Stellar Spectra and the "Virtual Observatory" Project

R. Elizabeth Griffin
Oxford, England

Present address: Dominion Astrophysical Observatory, Victoria, BC, Canada

Abstract Spectroscopy is the essential tool for investigating the nature of stellar atmospheres of all types of stars, standard as well as variable; monitoring the spectrum of a variable star provides vital information about the causes of observed photometric variability. Astrophysics, therefore, needs accessible data banks of spectra to complement those of photometric observations such as are held by the AAVSO. The IAU Working Group for Spectroscopic Data Archives is working towards the establishment of such data-banks from two approaches: (1) the collection and dissemination of newly-acquired digital spectra, and (2) the creation of a world plate store (the "Virtual Observatory") to facilitate the re-use of older photographic spectrograms.

1. Introduction

This meeting is focussing on "new frontiers" for collections of astronomical photometry, and I find it extremely heartening that an octogenerian organization like the AAVSO is able to approach the technological future with increasing vigor and freshness. The topic of this session is "variable objects and sky surveys," and although my contribution will have no obvious involvement with either variable objects or sky surveys, that is largely because spectroscopists have not yet gotten their act together in such a comprehensive and powerful way as has the AAVSO, and have yet to organize even the preliminaries. Nevertheless, "new frontiers" involve expansion—of opportunity, of scope, and of volumes—and expansion cannot be maintained without increased cooperation. It is certain that an effort to improve the present feeble condition of spectroscopic databases will generate research tools that have immediate relevance to photometry and, moreover, will soon become indispensable for it and mutually supportive of it.

The differences in the natural roles of photometry and spectroscopy can be summed up quite simply by saying that photometry describes *what* is happening while spectroscopy shows *how* things may be happening; it takes a combination of the two in order to explain *why*. The complementarity of the different approaches is also obvious from their products: it may take one hundred photometric observations to produce one paper, while one spectrogram (or, as Dr. Friedjung interjected from the audience, even just one spectral line) can generate one hundred papers. In this contribution I will not be addressing questions of *why*; that is the simple realm of astrophysics. I will concentrate on the much thornier problems of organizing and handling spectroscopic data, and will discuss the extent to which the spectroscopic data that are so essential to modern research are being, or could be, made accessible in useable forms by and for the astronomical community. My talk is therefore about archiving spectroscopic data.

2. Basic definitions

Before I launch into details, it is important to establish a set of uniform definitions of strategic words like "data," "archive," and "information," and also to reflect for a moment on the scientific processes which take part in the flow along the "corridor of wisdom" from photons to physics—see Figure 1. I actually visualize the corridor of wisdom as a physical entity occupying an activity space between telescopes at one end and localities of research at the other. One's exposure to this corridor is normally in the reverse direction from that taken by the photons, since one is likely to commence by hearing or reading the finished product, wisdom (= knowledge + experience), and later pursue investigations back to the sources of knowledge (catalogues), to their sources of information (data banks), or perhaps finally acquire new data (unadulterated photons).

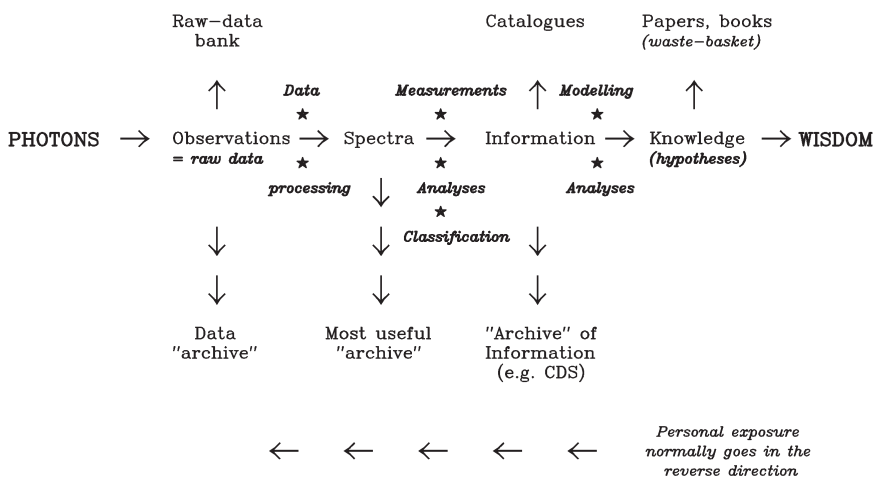

Figure 1. The photon-to-physics pipeline—the "Corridor of Wisdom."

One's first passage along the corridor may convey the impression that the links are all in place and that things work well; however, closer looks reveal uncomfortable gaps and ad-hoc solutions, and further probing shows that there is no coordinated concern, no managerial policies, and little investment of resources except in a few specific cases. Confusion also abounds, because every different activity within the corridor will have been branded "archiving" in some context or other, whilst everyone is also conscious of the civilian use of the term and its strong connotations with things put away in remote places undisturbed for long periods. It is therefore all the more important that we use the term from the outset to describe a processes of recalling (physically) observations associated with one's research with much the same ease and flexibility as one recalls (mentally) associated facts, history, and knowledge. As researchers, we have unfortunately grown accustomed to expect inertia when trying to recall previous spectroscopic observations, and have, as a direct consequence, been content to allow our research techniques to flourish in modes that avoid any need to oil rusty wheels or brush aside cobwebs. Access to adequate amounts of observing time seems also to have taught some astronomers to equate "old" with "inferior." However, all astronomical objects are varying on some time-scale (secularly, if not obviously periodically, too), and no one can deny the importance of rapid and easy access to historic (i.e., past) observations that span complete periods of variability; no one would knowingly discard relevant data, whatever its characteristics. In any case, while individual spectra may contain time-sensitive information that could be irreplaceable, together they also build up a body of information that increasingly assumes the responsibility of knowledge.

3. History of spectroscopic archives

A spectrogram is an extensive array of data-points, and carries the advantage that it can be reworked, remeasured, or reanalyzed for whatever reason. Photographic spectrograms have always been regarded as the property of the observatory where they were taken, and throughout the first three-quarters of this century nearly all spectra were recorded on photographic plates, and were diligently catalogued by plate librarians and stored in plate vaults along with log books and occasionally with card-index catalogues as well. Then the digital era introduced the electronic camera, the reticon, and finally the CCD, and for about twenty years ideas outran technology; as a consequence, the task of collecting up and saving digital spectroscopic observations, having started out as difficult, became ignored by habit. Whenever the question of "saving" observations was raised, people complained that it was too complicated and too costly in resources, that it was quicker and easier to make observations afresh, and that most researchers would not trust other people's observations anyway. The private observatories simplified the situation by holding the view that the data belonged to the observer and were to be taken away right out of sight, please, at the end of an observing run.

By about the late 1980s technology had evolved sufficiently that things were starting to change. As influential observatories became multi-national public organizations, so it was recognized that their data resided in public ownership and had to be treated respectfully as such. It also became clear that space observatories could in no way afford to lose sight of their data. The acute expense of their operations drove to the fore a powerful new economic factor: a space mission had a limited lifetime which should not be squandered on repeating what was already in its database if there was no scientific purpose in duplication. There was also the question of instrument calibration: the observations from a space mission comprised an indispensable measurement of technical performance. Thus the modern teams set the example in manners for their forebears to follow at their leisure (though there seemed little incentive to hurry until new technological developments drove them to do so).

After several false attempts, the International Astronomical Union (IAU) eventually passed a resolution in 1991 that archiving spectroscopic observations should be regarded as a serious duty, and set up a Working Group in 1992 with dual responsibilities: (1) to interpret the Resolution as best it could by encouraging the formation of archives of observations with a view to working towards collections of fully reduced spectra, and (2) to do something (but without specifying what) about the world's heritage of photographic spectra.

Members of the Working Group have encouraged, explained, and persisted in arguing that it does make good scientific sense to archive spectra. The arguments are fairly obvious, and their main points—scientific, technical, and economic—can be summarized as follows:

Scientific:

(i) All astronomical objects vary (quickly, slowly, periodically, secularly).

(ii) Better analyses use multi-frequency data (x-ray, UV, optical, radio...).

(iii) Statistical analyses are much more meaningful from larger samples.

Technical:

Extensive data-sets, especially of standards and calibrations, supply unique information on the performance of an instrument.

Economic:

(i) "Expensive" observations (e.g., from space) should not be repeated unnecessarily if suitable observations have already been made (even if by someone else).

(ii) The fact that observations are able to be used to the full improves the accountability of publicly-funded instruments in the eyes of the public.

4. Digital spectra

During its five years of existence, the Working Group has witnessed important changes in world opinion towards matters associated with the collection of new (digital) data, and now sees the need to include planning into its brief; this is because the thawing of previously frozen opinions through only a small rise in temperature is threatening certain inundation and subsequent chaos unless adequate plans are laid before the events rather than after them. Multi-object spectrographs and robotic telescopes are just two developments that promise to submerge astronomers in their *tsunami* of raw data. Experimentation with improved data-control techniques are being undertaken, and the design of new telescopes such as the European Southern Observatory Very Large Telescope (ESO VLT) and the Italian Galileo Telescope are putting resources into simulating the data flows and creating data-handling routines that will become part and parcel of the instruments' real-time activities.

The next question to be faced centers on the formats which reusable spectra should take (i.e., raw images or reduced spectra). The arguments in favor of the contents of an archive being *reduced* (i.e., fully calibrated) spectra are both scientific (procuring the best possible reductions through expert consensus) and economic (saving time). Again, it took the shining example of a space mission (the International Ultraviolet Explorer Uniform Low Dispersion Archive or IUE ULDA) to start convincing those more resistant to change. The "Saving the Bits" schemes that have been initiated are without doubt a magnificent start, but are not the end of the road; data-banks of raw images which require specific knowledge for their correct usage and demand inordinate effort from the analyst in every individual case are scarcely research-friendly. If resources were infinite, much more effort would surely be directed into the efficient conversion of spectral images into physically-meaningful forms; but they are not, and instead quite a scandalous number of research person-years is currently spent in repetitive, piecemeal extraction of spectra from images, and in the absence of much properly coordinated information about a given instrument, it has to be doubted whether all of that time is well spent. Curiously, that economic argument seems to be less powerful because it is somehow less direct.

However, the generation of some degree of automation for converting observed images into spectra ("pipeline reduction") is becoming more urgent now owing to different forces: the mounting complexities of instruments and of the modes in which they can be operated. Not surprisingly, initiatives for streamlining spectrum extraction tend to emanate from research projects which cannot otherwise flourish; Doppler-image tomography from rapid spectroscopy is but one such case in which the sheer volumes of complicated spectra to be analyzed are driving the development of highly automated data-reduction techniques. But these processes can—and should—be generalized, and when the Working Group recently gave careful and unbiased deliberation to the question, "Is it feasible to install pipeline reduction software at existing as well as at future telescopes and instruments?" It concluded that, yes, it is indeed feasible provided certain important criteria in the realms of data acquisition and handling are met.

The ultimate goal of this part of the Working Group's brief is therefore a set of inter-linked, fully reduced, instantly-accessible spectra. The peripheral requirements include a master catalogue that one can consult (rather like one can browse SIMBAD); the catalogue will have pointers to available photometry and to lists of spectra, along with their vital statistics (date, wavelength span, resolution, etc.). It will be possible to "view" selected spectra, and of course to request transfers of the relevant data files. The organization of the "world" spectroscopic archive will almost certainly be that of a distributed database; the merged catalogue will be made available centrally, but the creation and maintenance of each individual archive will be handled as close to its source instrument as possible.

5. Photographic spectra

Whilst these plans are being developed to enhance our ability to use and re-use spectroscopic observations acquired by digital detectors, the photographic plates that contain the world's heritage of spectra up to the 1970s or beyond are just collecting dust—literally, in some cases. Or they are languishing in filing drawers in offices, constituting "a lifelong extension to one's brain." There are about thirty-five plate stores around the world in just about every country where astronomical research flourished during the first half of this century, and most are now "closed" in the sense that they belong to an obsolete technology and are no longer the official responsibility of an observatory staff member. Few of the card-index catalogues have been transferred online, so to consult and possibly use such spectra (however important they prove to be for research) means that one has to visit the plate stores personally, be familiar with the characteristics of the plates found, and be in possession of the hardware, the software, and the expertise to extract (preferably digitally) the required information as correctly as possible. Clearly, the chances of being able to realize the potential of these spectra will decline from their present lowly rates to almost zero unless action is taken soon whilst experience in both digital and photographic techniques can be tapped symbiotically.

The Working Group's deliberated solution is to gather all simple spectroscopic plates into a World Plate Store, catalogue and arrange them, and commence digitizing them in order of scientific priority. The site chosen for this activity is the Haute Provence Observatory (OHP) in southeastern France, and letters of agreement to the principle have already been exchanged between the IAU and the Director of OHP. The existence of a physical location for a staffed archive of world photographic spectra also offers a natural headquarters for supervising the world archive of digital spectra described above, and the whole entity constitutes what I am labelling *The Virtual Observatory*.

Important practical questions remain, because whereas one can argue convincingly that the generation by an observatory of its own reduced spectra and the maintenance of accessible collections of them is a local responsibility whose costs should be absorbed as such, a world relocation for a collection of erstwhile distributed pieces is an isolated enterprise that must seek its own special funding and stand on its own feet financially. The level of funding is not in fact altogether astronomical, since we do not have to purchase land or construct buildings; we require modest start-up resources (for dispatch of plates, store-room refurbishment, and clerical assistance in sorting and cataloguing), plus running costs for clerical and technical assistance on an ascending scale: quite simply, the more funds we can raise, the more (and more quickly) we can generate a meaningful tool out of these historical resources. I recognize that the nature of the funding we seek places it on the fringes of the remit of national or trans-national agencies, and we know from experience how difficult it is to attract resources into supporting a service to *facilitate* scientific results when in strong competition with projects that tackle the science itself.

6. Outlook

On the positive side, the pendulum is swinging surely, if slowly, in the direction of collaborations, surveys, and databases. At the same time, our ability to carry out increasingly sophisticated measurements in astrometry, photometry, spectroscopy, and interferometry on bright stars is beginning to highlight some of the glaring gaps in our basic astrophysical knowledge. Unfortunately, the deleterious effects of those gaps and errors have spread unchecked, since the same astrophysical deductions are applied by extrapolation to situations in which objects are too faint for the associated laws to be verified accurately. It goes without saying that any improvements in the basic knowledge will therefore have a beneficial knock-on effect throughout astrophysics, but it does need to be said that it requires painstaking investigations of the spectra of the brighter stars to make good many of those deficiencies. It has also to be pointed out again that archives of stellar spectra constitute a supremely important, not to say unique, resource for such research. Wider acceptance of these truths should enhance the overall support for the types of world-wide effort described above.

7. Addendum, 2006

During the ensuing eight years since this symposium, several fundamental changes have been implemented to the original project design. The concept of a World Plate Store at Haute-Provence Observatory in France proved to be unsupportable. Inspection of the premises indicated that the buildings that could perhaps be used were not in fact sufficient, and would require more fundamental upgrade and maintenance than was proportional to the cost of the science to be gained.

A modified plan introduced the concept of a Scanning Laboratory, and to that end room space has been offered (and accepted) at the Dominion Astrophysical Observatory (DAO) in Victoria, British Columbia, Canada. The basic plan of the Scanning Laboratory is to install about three Photometric Data Systems microdensitometers (PDS), suitably upgraded, and to select and borrow, for scanning, plates from observatories that are willing to lend them on that basis for a short time. Two PDSs have already been offered.

Meanwhile, PARI (the Pisgah Astronomical Research Institute), a new institute in North Carolina, is designing a project to house, store and eventually digitize direct and objective-prism plates within its abundant storage space; it can also become the long-term repository for spectrograms. Efforts are continuing to raise funds for that community service. PARI has also become a shelter for endangered plates, and rescued the Michigan plate archive when it faced annihilation. In Europe, a tentative agreement has been reached with the Royal Observatory, Belgium (Brussels) to create a store there for European collections of plates that are no longer wanted by the parent Observatory. Progress is dependent upon the completion of a pilot project, involving the construction of a rapid and very accurate new scanner, to digitize part of the Observatory's own plates.

A pilot project to illustrate the value of historic stellar spectra has also been completed at the DAO: the development and confirmation of a method to measure the strength of the Earth's ozone from the UV Huggins bands, which are superimposed on all ground-based UV stellar spectra, at various past dates and geographic latitudes. The project was described in *PASP* **117**, 885, 2005 by R. E. M. Griffin, "The Detection and Measurement of Telluric Ozone From Stellar Spectra," and follow-up studies have been published in atmospheric-science journals. The ozone project demonstrated a high credibility for the results, and illustrated the undeniable value to mankind of researching historic astronomical data to those ends. Evidence of past ozone concentrations has largely been restricted to one European site, with only patchy contributions from other sites, so the input from astronomy is unique and will be highly prized.

It should be pointed out that the title of "Virtual Observatory" given to this project in 1998 was believed to be the first formal use of the phrase in Astronomy! However, because the title was not protected through publication at the time, the same name was adopted by the U.S.-born project to federate current major databases, and was also transferred to offspring organizations such as the "National Virtual Observatory," "European Virtual Observatory," "Astrophysical Virtual Observatory." In deference to mightier power we have had to rename our scanning laboratory the "Spectroscopic Virtual Observatory" (SVO).

8. Acknowledgements

I would like to thank NATO for awarding and renewing a grant (CRG.921015) in support of the activities of the IAU Working Group.