

Active optics

I. A system for optimizing the optical quality and reducing the costs of large telescopes

R. N. WILSON, F. FRANZA and L. NOETHE

European Southern Observatory, Karl-Schwarzschild-Strasse 2,
D-8046 Garching, F.R. Germany

(Received 1 September 1986; revision received 26 January 1987)

Abstract. A system of 'active optics' control for the optical imagery of astronomical telescopes has been under development in the European Southern Observatory for about ten years. Its first application will be in the 3.5 m New Technology Telescope (NTT) scheduled for operation in 1988. A model test with a thin 1 m mirror (aspect ratio 56) has given remarkably successful results which will be reported in Part II of this paper. Part I gives a complete presentation of the theoretical principles of this technique of active optics and its scope of application. The subject is treated from the viewpoint of the *temporal band-pass* of error sources, 'active optics' being concerned with the low-frequency band-pass. The high-frequency band-pass ('adaptive optics') is principally concerned with atmospheric correction and is only briefly referred to for comparison. 'Active optics' correction of the low-band-pass system errors should bring major improvements in image quality together with a large cost reduction. While its impact on ground-based telescopes seems beyond doubt, the most important application should be in space projects where permanent diffraction-limited performance should be possible with much relaxed tolerances.

1. Introduction

For over 300 years since its invention, the reflecting telescope has been *passive* in the sense that its optical system is designed, manufactured and set up according to a pre-determined design and subsequently only modified by some *off-line* maintenance intervention. Its optical quality at any given time is then subject to both the (invariable) manufacturing errors and the (variable) errors of adjustment and support. The latter have always been serious but are increasingly so with modern lighter, more flexible optics. An *active* telescope, in the sense we use the term here, has an in-built control system which enables the optical quality to be checked and optimized automatically at any time at the touch of a button. So far as stable or slowly varying telescope errors are concerned, the optical quality, even for the largest optical elements, can be made effectively diffraction limited *always*, while at the same time allowing major relaxation of manufacturing tolerances. The application to large telescopes is the most obvious use, but such a system can be used for any optical system whose elements have a size or flexibility justifying it.

A system for *active optics* control with general application to large telescopes has been under development at the European Southern Observatory (ESO) since about 1980. Although its origins go back to 1977 [1, 2] with the aim of improving the functional quality of existing large telescopes, the development of a complete system started with the decision by ESO in 1980 to build the 3.5 m *New Technology Telescope*

(NTT) [3]. This telescope envisaged technological innovations in three areas concerned directly or indirectly with the optics or optical quality:

- (1) a rotating building based on the concept of the MMT building in Arizona;
- (2) a metal (aluminium) prime mirror as well as a 'conventional' one in glass ceramic; and
- (3) active-optics control of image quality.

Of these features, the first is concerned with the impact of the local air conditions on the optical quality. The second (metal mirror) aspect was cancelled, purely for reasons of pressures on the time scale, and transferred to the subsequent ESO 16 m *Very Large Telescope* (VLT) project [4], for which more development time was available [5]. The third, the active-optics control system, is perhaps the most important new technology feature of the NTT and is the subject of this paper.

The basic principles of the active-optics system for the NTT have already been reported in the literature [6–8] and some details of the layout of the active prime mirror support have also been published [9]. The NTT is now in the fabrication phase and is scheduled for completion (first light) in the Autumn of 1988. In view of its importance and novelty, it was deemed necessary to test the system on a small-scale model mirror of 1 m diameter. These tests were successfully completed and a few essential results have been published in a brief form [10].

It is the purpose of the present paper to give a complete account of the principles, advantages and scope of application of this active-optics system; a detailed description of the results and conclusions of the confirmatory experiments with the 1 m test mirror will follow in Part II.

2. What 'active' optics can do: its complementary nature to 'adaptive' optics

2.1. *The limits to optical performance of telescopes*

2.1.1. *Diffraction*

The ultimate physical limit is, of course, set by *diffraction* and the classical formula $d_a = 2.44\lambda/D$ gives the diameter of the Airy disc in radians. For wavelength $\lambda = 500$ nm and aperture $D = 5$ m (Palomar), $d_a = 0.050$ arcsec.

2.1.2. *Atmospheric turbulence (seeing)*

In practice, with large ground-based telescopes, diffraction has been of no significance because *atmospheric turbulence (seeing)* d_b has set a limit at least an order of magnitude greater. Although better seeing is claimed for some modern sites [11], the accepted value for most of the world's observatories has been $d_b \geq 0.5$ arcsec. Because of the 'wings' of the function, d_b means, in practice, the diameter containing about 75–80 per cent of the energy in the image of a point source.

For apertures larger than the Fried parameter [12], d_b is effectively independent of the aperture D . Even for sites like La Silla, the ESO observatory, which are accepted as being very good, the Fried parameter rarely exceeds 30 cm [13], so for large telescopes d_b can certainly be taken to be independent of D .

2.1.3. *Telescope quality*

The optical specification to the manufacturer has hitherto set the limit to the image quality achievable. Large conventional telescopes built since 1945 have all had

specifications d_c roughly in accord with the best seeing expected so that the final instrumental performance of such telescopes has also been $d_c \simeq 0.5$ arcsec, also independent of D . Again, because of the 'wings' of the function, d_c means in practice the diameter containing about 75–80 per cent of the energy of the image of a point source.

Unfortunately, the fact that the optical manufacturer has met the above specification d_c for the basic optical elements does not mean that the functioning telescope has this quality. A large number of technical factors prevent most telescopes (large or small) from achieving their specification for more than a small fraction of their practical observing time. This is the basic problem addressed by 'active' optics (the second aspect it addresses is the improvement of the optical quality beyond that delivered by the optical manufacturer). To make this clear, the table lists ten sources of error which can afflict the optical image of a telescope. These are considered from the point of view of their *band-pass*, the temporal frequency range over which they can vary.

It is clear from the table that optical design and optical manufacture are once-for-all d.c. effects. In fact, optical design is only of significance for telescopes with significant fields: correction on axis is trivial. Optical manufacturing errors are, of course, by no means insignificant, even when the greatest care is taken, and they normally more or less absorb the seeing-related specification d_c . Of the theoretical errors of the supports or structure, the latter is principally due to lateral tube flexure or longitudinal sag (compression). The band-pass upper limit in this case is therefore set essentially by telescope movements of tracking and pointing. Maintenance errors are usually more serious than theoretical errors, indeed they are often the most serious of the lot! Typically they include all malfunctions of supports, misalignments of elements due to changeovers of focus station or other reasons, and other maladjustments such as focus or axial malpositioning of optical elements. The band-pass is very low because such degradations usually take place over weeks or months after an adjustment has been made.

Ten sources of error giving degradation of image quality in telescopes, and their corresponding band-passes.

Source of error	Band-pass (Hz)
Optical design	d.c.
Optical manufacture	d.c.
Theoretical errors of:	
Mirror supports	d.c. $\rightarrow 10^{-3}$
Structure (focus, centring)	10^{-3}
Maintenance errors of the structure and mirror supports	10^{-5} (weeks)
Thermal distortions:	
Mirrors	10^{-4}
Structure	10^{-3}
Mechanical distortion of mirrors (warping)	10^{-6} (years)
Thermal effects on ambient air	$10^{-3} \rightarrow 10^2$
Mirror deformation from wind gusts	$10^{-1} \rightarrow 2$
Atmospheric turbulence	$2 \times 10^{-2} \rightarrow 10^3$
Tracking errors	$5 \rightarrow 10^2$

Thermal distortions are essentially low band-pass since the mass and hence the thermal capacity involved in the mirrors and structures of large telescopes is bound to be considerable, even if measures are taken in modern telescopes to reduce mass compared with conventional ones. Mechanical distortion of mirrors has had little significance in the past because of the high form stability of glasses; but the consideration of other materials [5, 14] (metals or composites) draws attention to this potential source of error. Again, because the processes of stress release or structural change are very slow, the band-pass is very low.

The first error associated with the air is due to thermal effects on ambient air. Effects in the ambient air are normally referred to as 'telescope seeing', 'dome seeing', or 'site seeing' and are not always easy to distinguish from external seeing. The band-pass can range from quite low frequencies to high frequencies as associated with external seeing (e.g. at dome slits). For many conventional telescopes, this is one of the worst error sources. Mirror deformation from wind gusts is a problem which is of considerable importance for modern, very large telescopes with relatively flexible mirrors, operating in a fairly open environment to reduce errors due to the previously mentioned ambient air effect. In conventional telescopes with domes the problem scarcely existed, the ambient air effect being correspondingly more dangerous. Once again, since big mirrors have considerable mass and inertia, they act as low band-pass filters for the wind gust spectrum.

Atmospheric turbulence (seeing) is the effect which, given perfect maintenance and site conditions, would normally be the limiting factor. Of all the error sources, its band-pass is easily the greatest. Finally, tracking errors are mainly caused by vibration since lower frequencies are absorbed by the autoguider. The lower band-pass limit should always be pushed as high as possible by increasing the stiffness, whereby the amplitudes are reduced.

2.2. 'Active' optics and 'adaptive' optics: definitions

From a semantic point of view, the use of these terms is completely arbitrary and has led—and still leads—to considerable confusion. From its origins in 1976 [2], we have used the term 'active optics' for our *closed-loop* correction system for telescopes, using an image analyser. However, in certain circumstances, using calibration, this system may also be applied in open-loop systems. In 1979 Pearson, Freeman and Reynolds [15] proposed the general definition that the term 'adaptive' should imply closed-loop control while 'active' should be used for open-loop. In 1982 Woolf [16], referring more specifically to the application of such systems to telescopes, proposed the use of the term 'adaptive' for atmospheric seeing correction and the term 'active' for the correction of telescope errors. Since this definition by Woolf essentially agreed with our own, we have adopted it and define these terms somewhat more rigorously in the following way according to *band-pass*.

2.2.1. 'Active' optics

In the basic system for the NTT (d.c. $\rightarrow \frac{1}{30}$ Hz), the upper limit of $\frac{1}{30}$ Hz is simply a consequence of the fact that, with good seeing at a site like La Silla, 30 s is required to integrate out external seeing, that is to give an effectively circular image having the size of the seeing disc. This limit, the result of much practical telescope testing, thus allows the physical separation of the low-frequency terms from the seeing with any integrating detector.

This definition can be extended for wind gust effects for the VLT† unit telescopes (8 m): ca. 10^{-1} Hz \rightarrow ca. 2 Hz.

2.2.2. 'Adaptive' optics

With this definition, 'active' optics is concerned with all low band-pass error sources, 'adaptive' optics [17] with high band-pass sources (ca. 10^{-1} Hz \rightarrow ca. 10^3 Hz). There is an overlap band in the middle for which separation of effects, if possible at all, is difficult. However, such separation may not be necessary if one thinks strictly in terms of band-pass. Specifically, if the 'active' optics system for the VLT corrects not only wind buffet effects on the primary but also some low-frequency atmospheric effects, this is a gain, not a loss—provided, however, that the image analysis and the observed object are within the same isoplanatic field for the bandpass in question.

'Adaptive' optics is thus concerned from the table principally with errors due to atmospheric turbulence and with the upper frequency end of errors associated with thermal effects on ambient air. From the band-pass, there will be confusion with tracking error, although the physical origins are quite different. *All the other error sources of the table are in the band-pass of active optics.* This is a fact of great importance.

2.3. Reasons for the separation of 'active' and 'adaptive' optics correction systems

There are four basic reasons why systems for 'active' and 'adaptive' correction should be physically separated:

2.3.1. Field

In *active optics* none of the error sources in its band-pass have field limitation: the effects are essentially constant over the field. The logical point in the system for applying the correction is therefore the pupil. If the entrance pupil is used, the full field of the telescope is automatically corrected; whereas if a transferred pupil is used in, say, a Cassegrain telescope, only a small field can, in practice, normally be corrected.

Conventionally, the entrance pupil is placed at the primary mirror since most efficient use is then made of this, the most expensive element. Some telescopes, with strong emphasis on i.r. observation, place the pupil at the secondary. If the pupil is at the primary and the active correction is performed there, the beam moves over the secondary according to the point in the field observed. However, for the small fields we have in practice with large telescopes, this shift of the beam is negligible for the low spatial frequency terms corrected (see §4). In practice, then, active correction may be done either at the primary or the secondary, whichever is technically more convenient. In the NTT and for the VLT concept, it was concluded that the technical solution is easier at the primary, which is also defined as the pupil.

In the *adaptive optics* case, atmospheric seeing has an isoplanatic (coherence) angle in the visible wavelength band going from about 2 arcmin for low frequencies to 10 arcsec or less for high frequencies. Following Kolmogorov statistics, this angle increases with wavelength. In the i.r. range, this eases the problem, which is

† The ESO Very Large Telescope (VLT) project, still in the conceptual phase, envisages a linear array of four unit telescopes, each of 8 m aperture, giving an effective equivalent telescope of 16 m.

fundamental in the visible range, of finding a reference star in the isoplanatic field for getting correction information. A proposal has been made [18] to get round this problem with an artificial reference, but practical proof of its feasibility remains to be given. For the visible range, then, the most logical scheme for adaptive correction (and what has always been tried) is that shown in figure 1, using a transferred pupil.

The transfer optics and the correcting element are small because the isoplanatic field is small. Of course, in principle, one could aim to correct a large number of isoplanatic fields in parallel, covering the whole telescope field. However, such an enterprise for the full band-pass in the visible range would, even if the wavefront correction were physically possible, involve an information flow rate (of the order of 10^{11} wavefront slope information elements per second) which must remain quite utopian: space operation would probably be cheaper!

2.3.2. Temporal frequencies—band-pass

For active optics only a low band-pass is required so that a correction system with appreciable mechanical inertia is quite acceptable. Because of the low band-pass, then, active optics is easy.

Adaptive optics needs a high band-pass. Mechanical inertia is not acceptable which means a low-inertia correcting element (e.g. a deformable mirror) is essential. This fits in with the small-field concept of figure 1. Because of the high band-pass, adaptive optics is difficult, particularly in the visible wavelength band.

2.3.3. Wavefront amplitudes required for correction

There is a natural convergence of the amplitudes required and the temporal frequencies. While it may well be useful at the d.c. end of active optics to correct amplitudes up to hundreds of wavelengths, adaptive optics at high temporal frequencies will be concerned, at most, with a few wavelengths. The technical means of correction should therefore logically be adapted to these amplitude requirements.

2.3.4. Physical origin

Active optics is essentially concerned with effects which arise from the elastic behaviour of solids, whereas adaptive optics is largely dependant on the physics of the atmosphere, a gas.

Since the physical origin of the effects in the two cases is quite different, an algorithm which is appropriate to the one domain will not necessarily be the best for the other.

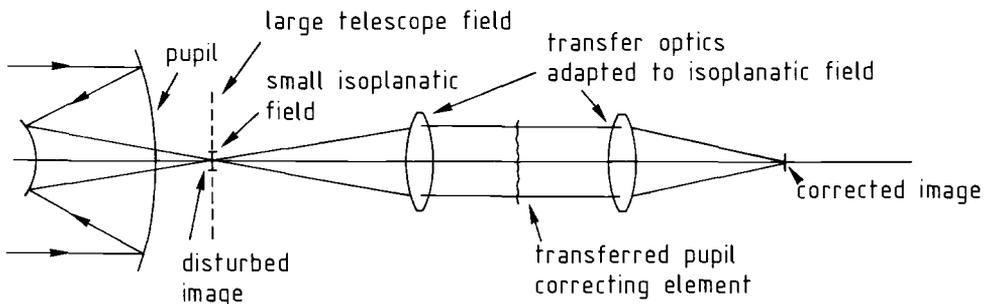


Figure 1. Adaptive optics (for high-frequency band-pass correction): normal basic technique using a transferred pupil and a small correction field adapted to the isoplanatic angle.

3. Active optics: the practical impact on telescope specifications

As a typical example of a conventional passive telescope, the ESO 3·6 m telescope had the following optical manufacturing specification for the image in the Cassegrain focus: 80 per cent of the geometrical optical energy (E_{80}) within 0·5 arcsec.

A passive telescope implies, of course, that adjustment can only be performed off-line from time to time, taking the telescope out of operation. There is normally no means in a passive telescope to monitor satisfactorily its optical quality on-line and no means of applying a correction on-line, even if this could be done.

The ESO 3·5 m NTT, an active telescope, had a dual specification for the manufacturer: $(E_{80})_{LF} \leq 0\cdot4$ arcsec, including low spatial frequencies (to be defined in §4); $(E_{80})_{HF} \leq 0\cdot15$ arcsec, for high spatial frequencies alone (e.g. ripple or zones).

The specification for its functional observing state is $E_{80} \leq 0\cdot15$ arcsec for all spatial frequencies, always (during the whole observing time, not just after alignment). The NTT tests itself as required and corrects its low spatial-frequency errors (the only ones that can vary—see §4.7) on-line. If the functional specification had been imposed on the optician, particularly bearing in mind the relatively thin primary with an aspect ratio of 15, the cost of optical figuring would have been vastly increased.

The functional specification implies, in practice, that the NTT will be very close to the normal diffraction limit (80 per cent Strehl intensity ratio). It will have three aspects of *active* control:

- (a) auto guiding (equivalent to active guiding, already used in many telescopes);
- (b) auto focus; and
- (c) auto correction of the image ((b) and (c) comprising the ‘active optics’).

It will thus permanently achieve its inherently optimum performance. We term this the *intrinsic quality* and will define it in §4.

The unit telescopes of the VLT will use the same system; but the dynamic range of correction will be much greater enabling a much greater relaxation of tolerances for the manufacturer, thereby giving even larger savings in costs and time.

4. The physical principles and practical function of the active-optics system

These principles have been reviewed in the cited literature [6–10], to which the reader is referred. For the sake of completeness, we shall briefly list them again here as a necessary introduction to a detailed analysis of modal calibration which we see as the essential key to simple and effective active optics control. A fuller description is given in the ESO Preprint (no. 484) of this paper.

4.1. Modal, closed-loop control using image analysis

Figure 2 shows the basic scheme. The prime mirror support is active and can be modified according to the algorithm to be explained. The other active element is the lateral and axial position of the secondary mirror. Light from a natural star (normally the offset guide star) is fed to the image analyser. This is of the Shack–Hartmann type [10, 19, 20] using a CCD detector. Combined with our software package for image analysis, we have named this complete test system ANTARES (analyser of transverse aberrations of ESO). The information is processed by the computer according to the quasi-Zernike test polynomial shown in figure 3. In the NTT the six low-spatial-frequency aberrations (indicated by †) are the terms corrected by the

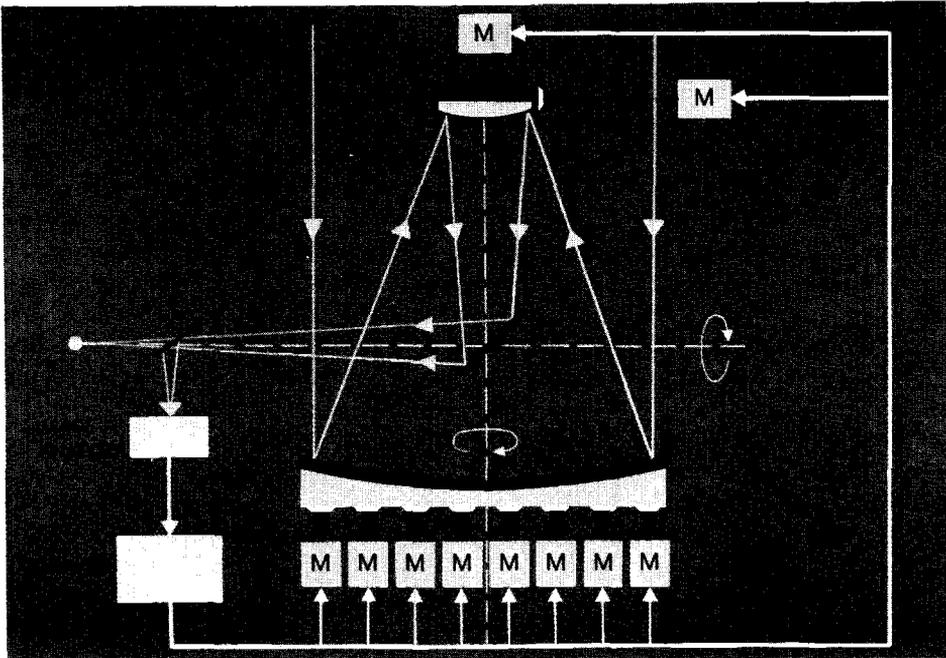


Figure 2. Active optics (for low-frequency band-pass correction): principle of the ESO closed-loop control technique for optimizing image quality.

Wavefront aberration $W = k\rho^m \cos(n\phi + q\theta)$ using:

$W = a$	constant	
$+ b\rho \cos(\phi + \theta_0)$	wavefront tilt (= lateral focus = pointing)	
$+ c\rho^2$	longitudinal focus	
$+ d\rho^3 \cos(\phi + \theta_1)$	decentring coma (3rd order Seidel)†	C
$+ e\rho^4$	3rd order (Seidel) spherical aberration†	S
$+ f\rho^6$	5th order spherical aberration	
$+ g\rho^2 \cos(2\phi + \theta_2)$	3rd order (Seidel) astigmatism†	A
$+ h\rho^3 \cos(3\phi + \theta_3)$	triangular coma†	△
$+ i\rho^4 \cos(4\phi + \theta_4)$	quadratic astigmatism†	□
$+ j\rho^4 \cos(2\phi + \theta_5)$	5th order astigmatism	
$+ k\rho^5 \cos(\phi + \theta_6)$	5th order coma†	

Figure 3. The quasi-Zernike test polynomial used at ESO for off-line telescope testing and on-line active-optics control of image quality.

active optics system. We shall show that it is a consequence of the principle of St. Venant in elasticity theory [21] that correction of the low-spatial-frequency terms is necessary and sufficient to correct those defects arising from elasticity in one form or another (see the table).

Figure 4 demonstrates the convergence due to the principle of St. Venant by a practical example. It shows Hartmann test results for the conventional (passive) ESO 3.6 m telescope produced just after its set-up and adjustment in 1976 [2]. Details are given in [6]. The left-hand point of each plot gives the actual telescope quality (80 per cent geometrical optical energy diameter in arcsec), the other points are fictitious, mathematical values showing what the telescope quality would be if the polynomial terms could be successively corrected. The left-hand part of the functions varies considerably depending on the state of maintenance and attitude of the telescope because the polynomial terms can vary; but the right-hand point is, within error of measurement, an invariant for a given telescope fixed by non-variable higher-spatial-frequency errors left by the manufacturer. We term this the *intrinsic quality* (IQ). In the case of the ESO 3.6 m telescope, the IQ is 80 per cent within 0.27 arcsec.

In the passive 3.6 m telescope, there is no physical means for realizing the correction process shown in figure 4. In the NTT, by contrast, the active control loop enables this to be done and the IQ should always be achieved and maintained (80 per cent within 0.15 arcsec).

4.2. Decentring coma

This is the first and most important term of the correction [9]: it is the curse of Cassegrain telescopes. Active adjustment of the secondary to correct third-order coma avoids the imposition of extremely hard tolerances required for passive performance, which are in practice very rarely met. In the NTT, coma is corrected by rotation of the secondary about its centre of curvature which avoids change of pointing.

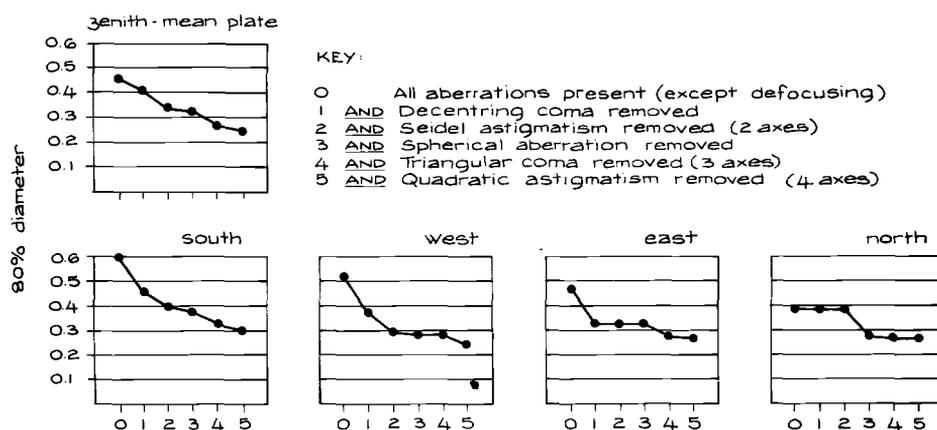


Figure 4. Telescope performance after successive removal of polynomial terms. Results of Hartmann test of the passive ESO 3.6 m telescope in 1976 [2], showing the theoretical improvement that would be attained by correcting low spatial frequency terms if the telescope were active. The mean right-hand point of the functions gives the intrinsic quality (IQ).

4.3. *Other low spatial terms*

All other correctable low-spatial-frequency terms (figure 3) are corrected at the prime-mirror axial support. After third-order coma, the commonest defects are third-order spherical aberration and astigmatism. The former is generated by 'matching' error of the aspherics on primary and secondary, axial position error of the Cassegrain image or of quasi afocal correctors. Astigmatism arises as the lowest mode immediately stimulated by any random support error; or curvature on any oblique plane mirror. Triangular coma arises from loading errors on the fixed points. Fifth-order coma is generated as a residual mirror sag effect in the primary radial support of the NTT.

4.4. *'Force'-based correction on 'soft' supports*

Our active-optics concept is fundamentally based on 'soft' supports. Historically, this is simply the conventional astatic lever support [7, 9]. The essential feature is that no attempt is made to establish the desired surface shape by position control; instead, it is achieved by using the natural elastic-flexure function of the mirror combined with optimum forces measured by load cells. In the NTT we use active astatic double levers (figure 5) but the active-optics principle could also be performed by other support types, e.g. hydraulic or pneumatic. This is a matter of technical convenience, an important aspect in favour of a lever solution being the cosine law of passive support force [9]. The NTT support of figure 5 also contains a fixed-tension spring which applies a constant force determined at telescope set-up for correction of d.c. effects (see the table). These forces might be re-adjusted after long time intervals to correct, say, errors due to the mechanical distortion of mirrors.

As has been emphasized by Citterio (O. Citterio 1985, private communication), active-optics correction could also be performed with a position-based (hard) support, but 'soft' supports seem to have all the advantages for low time frequencies and in particular for modal control.

For the basic (low band-pass) active-optics system, it will be sufficient to perform corrections from time to time [9]. In the NTT, the offset guide star is 'borrowed' for the required integration time of the image analyser (ca. 1 minute). The guide star is then returned to the auto-guider and the correction performed at the secondary and at the prime support. This in no way disturbs the observation: indeed the astronomer will be unaware it is taking place.

4.5. *Quasi-Zernike image analysis: flexure theory*

The basic equation of our test polynomial (figure 3) has the general form of a Fourier equation or of Zernike circle polynomials. The azimuthal orthogonality given by the parameter n is self-evident. The radial polynomials of Zernike are effectively best least-squares-fits of the essentially non-orthogonal individual radial terms in the radius ρ^m , the radial polynomials then also having orthogonal properties. In spite of the attractions of the radial polynomials, we prefer to use the basic (Hamilton) radial terms because these terms have physical significance and origins in practical telescopes, as indicated in §4.3. Another example is focus error: in practice, it is essential to separate focus error from spherical aberration because their correction actively requires different operations.

The analytical theory of mirror flexure has been treated exhaustively by Schwesinger [22, 23]. He uses essentially the same Fourier representation for the

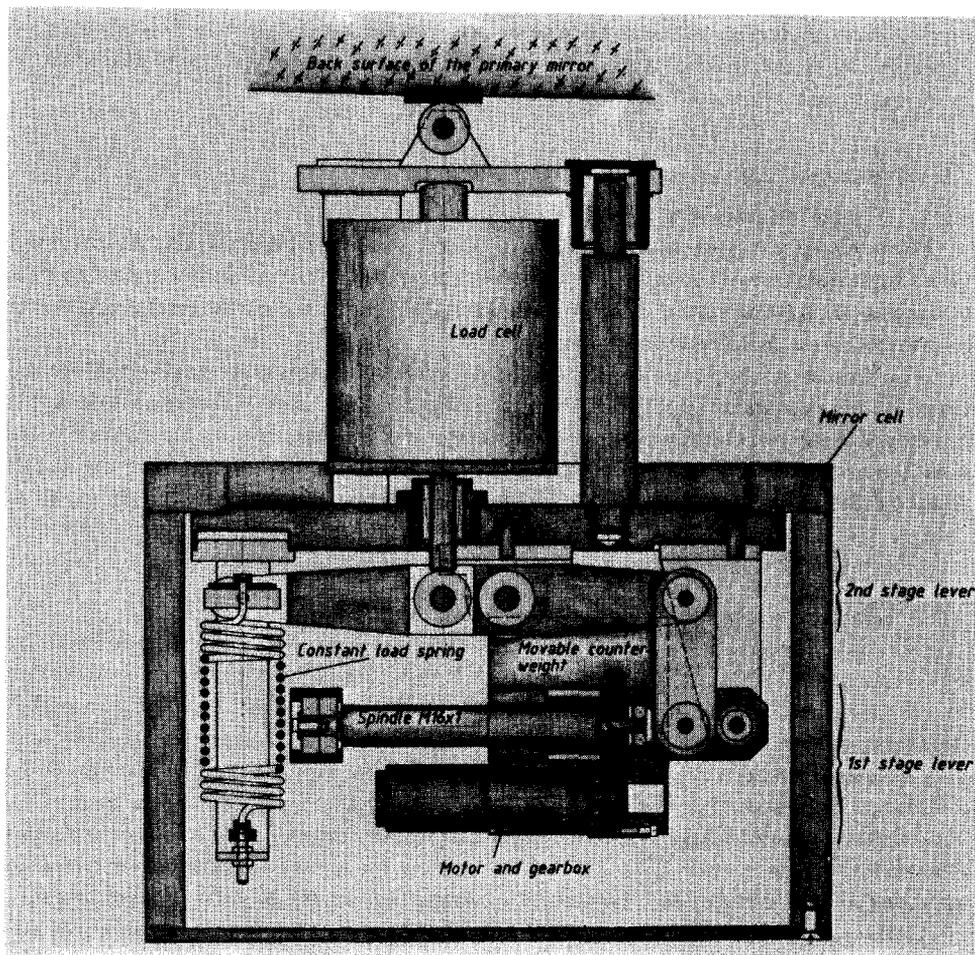


Figure 5. Conceptual drawing of one of the 75 active axial supports of the ESO NTT primary. The d.c. errors are corrected by the springs, other low band-pass errors by the moveable counterweights [9].

analysis of flexure modes. For uncored mirrors, the flexure modes with n -fold rotational symmetry then have the form

$$W_n(\rho) = kc_n(\alpha_n \rho^n + \beta_n \rho^{n+2} + \gamma_n \rho^{n+4} + \delta_n \rho^{n+6} \dots),$$

where k is a function of the density, the Young's modulus and the geometry of the mirror and c_n , α_n , β_n , γ_n and δ_n are dimensionless constants. For mirrors with a central hole the flexure modes W_n also involve functional terms with a singularity at $\rho = 0$, but this in no way invalidates the modal treatment. This confirms that modal control in the general Fourier sense is the natural way of correcting actively all defects originating in elasticity. The fact that optical aberration theory and elasticity theory both make use of a Fourier-type modal formulation is no accident of nature: it originates primarily from the circular symmetry of telescope mirrors and the theory of uni-axial optical systems. It may also be noted that the other theories of the

propagation of electromagnetic radiation, fashionable before the Michelson–Morley experiment, explained the propagation extremely well on the basis of an ether model as an extremely rigid, perfectly elastic solid!

4.6. *Active optics based on three laws of physics*

4.6.1. *The law of linearity (Hooke's Law)*

Passive support systems have always been calculated on the basis of linearity. Glass materials obey Hooke's Law *exactly* right up to fracture. Metals depart from it as they approach the elastic limit, but the dynamic range for active optics will never be more than a tiny fraction of this limit.

The linearity law has two important practical consequences for active optics. First, it allows a linear superposition of the effects of any sets of forces. Secondly, it implies that a given change of force distribution will always produce the same flexure change *independent of the initial state of the force field*, i.e. independent of the initial shape of the mirror. Modes can be controlled independently whatever their initial values.

4.6.2. *The law of convergence (the principle of St. Venant)*

We have referred to this important principle in §4.1. As formulated by Schwesinger [21] it explains the amplitude convergence of the flexure modes with increasing spatial frequency: the higher the spatial frequency of the mode, the higher the forces required to generate a given amplitude. This is intuitively obvious but is of crucial importance in active optics. Beyond a certain frequency, a mode cannot be generated by forces that can either occur naturally in the system or be generated actively in practice. This leads to a simple but important axiom: *if a mirror is flexible enough to develop a given elastic-error mode, then the same error can be corrected by applying active forces of the same order of magnitude as the passive support forces.*

Conversely, if a higher spatial frequency mode can never appear as an elastic error because the forces required are higher than can occur, then it will not be correctable by active optics. This is the case with 'ripple', an error generated by resonance effects in polishing which have nothing to do with elasticity. Active optics can do nothing about ripple: its amplitude must be kept low by the specification to the optician.

4.6.3. *The law of orthogonality*

This is the law of azimuthal orthogonality originating from the Fourier–Zernike polynomial (figure 3). All terms with a different parameter n are mathematically and physically orthogonal and independent and can be controlled independently without 'cross-talk'.

4.7. *Modal calibration (or pre-calculation)*

This feature of our active optics control is absolutely fundamental to it and distinguishes it, we believe, from other systems of active optics that have been proposed [24, 25], even if these use modal control. We consider that the full power and advantage of modal control are only utilized if modal calibration is performed. As will be shown in §5, it leads to a purely diagonal solution matrix and ensures that the dynamic range of correction for a given range of force variation is a maximum.

The principle of modal calibration is extremely simple. Instead of a general matrix-inversion procedure to determine the forces required to correct the total sampled wavefront, the force-variation distribution is determined in advance.

In the case of the NTT, the basic passive-support-system was calculated by Schwesinger and has four support rings. The forces (in newtons) are shown in figure 6. Schwesinger was then asked to determine a force-change distribution which would generate a coefficient of 500 nm for each of the controlled polynomial terms (figure 3) in as pure a form as could reasonably be achieved (this question of accuracy is important and will be discussed in § 5). The results for third-order astigmatism and spherical aberration are shown in figures 7 and 8.

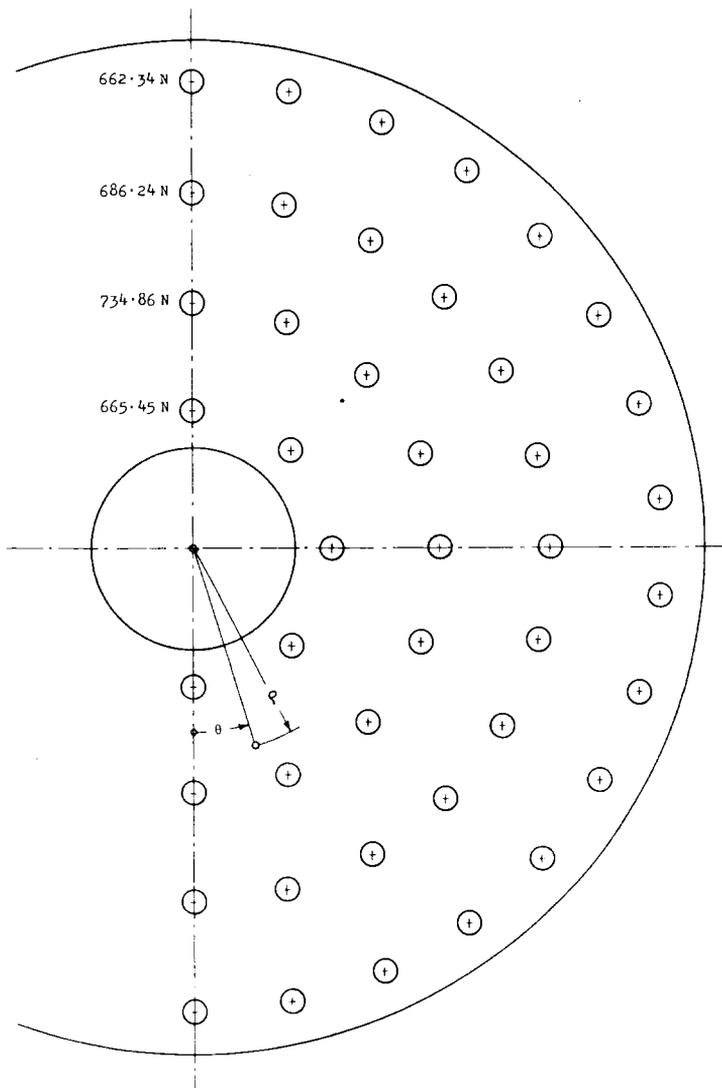


Figure 6. Force distribution (in newtons) for the basic (passive) axial support of the ESO NTT primary (calculated by Schwesinger).

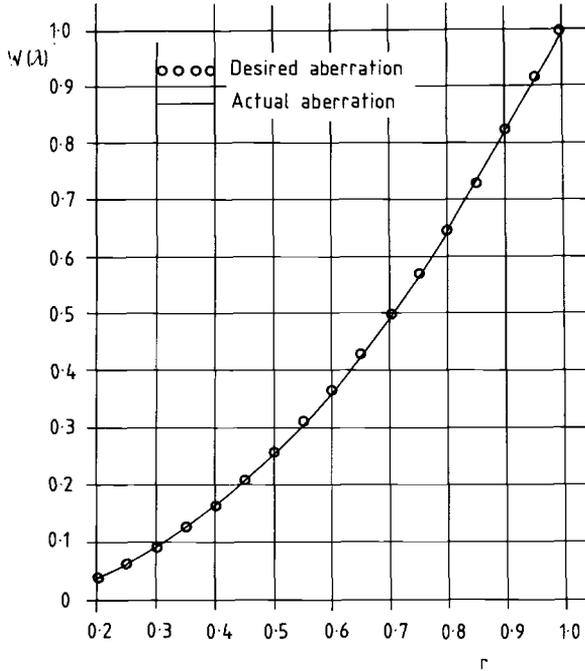


Figure 7. Calibration (pre-calculation) for the active correction of astigmatism in the ESO NTT. Force changes were determined (for the two outer support rings only) giving a change of third-order astigmatism coefficient of one wavelength (500 nm) with very high purity (calculated by Schwesinger). The desired aberration is shown by the open circles, the actual aberration by the solid curve.

Astigmatism requires the lowest forces for the generation of a given wavefront coefficient. It proved possible to generate it with high purity using only the two outer rings, the maximum force change being only of the order of 3 per cent of the mean passive load. Figure 9 shows the required force-change distribution. Spherical aberration is a harder mode to generate, requiring higher forces, the maximum force change being nearly three times greater, 8.8 per cent of the mean passive load. The discrepancies near the axis (figure 8) are negligible in practice for the dynamic range of correction of the NTT. Figure 10 shows the required force change distribution. All other calibrations were easier than spherical aberration.

Following St. Venant, the higher the order of the aberration, the higher are the force changes required to generate a given amplitude [9].

Each of these force-change distributions is stored in the computer. Because of the linearity and orthogonality laws, these force changes can be summed linearly by simple superposition, taking account of the azimuthal phase of the term (θ in figure 3). So the computer operation, using the pre-stored calibrations, becomes trivial: the image analyser determines the coefficient of each term as χ nm, the stored force distributions are simply multiplied by $\chi/500$; then, taking account of azimuthal phase θ , the force changes for the controlled terms are simply added up algebraically. After application of the total force-change distribution, all terms are corrected—the intrinsic quality is achieved.

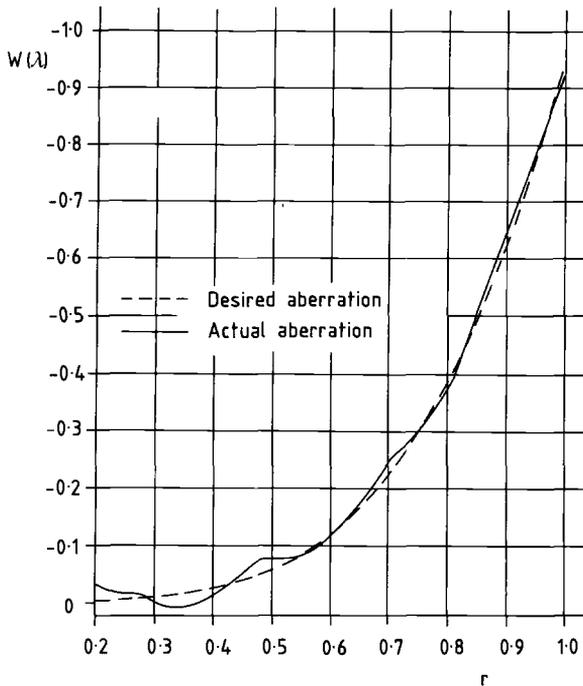


Figure 8. Calibration (pre-calculation) for the active correction of third-order spherical aberration in the ESO NTT. The calibration is analogous to that for astigmatism, but more difficult, requiring force changes on all four support rings and values some three times larger. The discrepancies from the theoretical function near the centre are negligible in practice (calculated by Schwesinger). The desired aberration is shown by the dashed curve, the actual aberration by the solid curve.

Of course, calibrations will not be perfect and will be subject to two types of error: errors of scale and 'cross-talk' to higher radial terms. Such errors have turned out to be remarkably small and are dealt with in detail in Part II of this paper. It suffices to mention here that any scaling error, in comparison with experimental measurements, can be immediately corrected by applying the appropriate scaling factor.

5. General representation and analysis of active modal control

5.1. Wavefront versus modal control

The law of linearity suggests the well-known matrix-inversion solution to the active-optics correction problem at the prime-mirror support. The formulation is particularly simple if a square influence matrix (in this case, stiffness matrix) is set up. For each support a standard force change is applied and the wavefront change established for the same number of sampling points as supports. Then, assuming linearity, we have the unique solution

$$\Delta F_j = A_{ji}^{-1} \cdot \Delta W_i, \quad (1)$$

where ΔF_j is the column vector component of force changes at the actuator points, ΔW_i the column vector component of desired wavefront changes at the sampled (actuator) points, and A_{ji} the component of the stiffness matrix. By modifying this

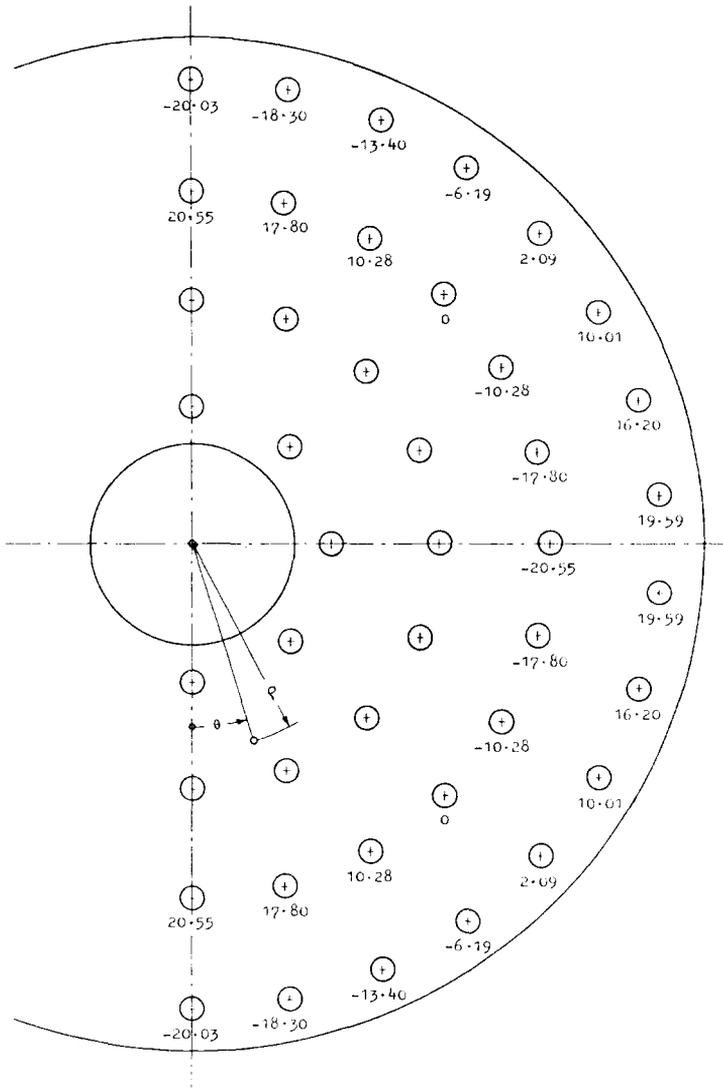


Figure 9. Force changes (in newtons) corresponding to figure 7 for the active correction of astigmatism. Astigmatism is the lowest energy mode: a change of coefficient of 500 nm can be generated by maximum force changes of only about 3% of the mean passive load [9].

basic algorithm, used commonly, for example, in optimization programs for optical design [26], the square stiffness matrix can be made rectangular in either sense (see § 5.3). With this wavefront approach, one would then establish the stiffness matrix in advance. The on-line correction operation would consist of an image analysis followed by a matrix inversion giving the force changes required to reduce the aberrations at the sampled points to zero.

In spite of its mathematical simplicity and apparent elegance which have often led to its proposal, we believe that this method has the following serious disadvantages which are immediately avoided by the modal approach.

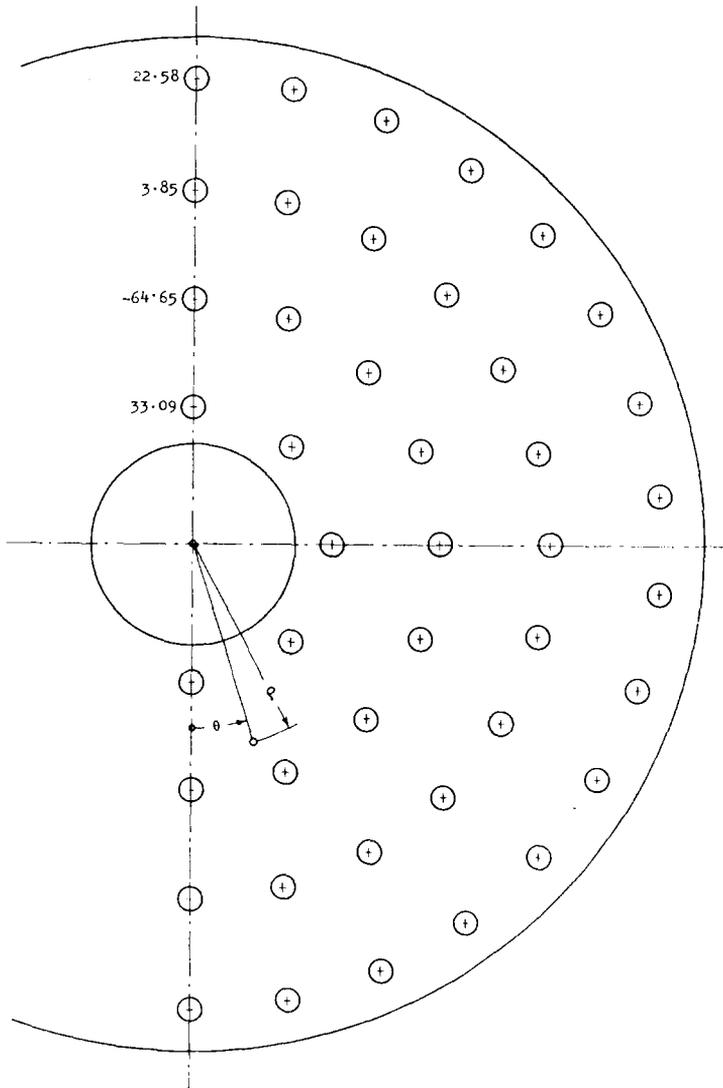


Figure 10. Force changes (in newtons) corresponding to figure 8 for the active correction of spherical aberration. A change of coefficient of 500 nm can be generated by maximum force changes of about 8.8 per cent of the mean passive load [9].

5.1.1. *Ill-conditioned solution matrix*

The requirement of zero aberration targets over an arbitrary mesh (usually rectangular) of sampling points takes no account whatever of the natural flexure modes of the mirror. If the normal case of $i_{\max} > j_{\max}$ (i.e. more sampling points than support points) is taken, spatial frequencies will have to be corrected, from the sampling theorem, up to about half the sampling mesh frequency. Since, without excessive forces, the mirror cannot bend in such high modes and since the support mesh is incapable of generating them, the matrix will become extremely ill-conditioned with extreme eigenvalue ratios. Such phenomena are very well-known in the optical-design field [26]. The result in the present case would be very high

force changes which would limit, in practice, the dynamic range of correction to a fraction of what it could be with modal control.

In extreme cases of two force parameters having virtually the same effect, the matrix solution becomes singular and the force changes infinite. Of course, such a situation can be ameliorated by reducing the sampling or selecting sampling points according to flexure modes. But blindly reduced sampling may not take sufficient account of the sampling information really necessary; and if sampling is chosen for flexure modes it is simpler and more efficient to use a modal method and calibration from the start!

Another way of alleviating the consequences of the ill-conditioning of the inverse stiffness matrix A_{ji}^{-1} is to accept wavefront changes $\Delta\bar{W}_i$ which differ slightly from the theoretically exact desired changes ΔW_i of equation (1), i.e. to use target tolerances. This well-known procedure will relax the requirement to correct exactly all the orders up to that limited by the sampling, with the result that small higher-order wavefront error residuals of negligible effect are accepted. Since the correction of these latter requires high forces, the corresponding force changes

$$\Delta\bar{F}_j = A_{ji}^{-1} \cdot \Delta\bar{W}_i \quad (2)$$

are reduced compared with ΔF_j of equation (1). It should be emphasized that one is still operating in the domain of general wavefront correction using the same basic stiffness matrix A_{ji} . Any gain in dynamic range of equation (2) over equation (1) arises therefore from the improvement produced by multiplying the inverse matrix A_{ji}^{-1} , which depends only on the basic elastic system and not in any way on the wavefront change desired, by a column vector $\Delta\bar{W}_i$ which leads to an easier and nearer solution in the force parameter space. However, the determination of $\Delta\bar{W}_i$ in the general wavefront case may pose a serious problem in practice.

It will be shown in § 5.2 that these basic problems of ill-conditioning of the inverse stiffness matrix A_{ji}^{-1} and in the determination of target tolerances, can be completely solved by modal decomposition and precalibration.

5.1.2. Heavy computer operation, no learning process

It follows from the considerations of the last section that the conventional wavefront, matrix-inversion method starts *every* on-line correction operation with the basic differential information and repeats every time the equivalent of our calibrations without ever learning or re-using the information. Such matrix operations always require considerably more computing than the determination of the coefficients and the simple superposition of our modal method. For very low band-pass corrections, this may not be of major importance; but at the upper end of the active-optics band-pass it becomes crucial to reduce the computing to a minimum. It is also senseless to repeat an operation many thousands of times that can perfectly well be done once, with storage of the result: an efficient process is always a *learning* process to a maximum extent.

5.2. Formal mathematical representation of modal active-optics control

For modal control, the total wavefront correction ΔW can be decomposed into aberration modes:

$$\Delta W(r, \phi) = \sum_{n=0}^{\infty} \sum_{m=n}^{\infty} \Delta W_{nm}(r, \phi). \quad (3)$$

In contrast to ΔW in equations (1) and (2), the ΔW_{nm} , where n and m characterize the aberration mode, are now simple, functionally-defined wavefront aberrations. For each of these aberration modes, the forces at the actuator positions j necessary to produce the desired change ΔW_{nm} at sampling points i could be determined by the matrix inversion

$$(\Delta F_j)_{nm} = A_{ji}^{-1} \cdot (\Delta W_i)_{nm}. \quad (4)$$

As in the general case treated in § 5.1, the stiffness matrix A_{ji}^{-1} only depends on the basic elastic system and not on the desired changes $(\Delta W)_{nm}$. As before, to avoid ill-conditioning and excessive forces, one can and should introduce tolerances on the desired wavefront changes $(\Delta W)_{nm}$, then allowing slightly different wavefront changes $(\Delta \tilde{W})_{nm}$ and thereby accepting negligible higher-order errors. In practice, this relaxation will also include acceptance of changes of lower-order terms in circumstances where these lower-order terms can be controlled by other means, e.g. focus change for a desired change of third-order spherical aberration, or wavefront tilt change and third-order coma change for a desired change in fifth-order coma.

In more general terms, target relaxation for an aberration mode (nm) means deliberately allowing a certain freedom of cross-talk from the basic radial mode m to higher terms ($m+q$), or to some lower terms ($m-q$), which are mathematically and physically not orthogonal. The limits to the extent of this cross-talk are set by the requirement that its effect on the imagery should be negligible. The computation of the force changes with such relaxation could again be achieved by the matrix inversion:

$$(\Delta \tilde{F}_j)_{nm} = A_{ji}^{-1} \cdot (\Delta \tilde{W}_i)_{nm}. \quad (5)$$

As in § 5.1, the stiffness matrix A_{ji} only depends on the general elastic system and not on $(\Delta \tilde{W})_{nm}$, which means that this matrix could be computed for a given system once and for all, off-line and in advance.

Formally, equation (5) represents the process which we term the modal precalibration of forces $(\Delta \tilde{F})_{nm}$ necessary to achieve with sufficient accuracy a desired change $(\Delta \tilde{W})_{nm}$ of a given aberration mode. The low-frequency modes which we wish to correct, such as astigmatism, are known in advance to be modes which are naturally easy to generate. Furthermore, in such modal control, ill-conditioning of the inverse stiffness matrix A_{ji}^{-1} is less critical, since the final product solution of equation (5) is bound to be well-conditioned, particularly if suitable radial-order target relaxation is introduced.

We must now consider how these precalibrations can best be achieved in practice. First, it should be noted that the functionally-defined aberration mode changes depend only on two *scalar* parameters, namely the magnitude ΔC_{nm} of the modal coefficient of $(\Delta W)_{nm}$ and, in the case of rotationally non-symmetric modes, on the orientation θ_{nm} . The actuator forces $(\Delta \tilde{F}_j)_{nm}$ will be proportional to the desired wavefront change $(\Delta \tilde{W})_{nm}$ and the orientation of the force pattern will directly follow the orientation of $(\Delta \tilde{W})_{nm}$, i.e. they will be a simple direct function of the angle θ_{nm} . It follows that the precalibration must derive the forces $(\Delta \tilde{F}^U(\theta=0))_{nm}$ necessary to achieve with sufficient accuracy a modal change $(\Delta \tilde{W}^U)_{nm}$ with a magnitude (coefficient) ΔC_{nm}^U and an orientation of zero degrees. Since the effects are linear, the magnitude ΔC_{nm}^U for precalibration is arbitrary and uncritical and has been fixed at a change of coefficient of 500 nm (1 wavelength) in the polynomial of figure 3.

Precalibration could be done, following equation (5), by the matrix inversion:

$$(\Delta \tilde{F}_j^U(\theta=0))_{nm} = A_{ji}^{-1} \cdot (\Delta \tilde{W}_i^U(\theta=0))_{nm}. \quad (6)$$

But there are other, more practical, methods which can *formally* be written as

$$[\Delta \tilde{F}^U(\theta=0)]_{nm} = [Q]_{nm} \cdot [\Delta \tilde{W}^U(\theta=0)]_{nm}, \quad (7)$$

where $[\Delta \tilde{W}^U]_{nm}$ represents the desired unit wavefront change, $[\Delta \tilde{F}^U]_{nm}$ the unit forces and $[Q]_{nm}$ the 'operator' or method used to derive the forces for the particular mode (nm). In equation (7) $[\Delta \tilde{W}^U]_{nm}$ would be represented by pre-defined values at the sampling points and $[Q]_{nm}$ would be the formal equivalent of the inverted stiffness matrix.

In the case of the NTT, the method used by Schwesinger (G. Schwesinger, 1982, private communication) was the exploitation of vast experience in the application of general elastic-plate theory and of some specific properties of circular plates and shallow shells to derive the precalibration force distributions $[\Delta \tilde{F}^U]_{nm}$. This gives target tolerance relaxation as with the analogous formal procedure of straightforward inversion of the stiffness matrix A_{ji}^{-1} in equation (6) and therefore yields more or less minimum forces, i.e. maximum dynamic range. Examples of the results were given in figures 7 and 8 which show the discrepancies (negligible in practice) which were deliberately accepted to avoid excessive forces (ill-conditioning). Since the process was done completely off-line in advance, and *once only*, the time required for the precalibrations was uncritical.

With the calibration results $(\Delta \tilde{F}_j^U)_{nm}$ stored in the computer, the on-line correction calculation (from linearity and orthogonality) for the aberration terms chosen for the NTT as an example—see figure 3—is simply:

$$\begin{aligned} \Delta F_j = & (\Delta \tilde{F}_j^U)_{04} \cdot \frac{\Delta C_{04}}{\Delta C_{04}^U} && \text{spher} \\ & + (\Delta \tilde{F}_j^U(\theta_{22}))_{22} \cdot \frac{\Delta C_{22}}{\Delta C_{22}^U} && \text{ast} \\ & + (\Delta \tilde{F}_j^U(\theta_{33}))_{33} \cdot \frac{\Delta C_{33}}{\Delta C_{33}^U} && \text{tri} \\ & + (\Delta \tilde{F}_j^U(\theta_{44}))_{44} \cdot \frac{\Delta C_{44}}{\Delta C_{44}^U} && \text{quad} \\ & + (\Delta \tilde{F}_j^U(\theta_{15}))_{15} \cdot \frac{\Delta C_{15}}{\Delta C_{15}^U} && \text{coma5.} \end{aligned} \quad (8)$$

For conformity of notation the aberration coefficients have been represented here by ΔC_{nm} but their significance is identical to those of the coefficients e, g, h, i and k of the polynomial shown in figure 3.

The total elastic correction ΔW_E achieved by applying ΔF_j to the mirror is then

$$\Delta W_E = \Delta W_{04} + \Delta W_{22} + \Delta W_{33} + \Delta W_{44} + \Delta W_{15}. \quad (9)$$

The calibration of ΔW_{15}^U , third-order coma, is achieved in equivalence to equation (7) by a trivial and well-known computation (formally the 'operator' $[Q]_{13}$) in optical design:

$$\Delta \tilde{x}^U(\theta=0) = [Q]_{13} \cdot [\Delta \tilde{W}^U(\theta=0)]_{13}, \quad (10)$$

giving $\Delta\alpha^U$, the rotation of the secondary mirror about its centre of curvature to generate a coefficient of coma³ of 500 nm. Then, again with strict linearity

$$\Delta\alpha = \Delta\alpha^U(\theta_{13}) \cdot \frac{\Delta C_{13}}{\Delta C_{13}^U}, \quad (11)$$

giving the required rotation $\Delta\alpha$ of the secondary from the scalar magnitudes of the precalibrated (unit) coefficient change ΔC_{13}^U and the coefficient change ΔC_{13} to be corrected. This achieves the required modal wavefront correction ΔW_{13} . The total correction is then

$$\Delta W = \Delta W_E + \Delta W_{13}, \quad (12)$$

giving the intrinsic quality as defined in § 4.

The general possibilities of modal decomposition have been studied by Creedon and Lindgren [24] and theoretical and numerical applications of their method have been reported by Howell and Creedon [25] for a thin meniscus with a diameter of 76.2 cm and a thickness of 1.27 cm. Ray and Chang [27] have used a similar approach for their 7.6 m (Texas) mirror. The major difference between this approach and our own is the basic method of optimization of the dynamic range. In the papers cited above, the main emphasis is on finding actuator *locations* which minimize cross-talk to higher-order bending modes or, equivalently, on optimizing the stiffness matrix to control more efficiently a certain set of aberration modes.

In contrast, our approach is that we fix the actuator locations *from the start* according to the requirements of the passive support system, the actuators of the NTT supports then being located on four rings in such a way that they carry similar loads. This effectively predetermines the stiffness matrix A_{ji} which may be highly ill-conditioned, since it is of a quite general nature (as indicated by equation (1)) and implies nothing in itself regarding a modal or any other solution principle. The ill-conditioning of the total system is removed, as has been shown above, by optimizing the aberration target correction conditions rather than the actuator locations.

5.3. The general problem of modal calibration

So long as a telescope mirror can be reasonably well modelled by a circular solid plate or shell, remarkably accurate calibrations (see Part II) can be obtained with relatively little effort, particularly in the case of the plate model (a good approximation for the NTT). If a mirror meniscus is thin (like our 1 m test mirror, see § 6), the 'shell effect' due to its curvature may become significant in the cases of spherical aberration and coma. These cases were therefore treated by shell theory, although the results obtained by plate theory were not far from being acceptable.

The problem of calibration becomes much more difficult with structured (lightweighted) blanks. Calibration then requires an equivalent model which can be used with analytical means. According to Schwesinger, such a model can only be established if the real flexure characteristics are known for certain cases which give then the basic parameters of the model.

The obvious approach is by finite element (FE) calculations. Such calculations are now common for support systems for structured mirror blanks [28]. Interesting attempts at modal calibration by FE for meniscus blanks using a stiffness matrix have been made by Ballio *et al.* ([29], G. Ballio, O. Citterio, R. Contro and C. Poggi 1984, private communication). Although, after discussions and corrections, reasonable

agreement with analytical results was obtained, it was revealed that such calculations are very difficult and require great care with definition of elements and boundary conditions. Above all, there is the danger of reducing the dynamic range by over-constraining the system: in certain modes, the FE calculations had little more than one third of the dynamic range of the analytical calibrations. Attempts with structured blanks only confirm this since they are much more difficult. The computing effort is also formidable.

In connection with the FE modal calibration attempts for the NTT 1 m test mirror [29], it was suggested by Ballio (G. Ballio, 1984, private communication) that the stiffness matrix for that mirror could readily be determined by us *experimentally*. The need to do this was removed at the time by the satisfactory agreement between analytical theory (Schwesinger) and FE calculations (Ballio). But the experimental stiffness matrix for modal calibration seems to us one of the most promising directions for modal calibration of structured blanks.

Reference was made in § 5.1 above to the application of optical-design algorithms for the inversion of rectangular stiffness matrices. The most generally used algorithm in optical design is the so-called 'damped-least-squares' (DLS) algorithm [26] which is applicable to a rectangular stiffness matrix A_{ji} in which the number of conditions i_p (wavefront sampling points in the Shack-Hartmann measurement) is greater than the number of force points j_p . The general solution is then

$$(\Delta F) = [(A)^t(A) + K(I)]^{-1}(A^t)(\Delta W), \quad (13)$$

where $(A)^t$ is the transpose of (A) , (I) the unit matrix and K is a damping factor introduced to reduce the solution range because of nonlinearities of aberration functions [26]. It also reduces the danger of singularities occurring in the basic least-squares matrix $[(A)^t(A)]^{-1}$. In the present case, which is strictly linear, it is unclear to what extent damping is needed.

If a thin meniscus solution were chosen for an 8 m mirror, it may be that the number of supports j_p could exceed the number of wavefront sampling points i_p . In this case, the Glatzel 'adaptive' algorithm [26] would be more appropriate:

$$(\Delta F) = (A)^t[(A)(A)^t]^{-1}(\Delta W). \quad (14)$$

This would allow, with linearity, each condition to be reduced to its target in one iteration, more efficiently than the least-squares reduction of DLS. Furthermore, by setting tolerance limits for the targets, as is done in optical design with this algorithm, a controlled freedom could be introduced into the calibration to avoid higher-order constraints which would limit the dynamic range of the correction system.

6. Aims of the 1 m active-optics experiment

The results of this experiment are given in detail in Part II of this paper. A brief review of the most important aspects has also been given by Noethe *et al.* [10]. Here, we will only give a brief introduction to the origins and purpose of the experiment.

The test mirror for the NTT has an aperture of 1.05 m, a thickness of 18.9 mm and an aspect ratio of 56 which is little more than a tenth of that of a conventional blank. It was scaled according to the gravity flexure law

$$W = k \frac{D^4}{d^2},$$

where D is the diameter and d the thickness, from the 3.5 m NTT blank, a meniscus with aspect ratio of 15. With this scaling, the passive support system maintains its geometry and all forces are scaled proportionally to the relative weights of the NTT and 1 m blanks.

A spherical form was chosen for simplicity of manufacture and test at the centre of curvature in autocollimation using the ANTARES (Shack–Hartmann) image analyser.

Relatively generous tolerances were given to the manufacturer (REOSC of Paris) in respect of the low spatial-frequency errors, since these were to be corrected afterwards actively. In terms of the polynomial coefficients in nm, these tolerances and the achieved quality were as follows, whereby the \pm sign only implies the cosine effect of doubling the wavefront aberration:

<i>Aberration</i>	<i>Tolerance coefficient</i>	<i>Achieved quality</i>
spher	1500	330
ast	± 1500	± 1160
tri	± 250	± 165
quad	± 125	± 285

Since the first three aberrations were all well inside tolerance, the achieved value of quad could be accepted, even though it was outside the formal tolerance. The dynamic range of motorised variation in the experiment was limited (for purely technical reasons) to ± 10 per cent of the mean passive load (much less than in the NTT itself [9]), although a hand adjustment enables an initial bigger change to be made in order to bring the quality within the automatic range.

As will be clear from the discussions in §4 above, the main aim of the experiment was to test the validity and precision of the calibrations. Of course, a demonstration of orthogonality and linear superposition was psychologically also necessary, although the validity of these laws could not be in doubt. Once the accuracy of calibration and orthogonality had been proven, the active correction of the 1 m mirror to reduce all the controlled terms to zero was a trivial operation (see Part II).

As given in Part II, the test results exceeded our expectations on the precision of the analytical calibrations of Schwesinger and the orthogonality (cross-talk) was within the precision of the Shack–Hartmann measurements.

With these results, we consider that the active-optics system presented above is a proven, practical system. The only limits to its use in any given case are the limits to getting reliable modal calibrations. We believe the methods outlined in §5.3 will provide the solution to this, even in difficult cases of large structured mirrors.

7. Conclusions

To summarize, the most important advantages of the active-optics system outlined are:

It allows the relaxation of optical and structural tolerances.

It allows thinner, lighter and more flexible mirrors, thereby reducing weight throughout the whole telescope.

Reduction of the mass of mirrors automatically reduces their thermal capacity, an important advantage for the thermal behaviour of the telescope or optical system.

It allows consideration of new materials for mirrors which may be subject to some warping.

The above four features should greatly reduce costs.

It allows the 'Intrinsic Quality' (effectively diffraction limited) of a telescope to be realized for low band-pass errors all the time, but without demanding this quality of the optician.

It allows enormous simplification of optical maintenance (the telescope maintains itself).

Because of the principle of modal calibration, the upper limit of the band-pass set by processing of image analyser data is much higher than if laborious matrix inversion or other calculation-intensive algorithms were used. The precision of correction is limited solely by the precision of the image analysis.

The hardware aspect of the active-optics control involves only application of known technology.

The most important application should be in space optics [30] where only active optics is needed for diffraction-limited systems (no adaptive optics needed). For very light and highly flexible space optics, the dynamic range of correction will be a fundamental parameter of active control and will determine the required purity of modal calibrations. Similar considerations apply to the ESO VLT [31].

For ground-based work, the quality achievable by active optics will force progress in adaptive optics (high band-pass), although the problems are immense by comparison and the adaptive correction field is always limited by the isoplanatic angle.

8. Acknowledgments

We owe a deep debt of gratitude to our consultant, G. Schwesinger, without whose major contributions the active-optics development would not have taken place; also to O. Citterio and his colleagues for the most successful design and manufacture of the support of the 1 m test mirror, and to the firm REOSC for excellent work and cooperation on the manufacture of the 1 m test mirror itself.

We are grateful for calculations by G. Ballio and colleagues and discussions with them and O. Citterio which were important for the calibrations. Our thanks are due to many colleagues in ESO who have made major contributions, particularly M. Cullum, D. Enard, P. Giordano, F. Merkle, S. Milligan, M. Moresmau, M. Schneermann and M. Tarenghi. We also acknowledge stimulating discussions with A. Meinel, R. Tull and F. Ray.

References

- [1] WILSON, R. N., 1977, *ESO Conference on Optical Telescopes of the Future*, Geneva (Geneva: ESO), p. 99.
- [2] FRANZA, F., LE LUYER, M., and WILSON, R. N., 1977, *3.6 m Telescope: The Adjustment and Test on the Sky of the PF Optics with the Gascoigne Plate Correctors*, ESO Technical Report No. 8.
- [3] TARENCHI, M., 1986, *Advanced Technology Optical Telescopes III*, Tucson, SPIE, Vol. 628 (Bellingham: SPIE), p. 213.

- [4] ENARD, D., 1986, *Advanced Technology Optical Telescopes III*, Tucson, SPIE, Vol. 628 (Bellingham: SPIE), p. 221.
- [5] WILSON, R. N., and MISCHUNG, K. N., 1987, *Astron. Astrophys.* (submitted).
- [6] WILSON, R. N., 1982, *Optica Acta*, **29**, 985.
- [7] FRANZA, F., and WILSON, R. N., 1982, *Advanced Technology Optical Telescopes*, Tucson, SPIE, Vol. 332 (Bellingham: SPIE), p. 90.
- [8] WILSON, R. N., 1983, *Proc. Workshop on ESO's Very Large Telescope*, Garching (Garching: ESO), p. 173.
- [9] WILSON, R. N., FRANZA, F., and NOETHE, L., 1984, *Proc. International Astronomical Union (IAU) Colloquium No. 79: Very Large Telescopes, their Instrumentation and Programs*, Garching (Garching: ESO), p. 23.
- [10] NOETHE, L., FRANZA, F., GIORDANO, P., and WILSON, R. N., 1986, *Advanced Technology Optical Telescopes III*, Tucson, SPIE, Vol. 628 (Bellingham: SPIE), p. 285.
- [11] WOOLF, N. J., and ULICH, B. L., 1984, *Proc. ESO Workshop on Site Testing for Future Large Telescopes*, La Silla, 1983 (Garching: ESO), p. 163.
- [12] FRIED, D. L., 1966, *J. opt. Soc. Am.*, **56**, 1372.
- [13] RODDIER, F., 1984, *Proc. ESO Workshop on Site Testing for Future Large Telescopes*, La Silla, 1983 (Garching: ESO), p. 193.
- [14] ENARD, D., MISCHUNG, K. N., SCHNEERMANN, M., and WILSON, R. N., 1986, *Advanced Technology Optical Telescopes III*, Tucson, SPIE, Vol. 628 (Bellingham: SPIE), p. 518.
- [15] PEARSON, J. E., FREEMAN, R. H., and REYNOLDS, H. C., 1979, *Applied Optics and Optical Engineering* (New York: Academic Press), Vol. 7, p. 262.
- [16] WOOLF, N. J., 1984, *Proc. IAU Colloquium No. 79: Very Large Telescopes, their Instrumentation and Programs*, Garching (Garching: ESO), p. 221.
- [17] MERKLE, F., 1986, ESO VLT Report No. 47.
- [18] FOY, R., and LABEYRIE, A., 1985, *Astron. Astrophys.*, **152**, L29.
- [19] PLATT, B., and SHACK, R. V., 1971, *Optical Sciences Center Newsletter* (University of Arizona), Vol. 5 (1), p. 15.
- [20] WILSON, R. N., FRANZA, F., NOETHE, L., and TARENGHI, M., 1984, *Proc. IAU Colloquium No. 79: Very Large Telescopes, their Instrumentation and Programs*, Garching (Garching: ESO), p. 119.
- [21] SCHWESINGER, G., 1968, *Proc. Symposium on Support and Testing of Large Astronomical Mirrors*, Tucson (Tucson: Kitt Peak National Observatory and University of Arizona), p. 11.
- [22] SCHWESINGER, G., 1954, *J. opt. Soc. Am.*, **44**, 417.
- [23] SCHWESINGER, G., 1969, *Astronom. J.*, **74**, 1243.
- [24] CREEDON, J. F., and LINDGREN, A. G., 1970, *Automatica*, **6**, 643.
- [25] HOWELL, W. E., and CREEDON, J. F., 1973, NASA Technical Note D-7090 (January 1973).
- [26] GLATZEL, E., and WILSON, R. N., 1967, *Appl. Optics*, **7**, 265.
- [27] RAY, F. B., and CHANG, J. H., 1985, *Appl. Optics*, **24**, 564.
- [28] RAY, F. B., and CHANG, J. H., 1986, *Advanced Technology Optical Telescopes III*, Tucson, SPIE, Vol. 628 (Bellingham: SPIE), p. 526.
- [29] BALLIO, G., CONTRO, R., POGGI, C., and CITTERIO, O., 1984, *Proc. IAU Colloquium No. 79: Very Large Telescopes, their Instrumentation and Programs*, Garching (Garching: ESO), p. 75.
- [30] BARKER, A. J., and CUTTER, M. A., 1985, SIRA European Space Agency Contract Report, Sira ref.: A/7123/00 October 1985, pp. 4, 58.
- [31] WILSON, R. N., 1986, *Second Workshop on ESO's Very Large Telescope*, Venice, ESO, Proc. No. 24 (Garching: ESO), p. 417.