LETTERS

Chemo-dynamical modelling with Schwarzschild's method

Richard John Long^{1,2} and Shude Mao^{3,1,2}

¹ National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; rjl@nao.cas.cn

² Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

³ Physics Department and Tsinghua Centre for Astrophysics, Tsinghua University, Beijing 100084, China

Received 2018 September 3; accepted 2018 October 16

Abstract We extend Schwarzschild's dynamical modelling method to model absorption line strength data as well as the more usual luminosity and kinematic data. Our approach draws on earlier published work by the first author with the Syer & Tremaine made-to-measure (M2M) dynamical modelling method and uses similar ideas to create a chemo-Schwarzschild method. We apply our extended Schwarzschild method to the same four early type galaxies (NGC 1248, NGC 3838, NGC 4452, NGC 4551) as the chemo-M2M work, and are able to recover successfully the 2D absorption line strength for the three lines we model (H β , Fe5015, Mg b). We believe that this is the first time Schwarzschild's method has been used in this way. The techniques developed can be applied to modelling other aspects of galaxies, for example age and metallicity data coming from stellar population modelling, and are not specific to absorption line strength data.

Key words: galaxies: abundances — galaxies: formation — galaxies: individual (NGC 1248, NGC 3838, NGC 4452, NGC 4551) — galaxies: kinematics and dynamics — galaxies: structure — methods: numerical

1 INTRODUCTION

In Long (2016), it was stated that "a galaxy's construction and evolution are imprinted in its kinematics and chemistry but require significant analysis to identify the contributing componentry". The paper then demonstrated successfully how the made-to-measure (M2M) method proposed by Syer & Tremaine (1996) for modelling stellar dynamical systems could be extended to model spectral absorption line strength data. Within the paper, it was noted that a similar approach to chemo-dynamical modelling may be possible using Schwarzschild (1979) modelling. We are pleased to report that we have now developed such a modelling method by a simple extension to Schwarzschild's method. This letter draws heavily on Long (2016) and should be considered as an extension to it. We believe that this is the first time in which line strength data have been employed in Schwarzschild modelling.

Schwarzschild's method pre-dates the M2M method by almost 20 years and its application to galaxies is more extensive and includes those described for the M2M method in Long (2016). Now that we have a chemo-dynamical Schwarzschild's method, we have a further tool to examine and analyse data from integralfield unit (IFU)-based galaxy surveys (for example, ATLAS^{3D}, Cappellari et al. 2011; SAMI, Bryant et al. 2015; MaNGA, Bundy et al. 2015), hopefully adding to our knowledge of the componentry underlying galaxies and how it was assembled.

As might be expected, our objectives in performing the current investigation mirror those in Long (2016) and are

to extend Schwarzschild's method to model absorption line strength data as well as the more usual luminosity and kinematic data, and to create a software implementation of the revised method,

- to apply the method to the same selection of external galaxies as Long (2016) and confirm that the criteria for a successful Schwarzschild model can be met (orbits weighted and observables reproduced), and
- to understand the limitations of the extensions and to identify areas for future work.

The structure of the paper broadly follows the objectives. In Section 2, we describe our enhanced Schwarzschild's method. In Sections 3 and 4, we apply the method to four galaxies taken from the ATLAS^{3D} survey and summarise our results respectively. In Sections 5 and 6, we discuss the results and draw conclusions identifying areas for further investigation respectively.

2 SCHWARZSCHILD'S METHOD

Schwarzschild's method was first described in Schwarzschild (1979). The method is concerned with weighting orbits in such a way that observations of a stellar system may be reproduced. Whereas in the M2M method particle weights are adjusted as the particles are being orbited, in Schwarszschild's method the orbit weights are calculated only after orbit creation has been completed. The method has been enhanced and applied by many researchers since 1979 (for example, Rix et al. 1997, Valluri et al. 2004, Cappellari et al. 2006, van den Bosch et al. 2008), and is very well described in those and other papers. Binney & Tremaine (2008) contains an overview. We therefore do not describe the full details of the method here but only give sufficient details so that it is clear how line strength data can be modelled.

In Section 2.1, we frame Schwarzschild's method in terms of matrices as a convex optimisation problem, and then in Section 2.2, using the same terminology, describe how line strength data can be handled. In essence, for every orbit, we introduce an additional value per orbit for each absorption line to be modelled.

2.1 Basic Theory - Orbit Weights

Calculation of the orbit weights in Schwarzschild's method is achieved using a 'least squares' approach to minimise the residuals between the model observables and the measured observables of a stellar system. Given that linear least squares is just a subset of convex function optimisation theory (Boyd & Vandenberghe 2004), we choose to think of the minimisation in those terms. In

matrix form we minimise

$$\|Dw - K\|_2^2$$
, (1)

where D is the 'design' matrix giving individual orbit contributions to model observables, w represents the orbit weights to be determined, and K contains the 'measured' observables. The L2-norm (Euclidean norm) is indicated by $\|\dots\|_2$. For our purposes, K is taken to contain both kinematic and luminosity measurements. Whether or not luminosity measurements should be taken as range constraints within the minimisation (for example, van den Bosch et al. 2008, Valluri et al. 2004) does not affect our arguments. Given that individual orbit weights should not be negative, the least squares algorithm required must be able to generate a non-negative least squares result (see Chen & Plemmons 2009 for techniques) and typically the implementation used is from Lawson & Hanson (1974). The orbit weights should sum to 1 and so D and K are adjusted to include a sum of weights constraint (a row of 1s in D with the corresponding element of K set to 1).

The regularised form of expression (1) that we use is

$$\|Dw - K\|_{2}^{2} + \lambda \|w\|_{2}^{2},$$
 (2)

where λ is a user tunable parameter. This quadratic form of Tikhonov regularisation (Tikhonov 1963) maintains the convex form of the expression to be minimised, and acts by suppressing high valued weights (see, for example, Vasiliev 2013 or Valluri et al. 2004). In this letter, as will be seen in Section 4, we will use it to increase the number of active orbits in our models and to improve the orbit weight distribution.

As a final point in this subsection, we turn the minimisation into a χ^2 minimisation by dividing the data elements in K by their errors and similarly the corresponding rows in D.

2.2 Extensions for Modelling Spectral Line Strength Data

In the previous section, we covered how orbit weights are generated by weighting model observables to match the measured observations. In this section, we generate the line strength values to be associated with each orbit by asking what model orbit contributions given the orbit weighting will enable the measured values to be matched. In other words, the roles of the orbit weights and the model's orbit contributions are swapped by comparison with expression (1), with the role of the fractional mass terms being unchanged. Again we will use a nonnegative least squares / convex optimisation approach, this time seeking to minimise

$$\|Cx - S\|_{2}^{2} + \lambda_{\rm LS} \|x\|_{2}^{2}, \qquad (3)$$

where C is the 'design' matrix giving the individual orbit weightings including the fractional mass terms, x represents the orbit line strength values to be determined, Scontains the target line strength values for a given spectral line, and $\lambda_{\rm LS}$ is the regularisation parameter which may be zero. Only active orbits (orbits with non-zero weights) are included in the design matrix. Note that different expressions (3) as above exist for each line to be modelled. Considering the summation constraint, for the spatial region that we have line strength data, we require that the model sum of orbit line strength values equals the sum of the input line strengths with C and S being adjusted accordingly.

Note that, even though we position this letter on modelling line strength data, the above techniques will work on other numerical attributes of orbits provided they are linear in superposition. Modelling age and metallicity from stellar population analyses could be handled in this way. Logarithmic data can also be modelled without conversion provided it is acceptable that values will be geometric means, not arithmetic means (the arithmetic mean of log data is in fact the log of the geometric mean of the data).

2.3 Software Implementation

The Python M2M implementation used in Long (2016) has been extended to perform Schwarzschild modelling as well. This approach means that we can take advantage of existing M2M software concerned with handling initial conditions, gravitational potentials, binning schemes, orbit integration and creation, and parallelisation. We are able to model surface brightness, luminosity density, mean line of sight velocity and mean line of sight velocity squared using Voronoi cells for binning kinematic data. Modelling with Gauss-Hermite coefficients of the line of sight velocity distribution (see Rix et al. 1997) has been implemented and used but is not reported on in this letter. Dithering of orbits is not currently implemented. Ignoring data preparation, execution of our software takes place in two stages. The first is concerned with orbit creation and collecting orbit data contributing to calculating model observables (the columns of the design matrices D and C). The second is concerned

with performing the convex optimisations with a variety of different methods and implementations, and different levels of regularisation. These different methods include non-negative least squares (CVXOPT¹ and Lawson & Hanson (1974) implementations), bounded variable least squares (Python *scipy* implementation), and the sequential coordinate-wise algorithm (Franc et al. 2005). Our preferred optimisation software is CVXOPT which has a parallelised linear algebra package. Note that the Lawson and Hanson algorithm can be parallelised - see, for example, Luo & Duraiswami (2011).

3 APPLICATION TO EXTERNAL GALAXIES

We use the same ATLAS^{3D} data² and surface brightness multi-Gaussian expansions (MGEs, see Emsellem et al. 1994) for the same galaxies as the chemo-M2M activity in Long (2016). Considering the constraining observables, the luminosity constraints are the same as the M2M work but, for the kinematic constraints, we substitute mean velocity squared for the velocity dispersion constraint. We model the same three spectral lines (H β , Fe5015 and Mg *b*) using both symmetrised and unsymmetrised line strength data. Relevant galaxy properties are shown in Table 1. If it is needed, the total number of observational constraints contributing to the orbit weights can be determined as 2 × the number of IFU points from Table 1 + 512 (for the luminosity constraints).

We create axisymmetric Schwarzschild models of our galaxies using the MGEs noted above in creating gravitational potentials. Initial conditions for the orbits are the same as for the chemo-M2M activity and include the same circularity adjustment for S0 galaxies to create more circular orbits. We use 8000 undithered orbits per galaxy. Units for modelling are effective radii for length, 10^7 years for time, and mass in units of the solar mass M_{\odot} with luminosity similarly so. Line strength data values are in Angstrom.

Both the summation constraints and regularisation involve manually tunable parameters and we use the following parameter values determined by experimentation. For the summation constraints, the values are 10^3 for orbit weights, and 1.0 for orbit line strength contributions. For regularisation, for orbit weights they are 8×10^{-2} for NGC 1248 and 2×10^{-1} for the other galaxies, and for line strength 10^{-5} for all four galaxies.

¹ http://cvxopt.org/

² http://purl.org/atlas3d

Galaxy	Morphology	Inclination	M/L Ratio	IFU data points
NGC 1248	SO	42°	2.50	297
NGC 3838	S0	79°	4.00	383
NGC 4452	SO	88°	5.20	489
NGC 4551	Е	63°	4.89	596

Table 1 Galaxy Properties

Notes: Galaxies and their properties relevant to our Schwarzschild models. These are the same as used in Long (2016) for chemo-M2M modelling.

Galaxy	SB	LD	\bar{v}	$\bar{v^2}$	${\rm H}\beta$	Mg b	Fe5015
Without regularisation							
NGC 1248	0.062	0.173	0.030	0.080	0.398	0.498	0.497
NGC 3838	0.087	0.365	0.023	0.037	0.399	0.174	0.337
NGC 4452	0.071	0.024	0.040	0.087	0.219	0.267	0.362
NGC 4551	0.040	0.075	0.032	0.065	0.436	0.267	0.380
With regularisation							
NGC 1248	0.063	0.296	0.059	0.318	0.193	0.185	0.807
NGC 3838	0.178	1.007	0.144	0.369	0.307	0.128	0.340
NGC 4452	0.436	0.117	0.189	0.286	0.207	0.194	0.428
NGC 4551	0.147	0.313	0.068	0.356	0.301	0.215	0.569

Table 2 CVXOPT - Mean χ^2 Values

Notes: Mean χ^2 values calculated using CVXOPT with and without regularisation. SB indicates surface brightness and LD, luminosity density.

4 RESULTS

It is not practical to show all results for all galaxies using figures. We therefore focus on a single galaxy NGC 4452 with figures but display results for the other galaxies in tabular form only. Unless stated otherwise, all analyses are conducted using CVXOPT.

4.1 Initial Models

Our initial models for the four galaxies do not use regularisation in the weight determinations. We show the mean χ^2 values we achieve in Table 2 top rows, and plots specific to NGC 4452 in Figure 1. It should be clear from the table and figure that we are able to model line strength data as well as luminosity and kinematic data. All the data in these initial models have been symmetrised.

We examine quite how the enhanced modelling scheme is behaving. The sum of weights constraints are met with no problems encountered. For example, the model sum of orbit weights is 1.00. As can be seen from Figure 2 (left panel) for NGC 4452, the orbit weight distribution shows that larger orbit weights (99% of total weight) are associated with a small fraction of orbits (10% of orbits) and that many orbits (90%) have low orbit weights. This issue is not specific to NGC 4452 and applies to the other three galaxies as well.

If we now include regularisation in our models, as can be seen from Figure 2 (right panel), the imbalance in the left panel is significantly reduced and the number of active orbits increases considerably (from 10% to 84%). The mean χ^2 values we achieve are shown in Table 2 bottom rows. As before, the sum of weights constraints are met.

4.2 Minimisation Methods

In Section 4.1, we used the CVXOPT software to analyse our models. This software uses what is known as an 'interior point' method to determine the orbit weights and line strength contributions. The Lawson & Hanson (1974) NNLS method uses an 'active set' method and deliberately zeroises weights and contributions for orbits it decides it derives no benefit from using. For convex function minimisation, if a solution exists, it is unique (this comes directly from the definition of a convex function). This means that, to within some numerical accuracy, the CVXOPT software and the Lawson & Hanson (1974) NNLS method should be returning the same val-



Fig. 1 NGC 4452 line strength and kinematic plots from chemo-Schwarzschild modelling. Units are as per Section 3. Coordinates X and Y give on-sky positions in effective radii.



Fig. 2 NGC 4452 orbit weight distribution with and without regularisation. The left panel (no regularisation) shows that larger orbit weights (99% of total weight - *black line*) are associated with a small number of orbits (10% - *blue line*) and that many orbits (90% - *blue line*) have low orbit weights. The black line indicates the fractional number of orbits per logarithmic weight bin and the blue line, the fractional weight contained in each bin. Using regularisation (*right panel*), the imbalance is reduced with more than 99% of the total orbit weight distributed across 84% of the orbits.



Fig. 3 NGC 4452 orbit weight distribution and line strength contribution comparison between CVXOPT (*top row*) and Lawson & Hanson (1974) (*bottom row*). The black line indicates the fractional (weight or line strength) distribution.

ues for the weights and contributions. We re-analyse our models using Lawson & Hanson (1974) and compare the results obtained with the CVXOPT results. The Lawson & Hanson (1974) mean χ^2 values are in Table 3. To three decimal places, the luminosity and kinematic χ^2 values are the same as those in Table 2 while the line strength values differ due to different numbers of active orbits.

We examine the numbers of orbits contributing to the χ^2 values and show the results in Table 4. From the table, for Lawson & Hanson (1974), the impact of regularisation on increasing the number of active orbits is quite clear. It is also clear that even though the line strength data are being reproduced (see Table 3), not all the active orbits are contributing to all the lines. If now we look at a comparison of the orbit weights and line strength contributions between CVXOPT and Lawson & Hanson (1974) for NGC 4452, we see that the orbit weights have the same distribution (Fig. 3). Although the line strength profiles are similar, they are not exactly the same. This has potential implications on just what the weighted orbits might be used for: results may be specific to the underlying techniques used (i.e. CVXOPT vs. Lawson & Hanson NNLS) and may not be totally representative of the stellar system being modelled.

4.3 Data Symmetrisation

Since Long (2016) makes the point that chemo-M2M does not perform well with unsymmetrised data, we have evaluated how our chemo-Schwarzschild method performs. We use symmetrised kinematic data with unsymmetrised line strength data, and our results are shown

in Table 5. By comparison with the symmetrised regularised models in Table 2, it is quite clear that unsymmetrised data do not yield acceptable models. This may limit the applicability of our approach and is discussed further in Section 5. We have not yet attempted to quantify asymmetry but simple mechanisms such as axis reflection are straightforward to implement. More sophisticated methods using, for example, the Radon transform (Stark et al. 2018) or symmetry pattern recognition techniques should also be considered.

5 DISCUSSION

From our results in Section 4, it is clear that we are able to model successfully symmetrised 2D line strength data using our extended Schwarzschild's method. However, none of the issues identified in the chemo-M2M work in Long (2016) is able to be resolved by using a chemo-Schwarzschild's method such as developed here. We have not attempted to consider 3D aspects of modelling with Schwarzschild's method (for example, the 3D distribution of orbit line strength values): the same, previously identified concerns of uniqueness and plausibility apply, and the follow up work anticipated in Long (2016) concerning the likely robustness of any 3D predictions is not complete. It now needs extending to include chemo-Schwarzschild modelling together with possible variations arising from different convex optimisation methods. In addition, the need to symmetrise data remains and cannot be resolved by the approach we have taken. Modelling of asymmetric data remains an outstanding issue. Perhaps chemo-Schwarzschild and chemo-M2M

Galaxy	SB	LD	\bar{v}	\bar{v}^2	${ m H}eta$	$\operatorname{Mg} b$	Fe5015
Without regularisation							
NGC 1248	0.062	0.173	0.030	0.080	0.370	0.463	0.474
NGC 3838	0.087	0.365	0.023	0.037	0.362	0.131	0.300
NGC 4452	0.071	0.024	0.040	0.087	0.220	0.220	0.309
NGC 4551	0.040	0.075	0.032	0.065	0.429	0.264	0.354
With regularisation							
NGC 1248	0.063	0.296	0.059	0.318	0.125	0.115	0.654
NGC 3838	0.178	1.007	0.144	0.369	0.235	0.075	0.255
NGC 4452	0.436	0.117	0.189	0.286	0.191	0.170	0.381
NGC 4551	0.147	0.313	0.068	0.356	0.226	0.135	0.454

Table 3 Lawson & Hanson 1974 - Mean χ^2 Values

Notes: Mean χ^2 values calculated using Lawson & Hanson (1974) with and without regularisation. The luminosity and kinematic values agree with those obtained using CVXOPT - see Table 2. The line strength values are not due to differing numbers of active orbits.

		CVXOPT	Lawson & Hanson NNLS			
Galaxy	# orbits	# active	# active	${ m H}eta$	$\operatorname{Mg} b$	Fe5015
NGC 1248	7994	7994	587	156	143	131
		6420	4521	3383	2216	406
NGC 3838	7982	7942	786	221	269	196
		7453	5824	4524	5296	3582
NGC 4452	7997	7997	781	246	287	239
		7623	6818	5712	6252	6178
NGC 4551	7979	7979	940	298	385	310
		7973	6496	4700	6030	3239

Table 4 Active Orbits Comparison

Notes: The numbers of active orbits contributing to the orbit weight calculation and the orbit line strength contributions. For each galaxy, the top row is without regularisation, and the bottom row is with regularisation. The # active columns give the number of orbits with non-zero weights. The spectral line columns give the number of orbits with non-zero contributions to the model line strength values. For CVXOPT, the number of spectral line orbits is the same as the number of active orbits and so the CVXOPT columns are not shown. For Lawson & Hanson (1974), the impact of regularisation on increasing the number of active orbits is quite clear.

should only be applied to early type galaxies with thoroughly mixed stellar populations until it is addressed.

Using Schwarzschild's method as the core modelling method does bring a new set of issues as well as identifying some additional concerns with M2M. The lack of visibility of the impact of using the Lawson & Hanson (1974) method without regularisation is a major concern: the low number of active orbits is rarely documented in journal papers. Based on our mean χ^2 results, Schwarzschild's method does seem to overfit the constraining observable data. Perhaps all data (not just luminosity data) should be modelled as range constraints where the range is set using the error on the observed data. Note also that the Schwarzschild extensions do contain an implicit assumption that the orbit weights generated from modelling kinematics (with or without regularisation) are also suitable for modelling line strength data. This needs confirmation or otherwise by modelling a larger set of galaxies.

The extent to which results achieved by both the chemo-Schwarzschild and chemo-M2M approaches are influenced by the initial orbit or particle conditions requires further investigation. Both approaches can only weight what they are provided with initially. It is quite possible to vary the orbit mix between circular, radial and box orbits, for example, and achieve a number of plausible models reproducing the observed data. Again, more research is required.

6 CONCLUSIONS

We have met the objectives we set out in the Introduction, Section 1. We have extended Schwarzschild's method into a chemo-dynamical method which is able to han-

Table 5 Unsymmetrised Line Strength Data - Mean χ^2 Values

(Galaxy	${ m H}eta$	Mg b	Fe5015
1	NGC 1248	2.51	1.12	7.47
]	NGC 3838	3.05	1.51	2.57
1	NGC 4452	1.76	1.22	2.77
1	NGC 4551	2.03	1.35	2.40

Notes: Mean χ^2 values resulting from using unsymmetrised line strength data. By comparison with the regularised models in Table 2, the values are significantly higher (> 1) implying that modelling performs better with symmetrisation.

dle luminosity, kinematic and absorption line constraints, and have successfully applied the extended method to four ATLAS^{3D} early type galaxies. However, as anticipated in the objectives, and as can be seen from the Discussion, Section 5, much remains to be investigated to understand the limitations of the current approach and possible alternatives before robust predictions on real galaxies can be made. This makes the need for follow up investigations a priority. Overall, notwithstanding our reservations, we believe another promising step has been taken in developing a capability to perform chemo-dynamical modelling. We now have chemo-Schwarzschild to add to our original chemo-M2M. For the future, we plan to apply these chemo-methods to additional observational data sets, such as MaNGA, to try and gain a deeper insight into the relationships between chemistry and kinematics within galaxies.

Acknowledgements Computer runs were performed on the *Venus* high performance computer system at the Tsinghua Centre for Astrophysics, Tsinghua University, China. This work is partly supported by the National Key Basic Research and Development Program of China (No. 2018YFA0404501 to SM), and by the National Natural Science Foundation of China (Grant Nos. 11333003, 11390372 and 11761131004 to SM).

References

- Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition (Princeton Univ. Press)
- Boyd, S., & Vandenberghe, L. 2004, Convex Optimization (Cambridge: Cambridge Univ. Press)
- Bryant, J. J., Owers, M. S., Robotham, A. S. G., et al. 2015, MNRAS, 447, 2857
- Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, ApJ, 798, 7
- Cappellari, M., Bacon, R., Bureau, M., et al. 2006, MNRAS, 366, 1126
- Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, MNRAS, 413, 813
- Chen, D., & Plemmons, R. J. 2009, in Symposium on the Birth of Numerical Analysis, World Scientific, eds. A. Bultheel, & R. Cools (World Scientific Press)
- Emsellem, E., Monnet, G., & Bacon, R. 1994, A&A, 285, 723
- Franc, V., Hlavac, V., & Navara, M. 2005, in Computer Analysis of Images and Patterns. CAIP 2005. Lecture Notes in Computer Science, 3691, eds. A. Gagalowicz, & W. Philips (Springer, Berlin, Heidelberg)
- Lawson, C. L., & Hanson, R. J. 1974, Solving Least Squares Problems (Prentice-Hall Series in Automatic Computation, Englewood Cliffs: Prentice-Hall, 1974)
- Long, R. J. 2016, RAA (Research in Astronomy and Astrophysics), 16, 189
- Luo, Y., & Duraiswami, R. 2011, SIAM J. Sci. Computing, 33(5), 2848
- Rix, H.-W., de Zeeuw, P. T., Cretton, N., van der Marel, R. P., & Carollo, C. M. 1997, ApJ, 488, 702
- Schwarzschild, M. 1979, ApJ, 232, 236
- Stark, D. V., Bundy, K. A., Westfall, K., et al. 2018, MNRAS, 480, 2217
- Syer, D., & Tremaine, S. 1996, MNRAS, 282, 223
- Tikhonov, A. N. 1963, Soviet Mathematics, 4, 1035
- Valluri, M., Merritt, D., & Emsellem, E. 2004, ApJ, 602, 66
- van den Bosch, R. C. E., van de Ven, G., Verolme, E. K.,
- Cappellari, M., & de Zeeuw, P. T. 2008, MNRAS, 385, 647 Vasiliev, E. 2013, MNRAS, 434, 3174