representative single exposure. Two cosmic-ray hits affecting the standard spectrum are eliminated by the optimal extraction.

The Na I λλ8183,8194 absorption doublet from the secondary star of Z Cha is distinctly visible when the 24 optimally-extracted spectra are averaged. (The spectra were individually shifted before averaging to remove the orbital motion of the secondary star, but the doublet lines are blended by the rotation of the secondary star, which is synchronous with the 107-min binary-orbit period.) When the same analysis is performed with the standard spectra, the presence of the Na I lines is obscured by the

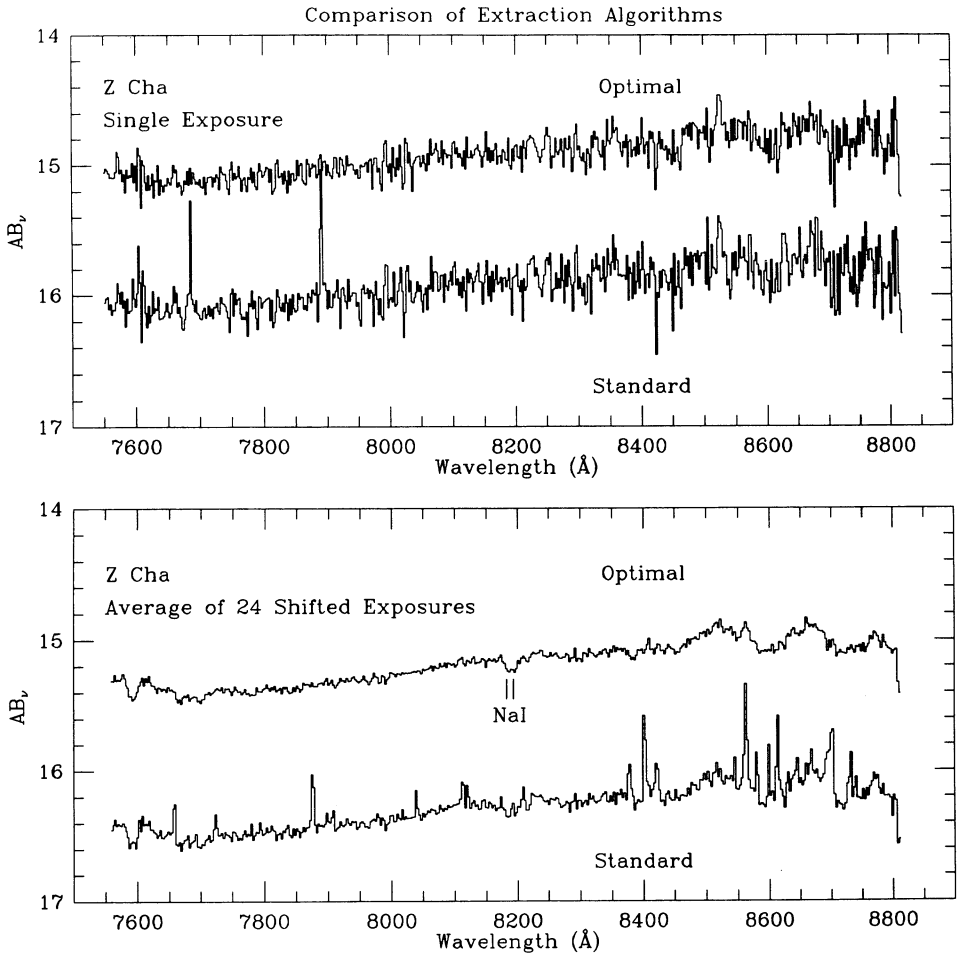

FIG. 2–Optimal and standard spectra of the dwarf nova Z Cha. Increased signal-to-noise ratio and rejection of cosmic rays are illustrated by the representative single exposure in the upper frame. The lower frame shows the average of 24 such spectra, each first shifted to remove the radial-velocity variation of the secondary star. The optimal spectrum reveals Na I absorption from the secondary star of Z Cha. In the standard spectrum, shifted by 1 magnitude, the doublet is lost in the noise.

larger statistical noise and the accumulation of cosmic ray hits. The optimal extraction algorithm was crucial to the success of this experiment, which will be reported in more detail elsewhere (Wade and Horne 1986).

Figure 3 compares the standard and optimal spectra extracted from a 3-s CCD spectrum of HD 64722, an early B star that was observed to aid in the removal of numerous telluric absorption lines from the Z Cha spectra. The two extraction methods give nearly indistinguishable results in this high signal-to-noise case, demonstrating explicitly the expected photometric capability of the optimal extraction procedure at a level considerably better than 1%. Glitches visible in the ratio of the two spectra correspond to several charge-transfer defects in the CCD. These defective pixels were automatically detected as outliers and rejected by the optimal extraction algorithm.

To indicate the computation time requirements of spectrum extraction, CPU times were recorded at various stages during the reduction of CCD spectra with the program PAMELA running on a VAX 11/780 computer at STScI. For 146 × 650 pixel CCD images, bias subtraction, including scaling to match the overscan region, required about 1 s, and division by the balance factor image took 1.9 s. These times are for 16-bit integer images, and would be roughly doubled for 32-bit floating-point arithmetic. The most costly step in this reduction was the creation of sky background images, which required 35 s to fit cubic polynomials to 69 sky pixels at each of the 576 wavelengths, and to evaluate the resulting sky polynomials over a 70 × 576 pixel region. At each wavelength the polynomial fit was performed at least twice, the first fit being used to reject bad pixels and to improve the variance estimates used for the second fit; a third fit was performed whenever new pixels were rejected by the second fit. On average about 100 pixels from the sky region were rejected. The sky evaluation was a rather larger task in these trials than would normally be the case because spectra were needed of two objects separated by 50 pixels on the slit.

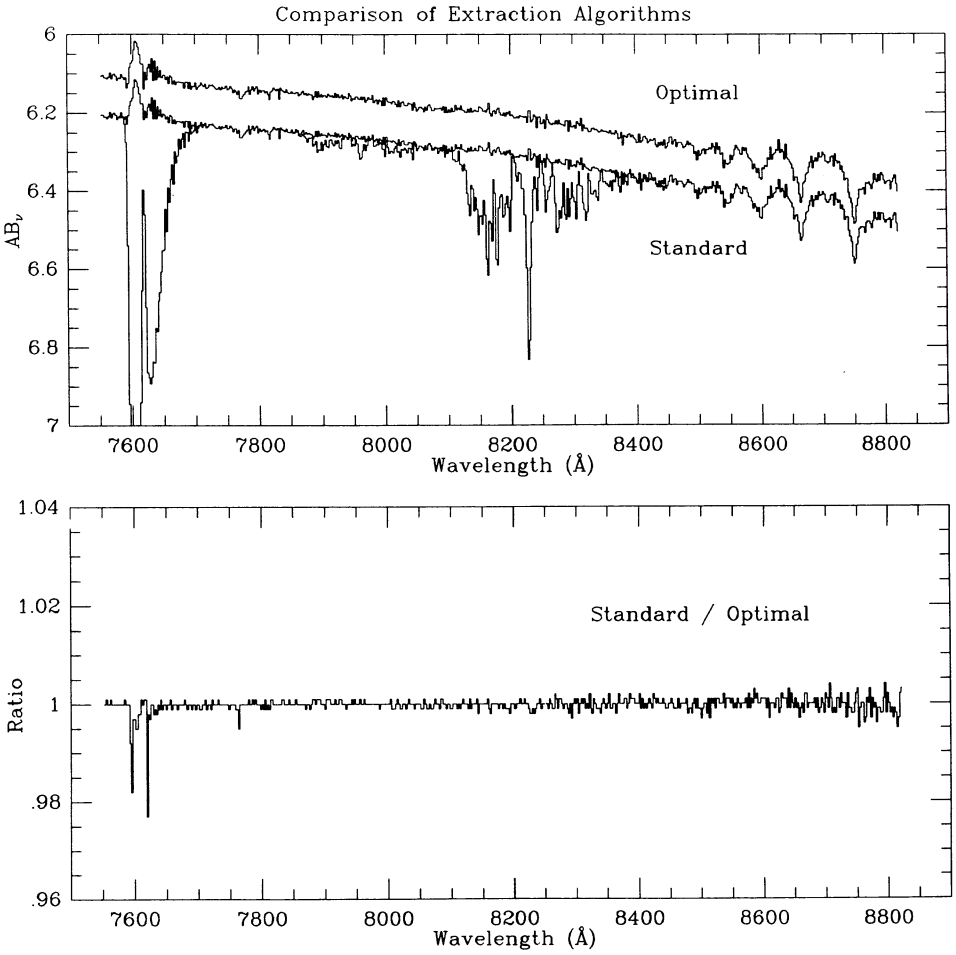

FIG. 3–Optimal and standard spectra of the early B star HD 64722. The optimal and standard extractions are nearly identical in this high signal-to-noise case, demonstrating the photometric capability of the optimal extraction algorithm. (The standard spectrum is shown both before and after removal of telluric absorption lines.)

The extraction of a one-dimensional spectrum from a 5×576 pixel region required 0.5 s by the standard extraction procedure. For this same extraction region, the optimal extraction procedure required 3 s plus 0.6 s for each rejected pixel. Most of this time is taken up by the polynomial fits that define the spatial profile at each wavelength. Cosmic-ray hits affecting the spectrum were detected in about half of the 5-min exposures, each hit resulting in the rejection of one to four pixels. As these results indicate, the computation time requirements for optimal extraction are not significantly greater than for the standard extraction procedure.

IV. Conclusion

We have described an optimal spectrum extraction algorithm and illustrated its performance with spectrum data obtained with a CCD. The main advantage offered by this optimal extraction method in comparison with conventional extraction procedures is improved signal-to-noise ratio, which can be the equivalent of a 70% increase in the exposure time. The performance gains are greatest in the regime of background-limited spectroscopy, such as spectroscopy of very faint targets against the night-sky background and time-resolved spectroscopy of rapid variables.

The improvement in signal-to-noise ratio is not achieved at the expense of spectrophotometric capability—the algorithm is thus suitable generally for both high and low signal-to-noise ratio spectra. The algorithm can reliably detect and eliminate the effects of cosmic-ray hits that fall on the object spectrum. Small geometric distortions of the spectrum are automatically accounted for, and the extracted spectra are not sensitive to the choice of object limits. These features make the algorithm particularly attractive for fully-automatic implementations in which spectrum reductions are to be performed with a minimum of human interaction.

The main limitation of the algorithm is its requirement that the spatial profile of the object is a *smooth* function of wavelength. Thus while ideal for spatially-unresolved targets, optimal extraction is not appropriate if spatial resolution is required, or if the spatial profile of the target varies rapidly with wavelength, as for objects with spatially-extended emission-line regions. Similarly, charge transfer effects at low exposure levels in CCD data may violate this assumption.

The computer time required for an optimal extraction is not significantly greater than that for a standard spectrum extraction. Most of the time is expended on weighted polynomial fits to the sky background and to the spatial profile. For most applications the modest increase in computer time is easily justified by the improved results.

We have given a comprehensive description of the optimal extraction algorithm and presented examples of its performance with CCD data in the belief that other astronomers will be interested in taking advantage of optimal extraction and to encourage its implementation in spectrum reduction packages intended for general use.

My interest in the subject of optimal spectrum extraction was initiated by Todd Boroson, who at that time was developing an idea from Steve Shectman into a working algorithm somewhat different from the one presented here. I have also benefited from discussions on this topic with John Biretta and Richard Wade, and from the thoughtful criticism of an anonymous referee.

REFERENCES

de Boer, K. S., and Snijders, M. A. J. 1981, *IUE NASA Newsletter*, **14**, 154.

Robertson, J. G. 1983, *Anglo Australian Observatory User Memo*, **11**, 1.

Wade, R. A., and Horne, K. 1986, *Ap. J.*, in preparation.