Astron. Astrophys. 326, 924–928 (1997)

ASTRONOMY AND ASTROPHYSICS

A statistic for describing pulsar and clock stabilities

Demetrios N. Matsakis[1], **J. H. Taylor**[2], **and T. Marshall Eubanks**[1]

[1] U.S. Naval Observatory, 3450 Massachusetts Avenue NW, Washington DC 20392-5420, USA
[2] Joseph Henry Laboratories and Physics Department, Princeton University, Princeton NJ 08544, USA

Received 23 September 1996 / Accepted 2 June 1997

Abstract. We propose a statistical measure suitable for comparing the rotational stabilities of pulsars with one another and with terrestrial time scales. By a simple extension of notation long used in the clock community, we call the statistic $\sigma_z(\tau)$. Defined in terms of third-order polynomials fitted to sequences of measured time offsets, $\sigma_z(\tau)$ is sensitive to variations in the frequency drift rate of the clock or pulsar. We apply the new statistical formalism to real pulsar data and terrestrial time scales, and show that over the past 10 years these two kinds of data have comparable stabilities.

Key words: time – pulsars:general – pulsars: B1937+21 B1855+09

1. Introduction

The possibility of using the exceptional rotational stability of millisecond pulsars to generate a time scale has long been of interest (Backer et al. 1982, Rawley et al. 1987, Guinot & Petit 1991, Taylor 1991, Kaspi et al. 1994, Petit 1995). A reliable statistical measure is needed for studying the physics of pulsar rotation and comparing pulsar stabilities with those of terrestrial clocks. Clock data are commonly analyzed using a statistic called σ_y, the square root of the "Allan variance," which can be computed from second differences of a table of clock offset measurements. Second differences are used because all but a few specially designed standards are likely to have significant frequency biases, whereas most good clocks have very small frequency drift rates (changes in those biases). The frequency drifts of terrestrial clocks are analogous to pulsar spin-down rates, which are meaningless from a timekeeping point of view because their magnitudes vary widely and they must be estimated separately for each pulsar, using the timing data themselves.

Following Taylor (1991), we therefore suggest use of a statistic we shall call σ_z, related to *third* differences (or third-order polynomial variations) of the timing residuals. These are the lowest-order deviations remaining in a pulsar time series after the phase, frequency, spin-down rate, and astrometric parameters have been determined by comparison with terrestrial time, and their effects removed. Since it is inherently insensitive to effects that pulsars cannot measure, and, as we shall show, is more sensitive to redder noise than other commonly used measures, σ_z is ideally suited for comparing pulsar stabilities with those of other time scales.

Send offprint requests to: D. N. Matsakis

2. Choice of statistic

Standard procedures for characterizing frequency stability were reviewed in a classic paper by Barnes et al. (1971). In their notation, a nearly periodic signal of nominal amplitude V_0 and frequency ν_0 can be defined by the relation

$$V(t) = [V_0 + \epsilon(t)]\ \sin[2\pi\nu_0 t + \varphi(t)]\,, \tag{1}$$

and by counting its cycles one can make a clock. The functions $\epsilon(t)$ and $\varphi(t)$ represent random amplitude and phase fluctuations about the ideal, and it is assumed that $\epsilon \ll V_0$ and $\varphi \ll 1$. The phase deviations of the clock are further characterized by two functions

$$x(t) \equiv \frac{\varphi(t)}{2\pi\nu_0}\,, \tag{2}$$

$$y(t) \equiv \frac{\dot\varphi(t)}{2\pi\nu_0}\,, \tag{3}$$

which measure the instantaneous time offset and fractional frequency offset.

The stability of the clock is often measured by means of the Allan variance, defined by

$$\sigma_y^2 = \left\langle \frac{1}{2}(\bar y_n - \bar y_{n-1})^2 \right\rangle\,. \tag{4}$$

Here $\bar y_n = (x_n - x_{n-1})/\tau$ is the average fractional frequency offset during the nth measurement interval of length τ, and the angle brackets denote averaging over all available intervals of that length. The Allan variance can also be written as the mean square of normalized second differences of the clock offset measurements, defined by

$$D_2(t,\tau) = \frac{x(t+\tau) - 2x(t) + x(t-\tau)}{\sqrt{2}\,\tau}\,. \quad (5)$$

Here the normalization of D_2 has been chosen so that for zero-mean white frequency noise, σ_y, defined as $\langle D_2^2 \rangle^{1/2}$ equals the root-mean-square fractional frequency deviation. A related statistic, the modified Allan variance, replaces each $x(t)$ by its average over the appropriate sub-interval (Allan 1987), which adds sensitivity to the difference between white and flicker phase noise while retaining the spectral characteristics of the Allan variance for redder forms of noise.

The Allan variance is well suited to characterizing the performance of manmade clocks, even if they have fixed frequency offsets (which can be measured and removed). However, it is not particularly well suited for pulsar data (Taylor 1991, Petit & Tavella 1996), because pulsars also have sizable, *a priori* unknown frequency drifts. The second-difference procedure discards all information about initial phase and frequency offsets between the clocks. For comparison with pulsar data we wish also to ignore fixed frequency drifts, which suggests using normalized third differences:

$$D_3(t,\tau) = \frac{x(t+\frac{\tau}{2}) - 3x(t+\frac{\tau}{6}) + 3x(t-\frac{\tau}{6}) - x(t-\frac{\tau}{2})}{2\sqrt{5}\,\tau}. \quad (6)$$

Here the normalization has been chosen so that for white phase noise the rms of D_3 will equal the rms of the original phase data, divided by τ, which is here redefined so as to symmetrically encompass the entire interval. Although this definition is chosen so that both D_2 and D_3 will be dimensionless, one should keep in mind that they characterize fundamentally different orders of variation.

Pulsar timing observations tend to be made at irregular intervals for which the differencing technique is not directly applicable. Taylor (1991) used the fact that fitting cubic polynomials to timing data segments of length τ is computationally convenient, and for equally spaced data leads to equivalent results (see also Deeter & Boynton 1982, Deeter 1984). For correspondence with the third difference defined in Eq. (6) the cubic term should be multiplied by a factor $\tau^2/(2\sqrt{5})$, and allowance should be made if τ is defined as the time between individual data points. We define $\sigma_z(\tau)$ as the weighted root-mean-square of the coefficients of the cubic terms fitted over intervals of length τ. This definition differs from that less formally presented by Taylor (1991) by a scale factor and in emphasizing the importance of using a weighted rms.

A significant advantage of the polynomial approach is that the fitted terms are sensitive to all values of x in each measurement interval, instead of merely to four fiducial points spaced by $\tau/3$. This allows σ_z to distinguish between white and slightly "blue" phase noise in a manner analogous to modified Allan variance.

It is easy to describe how the computed $\sigma_z(\tau)$ of a time series will behave if the underlying function has known power-law spectral characteristics. Following the techniques and notation of Barnes et al. (1971), it is easy to show that, if the power spectral density can be modeled as

$$S_x(f) \propto f^{\alpha-2}\,, \quad (7)$$

then σ_z^2 will also follow a power law,

$$\sigma_z^2(\tau) \propto \tau^{\mu}\,, \quad (8)$$

with the exponents α and μ related by

$$\mu = \begin{cases} -(\alpha+1) & \text{if } \alpha < 3, \\ -4 & \text{otherwise.} \end{cases} \quad (9)$$

It follows that σ_z is a good statistic for the analysis of low-frequency-dominated "red" or "pink" phase noise, such as that typical of pulsar timing residuals, but not for "blue" noise processes with $\alpha > 3$.

At the risk of overstating the obvious, we shall mention that statistics such as σ_y and σ_z which operate in the time domain are in many ways complementary to power spectral analysis in the frequency domain (IEEE 1988). It is important to note, however, that conventional Fourier techniques necessarily fail in the presence of steep power-law spectra (exponents > 2 in absolute value), because of spectral leakage between frequency bins. In contrast, the effective "filters" corresponding to the $\sigma_z(\tau)$ technique have steep cutoffs (proportional to f^6) on their low-frequency sides, and hence are ideal for analyzing red noise (Stinebring et al. 1990).

3. Recipe for computing σ_z

Our recipe for the computation of σ_z is as follows:

1. Make a chronological list of times, residual clock differences, and measurement uncertainties, namely $t_i, x_i, \sigma_i, i = 1, \ldots, N$. The total interval spanned by the data is $T = t_N - t_1$.

2. To find $\sigma_z(\tau)$, divide the data, according to time of observation into subsequences defined by continuous intervals of length τ and arbitrary offset time t_0. Fit the cubic function

$$X(t) = c_0 + c_1(t-t_0) + c_2(t-t_0)^2 + c_3(t-t_0)^3 \quad (10)$$

to the data in each subsequence by minimizing the weighted sum of squared differences, $[(x_i - X(t_i))/\sigma_i]^2$. Then set

$$\sigma_z(\tau) = \frac{\tau^2}{2\sqrt{5}} \left\langle c_3^2 \right\rangle^{1/2}\,, \quad (11)$$

where angle brackets denote averaging over the subsequences, weighted by the inverse squares of the formal errors in c_3. Since $X(t)$ has units of time, σ_z will be dimensionless.

Insist that for each valid subsequence there be at least four measurements, and that the interval between first and last measurement be at least $\tau/\sqrt{2}$. To allow for unmodeled error sources it may be desirable to weight all measurements equally by setting σ_i to a constant; in any case, the weighting scheme used

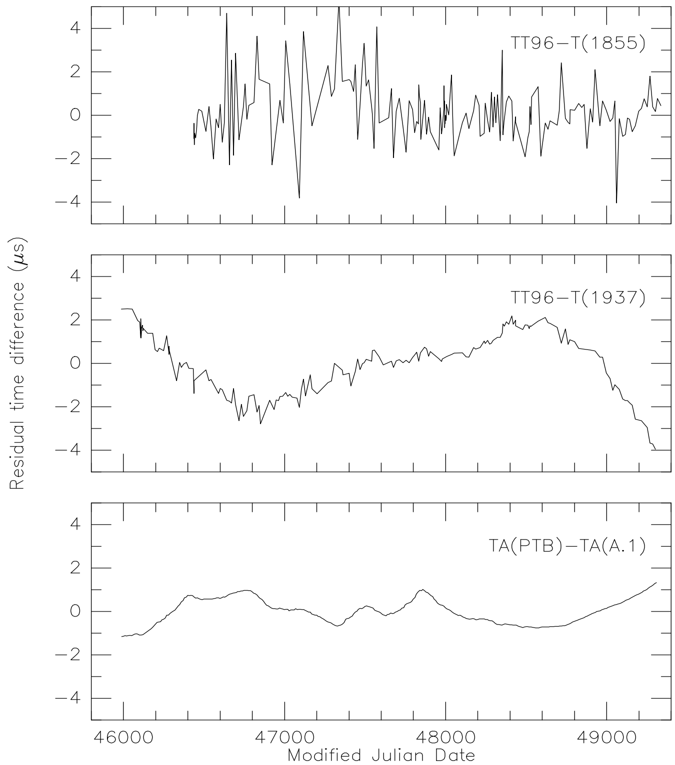

Fig. 1. Residual time differences for PSRs B1855+09 and B1937+21 with respect to TT96, and between the free-running time scales TA(PTB) and TA(A.1), after removal of best-fitting parabolas

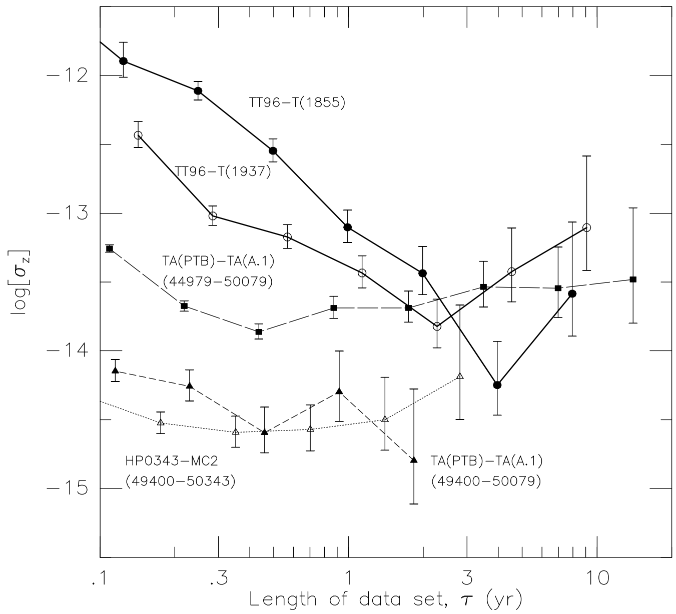

Fig. 2. Fractional stabilities $\sigma_z(\tau)$ for the three data sets illustrated in Fig. 1, and for two shorter data sequences of atomic clock differences: TA(PTB)−TA(A.1) extending from MJD 49400 to 50079, and the difference between a particular Hewlett Packard Model 5071 cesium standard and the USNO Master Clock, which is steered daily to approximate TAI.

should be reported. For computational reasons, t_0 is best chosen to lie near the midpoint of each subsequence.

3. While this procedure will work for any τ, we recommend computing $\sigma_z(\tau)$ only for $\tau = T, T/2, T/4, T/8, \ldots$ since other values of σ_z are not independent of these. We also recommend that the set of intervals be only the adjacent non-overlapping ones covering the full time range. In some cases a formally more accurate estimate can be obtained by also including overlapping intervals of length τ, but in the presence of red noise the derived values of σ_z would not be significantly improved and the error analysis would be more complicated. If such techniques are applied, they should be explicitly identified.

If the long-term average values of $\langle c_3 \rangle$ for the underlying stochastic noise process are zero, and if the recommended binning restrictions are followed, the computed values of σ_z^2 will have χ^2 distributions with n degrees of freedom, where n is the number of squared values of c_3 appearing in the average in Eq. (11). For nearly all cases of interest the resulting sampling errors will dominate the uncertainties. In Appendix A we outline how to compute these uncertainties.

To promote standardization we have written a Fortran program to compute σ_z and its statistical uncertainties, and have made it publicly available by anonymous ftp to URL ftp://tycho-.usno.navy.mil/pub/sigma_z. It can also be found on the World Wide Web at the USNO pulsar home page http://tycho.usno-.navy.mil/pulsar.html.

4. Observed σ_z of pulsars and time scales

We have applied the recipe of Sect. 3 to the data in Fig. 1, which includes timing residuals of the two longest-observed millisecond pulsars, PSRs B1855+09 and B1937+21 (all but the most recent of which are publicly available from Princeton, see Kaspi et al. 1994), and terrestrial clock data obtained by anonymous ftp from the Bureau International des Poids et Mesures (BIPM 1996). The pre-fit pulsar data were referenced to the terrestrial time scale TT96 (see Guinot, 1994), and we designate the residuals TT96−T(1855) and TT96−T(1937), respectively. In Fig. 2 we present values of σ_z for all three time series, computed with equal weighting for all measurements. For both pulsars the slope of the log-log graph is close to -1.5, at least up to intervals τ of several years, as expected for residuals dominated by uncorrelated measurement errors. For the clock differences, and for the pulsars at $\tau > 1$ year, the curves are dominated by low-frequency noise believed to be intrinsic to the pulsars and the atomic time scales.

Also included in Fig. 2 are values of σ_z for the differences between several terrestrial time scales (BIPM, 1992), again using equal weighting for all measurements. Note that at large τ the pulsar stabilities are comparable with that of the difference between two independent, free-running time scales: TA(A.1) maintained by the U.S. Naval Observatory, (whose 50-clock ensemble currently contributes about 40% of the weight to the BIPM's determination of TAI), and TA(PTB), maintained by the Physikalisch-Technische Bundesanstalt (PTB) in Germany, which maintains two standards optimized for calibration and

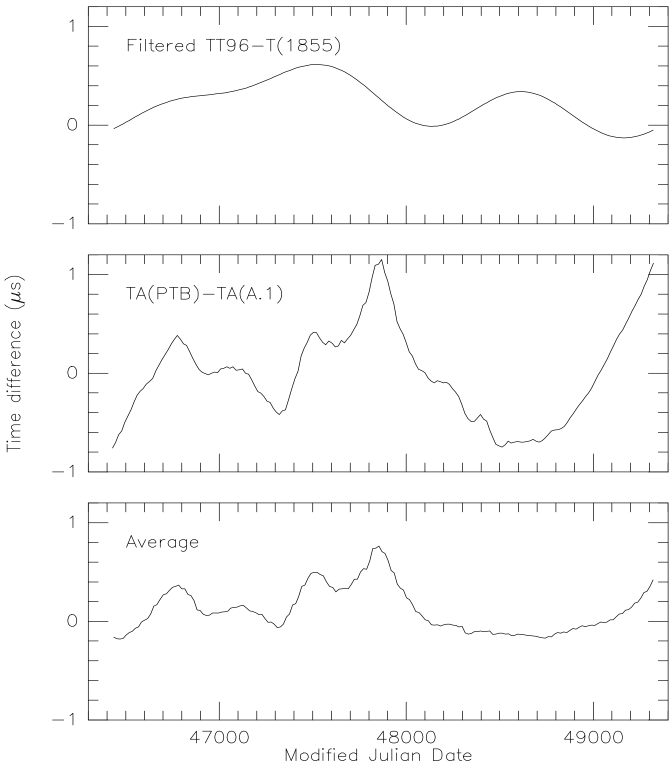

Fig. 3. An application of pulsar data to create an improved time scale. *Top:* timing residuals of PSR B1855+09 with respect to TT96, after filtering out all frequencies above 1/(970 days). *Center:* Residual difference TA(PTB)−TA(A.1) after fitting out a parabola. *Bottom:* Improved time scale generated by simple averaging of TT96-T(1855) and TA(PTB)-TA(A.1).

long-term accuracy. For this time period, the frequencies of these PTB standards were precise to about 3×10^{-15} when averaged over a year, and accurate to about 2.2×10^{-14} in an absolute sense (Bauch et al. 1987).

The 1996 computation of Terrestrial Time, TT96, was derived using TA(PTB) for the long-term information. Consequently for large τ, TA(PTB) can be considered equivalent to TT96. Recently there has been a dramatic improvement in terrestrial frequency standards (Breakiron & Koppang 1996) owing to increases in both the number and quality of atomic clocks. To illustrate this improvement we also include in Fig. 2 a σ_z curve based on just the latest TA(PTB)−TA(A.1) data, and also one showing the difference between a single Hewlett Packard Model 5071 cesium standard and the USNO Master Clock (Matsakis & Josties, 1996).

5. An application

The comparison of recent terrestrial clock data TA(PTB)-TA(A.1) with older data from the same series in Fig. 2 shows how terrestrial clock technology has improved over the last ten years, but it also suggests that long-term errors in the terrestrial time scales of the 1980's are making non-negligible contributions to pulsar timing residuals, and thus that pulsar data could be used to improve the terrestrial time scale. To test this hypothesis we used the technique outlined in Blandford et al. (1984) and Matsakis & Foster (1996) and applied a crude Wiener filter to the time series TT96−T(1855). We spline-filtered this 177-point data set to yield 177 equally spaced pseudo-measurements over the 8-year interval, and then Fourier transformed them. We made an approximate allowance for the frequency-dependent signal-to-noise ratio by discarding all but the three independent Fourier frequencies for periods larger than 800 days, and then transformed the data back to the time domain. The filtered data from TT96−T(1855) were then averaged with those from TA(PTB)−TA(A.1). Curves created by averaging TT96−T(1855) with TT96-TA(A.1) generate a similar average. Since our σ_z analysis indicated the absence of extremely red noise, spectral leakage was not an issue in this case.

Although this technique is not optimal, Fig. 3 shows that averaging the smoothed pulsar data with TA(PTB)−TA(A.1) has made a noticeable improvement in stability of their combined time scale. With a longer data set the improvement should be even more significant. Over long time intervals TA(PTB) and hence TT96 are each known to be more stable than TA(A.1), so the improvement shown here (and the smaller variances of the top and bottom panels of Fig. 3, compared with the center panel) can be taken as demonstrations that T(1855) also was more stable than TA(A.1) over the time range analyzed. A similar conclusion was reached by Kaspi et al. (1994).

6. Conclusion

The stastistic σ_z can complement more commonly used measures in the analysis of pulsar and terrestrial clock data because it is insensitive to the phase offset, average frequency, and average drift while allowing one to study the long-term noise characteristics of the data.

Acknowledgements. We thank Julie Zachariades for assistance in analyzing the USNO clock data, and Steve Thorsett and Roger Foster for many helpful discussions.

Appendix A: statistical errors in σ_z computations

As stated in the text, it will usually be the case that the computed values of σ_z^2 will have χ^2 distributions with n degrees of freedom, where n is the number of squared values of c_3 appearing in the average. For small n the distributions are significantly skewed, and most estimates of σ_z will be biased low because the median of the distribution is less than the average. A table of the relevant biases and the ranges of "1σ error bars" that enclose 68% of the expected distributions can be generated by evaluating incomplete gamma functions $P(a,x)$ (see, for example, Press et al. 1986). In particular, for each n we want to find the values of x_{16}, x_{50}, and x_{84} that satisfy the relations

$$P(0.5n, 0.5nx_{16}) = 0.16\,, \tag{A1}$$

$$P(0.5n, 0.5nx_{50}) = 0.50\,, \tag{A2}$$

$$P(0.5n, 0.5nx_{84}) = 0.84\,. \tag{A3}$$

For a base-10 logarithmic plot of σ_z, the bias correction

$$b = -0.5 \log x_{50} \tag{A4}$$

should be added to the biased estimate given by Eq. (11). Positive-going and negative-going error bars centered on the corrected value of σ_z should then have lengths

$$\delta_+ = -0.5 \log x_{16} - b\,, \tag{A5}$$
$$\delta_- = 0.5 \log x_{84} + b\,, \tag{A6}$$

respectively. Note that these are defined so as to specify what range of values of σ_z are consistent with the data rather than what range of observed σ_z would be obtained if the actual value were as measured.

Good approximations for b, δ_+, and δ_- can be obtained from the simple relations

$$b \approx 0.17/n\,, \tag{A7}$$
$$\delta_+ \approx 0.31/\sqrt{n-1} \text{ if } n > 1, \text{ 0.52 otherwise}, \tag{A8}$$
$$\delta_- \approx 0.31/\sqrt{n}\,. \tag{A9}$$

Because of the decidedly non-gaussian statistics of the underlying errors, these equations are not useful for making detailed comparisons of data with a model. We emphasize that they are intended solely for the purpose of making informative graphs, and note that in the limit of large n,

$$\delta_+ \approx \delta_- \approx 0.31/\sqrt{n}\,. \tag{A10}$$

References

Allan D. 1987, IEEE Trans. Ultrasonics, Ferroelectrics, and Frequency Control 34, 647

Backer D. C., Kulkarni S. R., Heiles C., Davis M. M., Goss W. M. 1982, Nat 300, 615

Barnes J. A., Andrew R. C., Cutler L. S., et al., 1971, IEEE Trans. Instrum. Meas. IM-20, 105

Bauch A., Dorenwendt K., Fischer B., et al., 1987, IEEE Trans. Instrum. Meas. IM-36, 613

Blandford R., Narayan R., Romani R.W. 1984, JA&A 5, 369

Breakiron L. A., Koppang P. 1996, Proc. 1996 IEEE Frequency Control Symposium, in press

Bureau International des Poids et Mesures 1992, Annual Report, Sevres, France

Bureau International des Poids et Mesures 1996, anonymous ftp to 145.238.2.2, June 10, 1996

Deeter J.E. 1984, ApJ 281, 482

Deeter J.E., Boynton P.E. 1982, ApJ 261, 337

Guinot B. 1994, Metrologia 31, 431

Guinot B., Petit G. 1991, A&A 248, 292

IEEE Standard Definitions of Physical Quantities for Fundamental Frequency and Time Metrology, 1988, 1139, Institute of Electrical and Electronics Engineers, Inc., New York.

Kaspi V. M., Taylor J. H., Ryba M. F., 1994, ApJ 428, 713

Matsakis D. N., Foster, R. S. 1996, Amazing Light, Ed. R. Chiao, Springer Press, New York, 445

Matsakis D. N., Josties, F. J. 1996, Proceedings of the 27th Annual Precise Time and Time Interval Symposium, in press.

Petit G. 1995, Proceedings of the 26th Annual Precise Time and Time Interval Symposium, 387

Petit G., Tavella P. 1996, A&A 308, 290

Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T., 1986, Numerical Recipes, Cambridge University Press, Cambridge.

Rawley L.A., Taylor J.H., Davis M.M., Allan D.W. 1987, Sci 238, 761

Stinebring D.R., Ryba M.F., Taylor J.H., Romani, R.W. 1990, Phys. Rev. Lett. 65, 285

Taylor J.H. 1991, Proc. IEEE 79, 1054

This article was processed by the author using Springer-Verlag TEX A&A macro package version 3.