A Numerical Method For Recovering The Original Design Parameters Of A Sundial[1]

Gianpiero Casalegno (Castellamonte, Italy)

When studying or restoring old dials it becomes necessary to retrieve some of the original dial lines that could have been heavily deteriorated by time. This process requires that the original design parameters of the dial are computed. As an alternative to classical methods that are described in the literature, this article suggests a different numerical algorithm that could be applied to any type of dial and hour line.

Introduction

Sooner or later every dialist will face the problem of computing the original design parameters of a sundial. This requirement could simply come from the curiosity of knowing the declination of the wall on which the dial is placed or verifying the type of hour system that is used in the dial (*e.g.* normal or *da campanile* italic hours). A scholar of history can be interested to find the original latitude of a portable dial and then to discover the construction place and maybe the name of the manufacturer. A professional restorer often needs to recover the original lines of an old dial, starting from a few weak traces, in order to be able to restore the dial to its original shape.

In all these situations the problem is solved if the following dial parameters can be found:
- latitude and longitude of the place
- declination and inclination of the wall
- position and length of the style.

Depending on the situation and on the type of dial, some of these parameters could be not applicable (*e.g.* longitude for hour lines that only depend on local time such as italic, temporary *etc.* or declination and inclination for a horizontal dial) or they could be already known (*e.g.* a hole in the wall can be the sign of a style that was placed in that position).

Classical approach

The following logical flow is followed when the unknown parameters of a dial are looked for:
- all the useful elements in the dial are identified: equinox line, meridian line, points of intersection with the hour lines *etc*.
- related angles and distances are measured
- the mathematical formulae giving as a result these parameters starting from the unknown dial parameters are identified
- these formulae are inverted in order to obtain the unknown parameters starting from the measured values.

For instance it is well known that in a vertical declining dial the slope μ of the equinox line can be computed from the latitude φ and the wall declination d as:

$$\tan \mu = \sin d / \tan \varphi$$

It is thus immediate to compute the wall declination when the latitude φ is known and the angle μ is measured[2].

---

[1] This paper was originally released as "Un metodo numerico per il recupero dei parametri costruttivi originali di un quadrante" in the Italian publication "Orologi Solari" , Year 5 , n. 16 , November 2008 , pp. 26-32.

[2] Wall declination could of course be obtained by means of one of the several measurement methods that are available in the literature, but actually what is here needed is the original declination value that was used by the dialist as it could differ from the true one (because of the movement of the dial from the original place or because of a measurement error by the dialist himself).

It is also possible for each case to find the trigonometric formulae that allow one to compute the unknown parameters starting from the measured geometric elements. (see [1] for the analysis of some specific situations).

This way of proceeding, although the only one available when just a hand calculator or a slide-rule was available and it was impossible to solve complicated mathematical problems, suffers from two drawbacks.

The first inconvenience is procedural: the method has to be adapted to each different faced situation (dial type, available lines …) and related formulae have to be found and inverted in order to come to a simple and fast solution to the problem.

The second inconvenience is substantial: the error that is unavoidably introduced when measuring a geometric element (angle, position, distance) has a direct effect on the computed parameter and then, with an avalanche effect, on all the following unknown parameters.

In other words: only one geometric element is used at a time, while we could take advantage of all the information that can be extracted from the dial in order to reduce, with some sort of averaging process, the resulting error in the results.

For instance in the previous example the wall declination is directly obtained from the slope of the equinox line, while the declination is of course related to all the dial's geometric elements and from them all it could actually be extracted.

Alternative approach

Considerable computation power is today available for everyone and it is therefore possible to deal with this problem using a different technique.

Let us consider a directional sundial on a vertical declining wall. This assumption is taken in order to fix the details of the problem and to simplify the following explanation. However the described method can be used for any dial type (azimuth, elevation ...) on any type of surface, not necessarily planar.

The x/y coordinates (X_i,Y_i) of a characteristic point i of a dial (*e.g.* the intersection between an hour line and a day line) can be computed by means of known non linear equations that depend on the following parameters:

ω_i	= hour angle for point i	δ_i	= sun declination for point i
φ	= latitude	λ	= longitude
d	= wall declination	L	= orthostyle length
x, y	= coordinates of the orthostyle		

Suppose that the positions of N/2 points[3] are measured on the dial, each of them corresponding to known values of sun declination δ_i and hour angle ω_i. The following system of N equations in 6 variables can then be written:

$$\begin{cases} X(\omega_i , \delta_i , \varphi , \lambda, d , L , x, y) = X_i & i = 1 \div N/2 \\ Y(\omega_i , \delta_i , \varphi , \lambda, d , L , x, y) = Y_i & i = 1 \div N/2 \end{cases}$$

where coefficients X_i and Y_i are the coordinates of the N/2 points as measured in an arbitrary x/y reference system.

By solving this system of equations, the 6 unknown values for φ, λ, d, L, x and y could be found.

[3] It can happen, when no day lines nor the meridian line are traced, that only straight lines are available in the dial without any intersection point. The method here described can still be used if some dial points, even if not present on the dial, are forced to lie on the measured line, or if the computed and the measured slopes are forced to be equal.

If X_i & Y_i are moved to the first side and a vector notation is used, the system reduces to

$$\underline{\varepsilon}(\underline{p}) = 0$$

where :

$\underline{\varepsilon}$ = $\{\varepsilon_1, \varepsilon_2 \ldots \varepsilon_N\}$ is the vector containing the errors between computed [X() , Y()] and measured $[X_i\ ,\ Y_i]$ coordinates

$\underline{p}$ = $\{p_1, p_2 \ldots p_M\}$ is the vector containing the M unknowns (in this case we have M=6 and $\underline{p}$={ φ, λ, d, L, x, y})

Although N=M equations are enough in order to find the desired solution (the vector $\underline{p}$ that brings to zero the vector $\underline{\varepsilon}$) it is convenient, as said before, to make use of as much information as possible in order to average any error introduced in the measurements. It is so suggested to have N >> M using the largest possible number of measured points X_i & Y_i and to solve the system looking for the vector $\underline{p}$* that can minimize "in some way" the vector $\underline{\varepsilon}$ (and so minimize the errors ε_i between computed and measured points).

One of the possible measurements of the size of vector $\underline{\varepsilon}$ is its Euclidean squared norm that can be computed as :

$$E(\underline{p}) = \left\|\underline{\varepsilon}(\underline{p})\right\|^2 = \varepsilon_1^2 + \varepsilon_2^2 + \ldots \varepsilon_N^2$$

that is also equal to N times the mean squared value of the errors ε_i. The solution to the system is thus the vector $\underline{p}$* that minimizes the function E($\underline{p}$).

As a summary: we are looking for the parameters that would provide a set of dial lines as close as possible to the measured points, or in other words that would give the minimum possible value for the mean squared value of the errors ε_i between computed and measured coordinates.

The algorithm

Several different methods exist for finding a solution to this kind of problem (see for example [2]). The results that are shown in the following paragraphs have been obtained with the Levenberg-Marquardt method [3] that is now briefly described.

This is an iterative method: at step i a new estimate $\underline{p}_{i+1}$ of the solution is computed from the previous estimate $\underline{p}_i$:

$$\underline{p}_{i+1} = \underline{p}_i + \underline{\Delta}_p \qquad (1)$$

where $\underline{\Delta}_p$ is a vector such that the value of the function E($\underline{p}$) is reduced.

The vector $\underline{\Delta}_p$ is computed by solving the following linear system

$$(J^T J + \mu I)\, \underline{\Delta}_p = - J^T \underline{\varepsilon} \qquad (2)$$

where J is the Jacobian matrix $(J(\underline{p}))_{ij} = \partial\varepsilon_i/\partial p_j$. I is the identity matrix and $\mu \geq 0$ is a damping factor.

The algorithm is auto-adaptive and the value of μ is computed at each step with the following rule :

- a large value of μ is used when the estimate is still far from the solution; equation (2) can thus be approximated as

$$\underline{\Delta}_p \approx - (1/\mu)\, J^T \underline{\varepsilon}$$

  that is a small step in the direction of the maximum slope of the function (this is the same as the *steepest descent method*). By this assumption the convergence to a local minimum of E($\underline{p}$), although very slow, is indeed guaranteed.
- a small value of μ is used when the solution is close; equation (2) can thus be approximated as

$$(J^T J)\, \underline{\Delta}_p \approx - J^T \underline{\varepsilon}$$

that is the solution according to the Gauss-Newton method, much faster but less robust than the previous one.

The iterative process (1) ends when :
- the residual error $E(\underline{p}^*)$ is less than an arbitrary small threshold
- the Euclidean norm of vector $\underline{\Delta}_p$ is less than an arbitrary small threshold
- the maximum allowed number of iterations has been reached.

In the following examples the linear system (2) has been solved with an SVD (*Singular Value Decomposition*) method and the Jacobian matrix has been approximated in each iteration with the aid of finite differences.

It is worthy to note that the algorithm will usually converge to the local minimum that is closest to the initial estimate $\underline{p}_0$ and this is not necessarily the required absolute minimum of the function $E(\underline{p})$.

This means that only by starting from "good enough" values of the unknown parameters can there be a good confidence of finding the required result.

In order to avoid this dependence on the initial estimate, following results have been obtained by means of several different runs of the algorithm, each of them starting from different values of φ, λ, and d computed as

$$\varphi = -90 + j * 15 \qquad j = 0 \div 12$$
$$\lambda = -180 + k * 60 \qquad k = 0 \div 6$$
$$d = -180 + l * 15 \qquad l = 0 \div 24$$

in such a way that all the definition domain of the three variables is scanned[4]. Then the solution with the minimum resulting mean squared value ($E(\underline{p})/N$) of the errors has been extracted from the 13*7*25=2275 results.

<u>*Simulation results*</u>[5]

In order to verify the performance of the above method, let's consider a vertical declining dial showing italic hours (φ = 40° N, d = 30° W, L = 25) where the following elements are available and have been measured:
- case 1: the three hour lines 17, 19 and 21 (each of them being defined by means of two points)
- case 2: the intersection points between the equinox line and the same three hour lines
- case 3: the intersection points between the 17 line and the equinox line and between the 19 line and the solstice lines
- case 4: all previous elements together

Obtained results are shown in the following table :

	true	case 1	case 2	case 3	case 4
φ	40.00°	40.03°	40.01°	40.02°	40.00°
d	30.00°	29.92°	30.05°	30.00°	30.01°
L	25.00	25.19	24.98	25.01	25.00
x	125.00	125.36	125.06	125.02	125.01
y	37.00	37.11	37.01	37.00	36.99

[4] The step used in the analysis of the volume included between the extreme values of φ, λ, and d was determined by experiments.

[5] All the results here exposed have been obtained by means of the freeware software program *Orologi Solari* that is available at the author's web site http://digilander.libero.it/orologi.solari. A copy of the input data used in these examples can be asked to the author at the address sun.dials@libero.it.

Figures 1 to 3 show the graphs that are obtained using the first 3 case results (cross signs show the position of the points used as the input to the problem).

In order to verify how the method behaves in the presence of measurement errors, a random error (between +1 and -1, corresponding to 1/25 of the style length) has been added to the measured coordinates. Obtained results are shown in the following table:

	true	case 1	case 2	case 3	case 4
φ	40.00°	38.57°	35.93°	34.83°	39.96°
d	30.00°	33.31°	25.22°	25.70°	29.60°
L	25.00	17.46	24.36	22.58	25.10
x	125.00	107.93	121.69	119.63	124.89
y	37.00	31.61	34.81	35.59	37.00

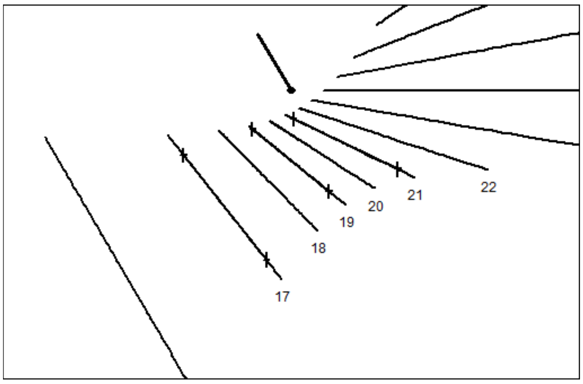

Figure 1

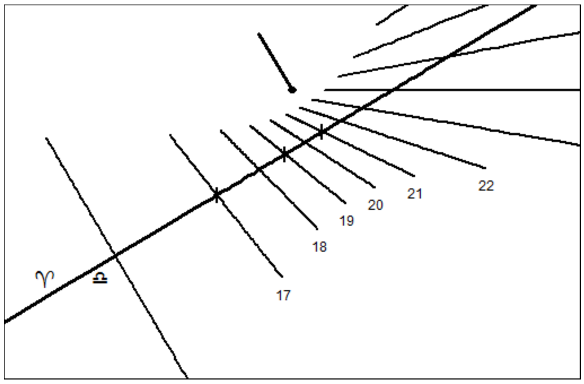

Figure 2

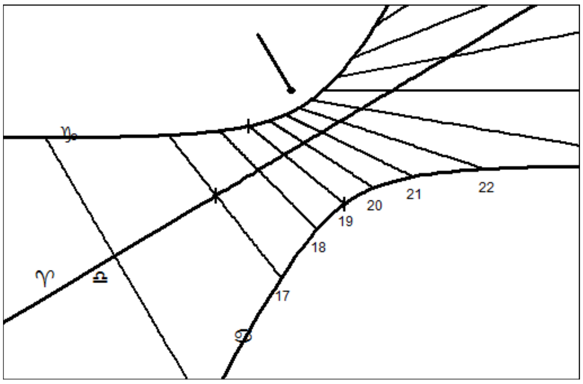

Figure 3

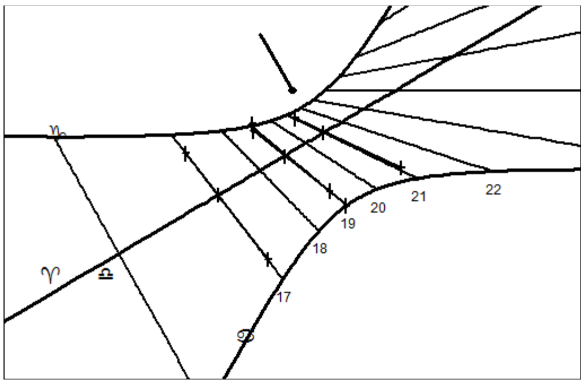

Figure 4

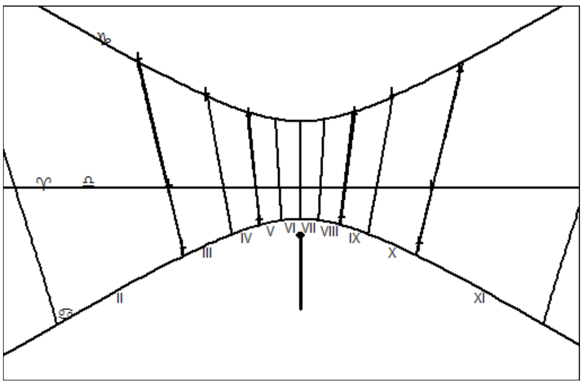

Figure 5

In the first three cases the effect of the measurement error is evident, but in the last one the resulting error is significantly smaller because of the larger amount of information provided and because of the averaging process that is inherent in the method.

Results are thus fairly aligned to the theoretical values (see fig. 4).

In a similar way, consider now a horizontal dial showing temporal hours (φ = 40° N, L = 25) where the following elements can be measured:

- case 1: the four hour lines[6] 2, 4, 8 and 10 (each being defined by two points)

- case 2: the intersection points between the equinox line and the 2 and 10 hour lines and between the winter line and the 3 and 9 hour lines

- case 3: all previous elements together

Obtained results are shown in the following table:

	true	case 1	case 2	case 3
φ	40.00°	38.62°	39.99°	39.60°
L	25.00	25.45	25.01	25.16
x	0.00	0.00	0.00	0.00
y	0.00	-0.41	-0.02	-0.42

If a random error between +1 and -1 is again added to the measurements, the following results are obtained:

	true	case 1	case 2	case 3
φ	40.00°	36.64°	39.79°	39.15°
L	25.00	25.47	25.26	25.67
x	0.00	0.11	-0.18	0.16
y	0.00	6.48	-0.42	-0.29

Figure 5 shows the result that is obtained for case 3. Here again the availability of more input data allows for a better result[7].

As a last example consider a vertical declining dial showing astronomical hours corrected for the longitude error (φ = 45° N, λ = 8° E, d = 60° E, L = 25) where the following elements have been measured:

- case 1: hour line 10:00 (defined by two points) and the intersection points between the 8:00 and 12:30 hour lines with the equinox line
- case 2: the intersection points between the winter line and the 8:30 and 10:30 hour lines and the intersection points between the summer line and the 11:30 hour line
- case 3: all previous elements together

	true	case 1	case 2	case 3
φ	45.00°	45.04°	45.05°	45.02°
λ	-8.00°	-7.99°	-7.93°	-7.98°
d	-60.00°	-60.05°	-59.53°	-60.03°
L	25.00	25.00	25.19	25.01
x	0.00	-0.01	0.21	-0.01
y	0.00	0.03	0.02	0.01

	true	case 1	case 2	case 3
φ	45.00°	44.63°	32.99°	45.19°
λ	-8.00°	-10.10°	-18.59°	-8.73°
d	-60.00°	-57.71°	-107.7°	-59.36°
L	25.00	24.18	0.58	24.72
x	0.00	-0.18	2.33	-0.24
y	0.00	0.36	1.01	0.30

Obtained results are in the table above.

When adding a measurement error, obtained results are as in the table to the left.

[6] Although temporal lines are actually curved lines, they are here approximated by straight lines: the smaller the place latitude is, the more accurate the results are.

[7] It should be noted however that the straight line assumption for temporal hour lines introduces additional errors in the results in cases 1 and 3, both in the absence or presence of measurement errors. The best results are obtained in case 2 where only intersection points are considered.

Figure 6 shows the result that is obtained in case 3 when all the available data are used.

It is possible to find situations where measured elements are not enough to exclusively define the design parameters of the dial and the system of equations admits an infinite number of solutions.

Consider for instance the vertical declining dial in fig. 7 showing babilonic hours (φ = 40° N, d = 30° W, L = 25). The only knowledge of the intersections between the 5, 6 and 7 hour lines with the meridian line does not permit to find the correct values.

The result shown in fig. 8 (φ = 40.02° N, d = 80.89° E, L = 4.58) is just one in an infinite number of possible solutions: all of them are correct from a mathematical point of view but very different from the required result.

The correct (required) solution can only be obtained if more measured elements (points or

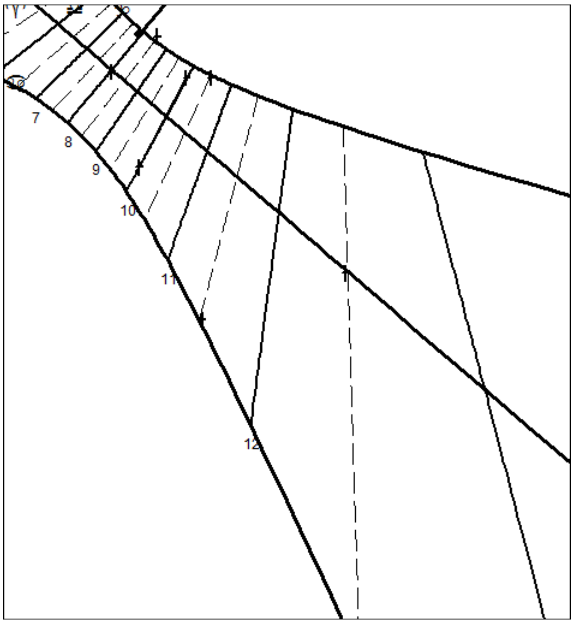

Figure 6

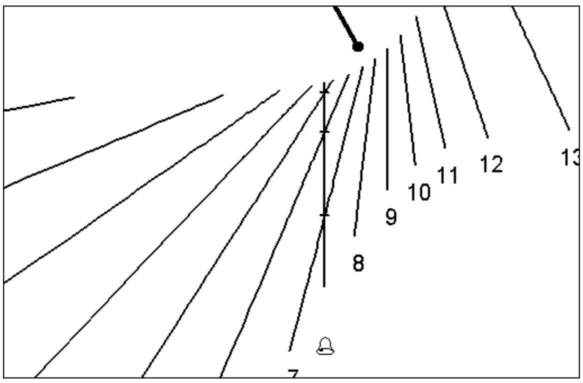

Figure 7

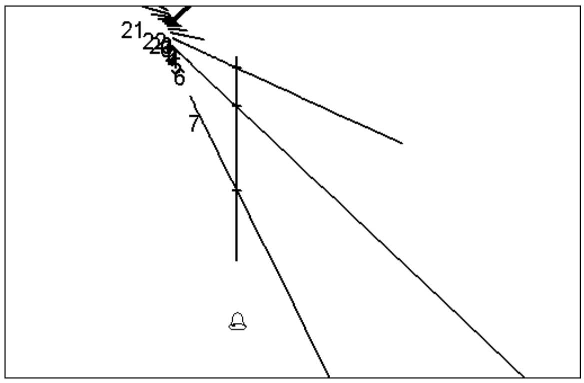

Figure 8

lines) are introduced or if the number of unknown variables is reduced: *e.g.* the knowledge of the position or the length of the style is often sufficient in order to get the desired solution.

As an alternative it can be useful to start from an estimate $\underline{p}_0$ of the unknown vector that is close enough to the solution and then to look for a local minimum of E($\underline{p}$) in the neighborhood of $\underline{p}_0$.

Conclusions

The above method can be useful in order to obtain the original design parameters of a sundial when only a few elements are left.

Simulation cases suggest the following practice:
- as much as possible input elements (points and lines) should be used
- points (intersections between hour lines and day lines, meridian line etc.) instead of lines should be preferred
- it is convenient to reduce the number of unknown variables in the problem when some parameter (latitude, style position *etc.*) is already known

- it can be useful to look for a local solution starting from a “good enough” (apparently close to the required) set of parameters
- every result should be assessed with the help of the resulting graph (including input data) in order to filter out any result that can be correct from the mathematical point of view but is not realistic for that dial.

References

[1] FERRARI G. (1998), *Relazioni e Formule per lo studio delle Meridiane Piane*, self-published, Fifth Chapter, pp. 146-157.
[2] MADSEN K. et al. (2004), *Methods for non-linear least squares problems*, Informatics and Mathematical Modelling, Technical University of Denmark, 2nd Edition.
http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3215/pdf/imm3215.pdf
[3] MANOLIS I. et al. (2005), *A Brief Description of the Levenberg-Marquardt Algorithm Implemented by levmar*, Institute of Computer Science, Heraklion, Crete, GREECE.
http://www.ics.forth.gr/~lourakis/levmar/levmar.pdf

Gianpiero Casalegno - Via M. D'Azeglio 179 – 10081 Castellamonte (TO) – Italy   sun.dials@libero.it

Digital Bonuses

This issue of *The Compendium* includes 4 bonuses.

The first is version 21 of Gian Casalegno's program *Orologi Solari*. This application permits the design of a vertical declined / inclined / horizontal sun dial with french, italic, babilonic, or twilight hours and with day lines, *etc.* It also includes horizontal and vertical analemmatic dials, azimuth ortho/stereographic, monofilar, ascendant lines, ecliptic hours, azimuth and elevation lines and more. It includes the computation of wall declination; it simulates the behavior of the dial; and it includes a screensaver based on the result of your design. Following the subject of Gian's article in this issue, the program can also do some reverse engineering: starting from the measured elements of an existing vertical declining /horizontal dial, you can obtain the original unknown parameters of the dial: latitude, wall declination, style length *etc.*

The program does not need to be installed; it can be run simply by double clicking on the file orologisolari.exe   To install the screensaver, copy *SunDialSaver.scr* and all the .gnm files to the Windows directory, then select the screensaver in Desktop Properties and set your desired .gnm file in screensaver setup.

The second and third bonuses are the PowerPoint files that accompany Tony Moss' article *Angling for Precision* and Don Petrie's article *Development Of The Christ Church Sundial*.

The fourth bonus is the folder DialistsCompanion described in The Tove's Nest in this issue. This folder uses the portable version of *DOSBox* to run *The Dialist's Companion* in any Windows environment, including 64-bit Vista. [Place this folder in the root directory of your drive.]

[For a limited time, these files (with the exception of the Orologi Solari program) are available to all members at http://drop.io/NASSbonus1. Use password: 162. The DialistsCompanion folder is stored as a single zip file. Members may also always find the current version of Orologi Solari at Gian Casalegno's site: http://digilander.libero.it/orologi.solari/download/download_enu.html.]