manuscripta geodaetica (1993) 18:46–52

manuscripta geodaetica

Fast algorithm for the formation of normal equations in a least-squares spherical harmonic analysis by FFT

Cheinway Hwang

Department of Geodetic Science and Surveying, The Ohio State University, 1958 Neil Ave., Columbus, Ohio 43210, USA
Present address: Department of Earth Sciences, Oxford University, Oxford OX1 3PR, England

Accepted November 17, 1992

ABSTRACT. An FFT method has been developed to efficiently form the normal equations in a least-squares spherical harmonic analysis given an incomplete set of data on a sphere or an incomplete set of surface gravity anomalies in equiangular blocks. In the method, the elements of normal equations are analytically expressed in a form suitable for FFT processes. Based on the method, two programs have been designed separately to handle the case where data are given on a sphere and the case where the surface gravity anomalies are given. The results show that the FFT method achieves a remarkably high speed in forming the elements as compared to the straightforward method used in the past. In particular, the FFT method performs 163 times faster than the conventional method when the maximum degree of expansion is 70. This method has made possible efficient selection of data weighting strategies in the least-squares adjustment.

Introduction

When an incomplete set of data is given on a sphere, a least-squares approach can be used to perform spherical harmonic analysis, such as the sea surface topography analysis carried out by Engelis (1987). In a combination solution of geopotential coefficients for which surface gravity anomalies, satellite tracking data and altimeter data are used, a least-squares approach again is employed, such as Rapp (1989). In these techniques, an important element is the formation of normal equations, which requires considerable computer time when an expansion of relativley high degree is desired (Rapp, 1989, p. 278; Pavlis, 1988, p. 89). An algorithm for analysing unevenly sampled point data on a sphere has been given by Colombo (1981, pp. 57-61). In this study, a method which reformulates the structure of the normal equations and subsequently utilizes an FFT technique will be developed to reduce substantially the computer time expended in forming the normal equations. This paper emphasizes the numerical aspects of the spherical harmonic analysis, and only a brief introduction of the origin of normal equations is given. Two cases will be studied separately, although the techniques for both are almost the same. The method is largely based on Hwang (1991, Section 6.2.2). Colombo's (1981) similar method will be compared with the one developed in this paper.

Harmonic Analysis Given Data on a Sphere

In this case, a function which must be at least square integrable is expanded into surface spherical harmonics in the form

$$f(\theta, \lambda) = \sum_{n=0}^{\infty} \sum_{m=0}^{n} (\bar{a}_{nm} \cos m\lambda + \bar{b}_{nm} \sin m\lambda) \bar{P}_{nm}(\theta) \tag{1}$$

where θ is colatitude, λ is longitude and $\bar{P}_{nm}(\theta)$ is the fully normalized associated Legendre function (Heiskanen and Moritz, 1967). For this study, the problem is defined as follows: given mean values of $f(\theta, \lambda)$, $\{\bar{f}_{k\ell}\}$ on equiangular blocks which may not cover the entire sphere, find the expansion coefficients $\bar{a}_{nm}$ and $\bar{b}_{nm}$. For such a problem, we obtain from (1)

$$\begin{aligned} \bar{f}_{k\ell} + e_{k\ell} &= \frac{1}{\Delta\sigma_k} \int_{\theta_k}^{\theta_{k+1}} \int_{\lambda_\ell}^{\lambda_{\ell+1}} f(\theta, \lambda) d\sigma \\ &= \frac{1}{\Delta\sigma_k} \sum_{n=0}^{N\max} \sum_{m=0}^{n} (\bar{a}_{nm} IC_m^\ell + \bar{b}_{nm} IS_m^\ell) \overline{IP}_{nm}^k \end{aligned} \tag{2}$$

where k is the index for colatitude, ℓ is the index for longitude, $\bar{f}_{k\ell}$ is the mean value at k, ℓ, $e_{k\ell}$ is the approximation error due to limited expansion degree Nmax, $\Delta\sigma_k$ is the area of the (k, ℓ) equiangular block which depends on colatitude only, and

$$\left\{ \begin{array}{c} IC_m^\ell \\ IS_m^\ell \end{array} \right\} = \int_{\lambda_\ell}^{\lambda_{\ell+1}} \left\{ \begin{array}{c} \cos m\lambda \\ \sin m\lambda \end{array} \right\} d\lambda \tag{3}$$

$$\overline{IP}_{nm}^k = \int_{\theta_k}^{\theta_{k+1}} \overline{P}_{nm}(t)dt \quad , \quad t = \cos\theta \tag{4}$$

To find the coefficients, we minimize the square of the error norm $\|e\|^2$ as follows:

$$\varphi = \|e\|^2 = \sum_{k=0}^{N-1} \sum_{\ell=0}^{2N-1} w_{k\ell} e_{k\ell}^2 = \text{a minimum} \tag{5}$$

where $N = \pi / (\theta_{k+1} - \theta_k)$, and $w_{k\ell}$ is an index function defined as

$$w_{k\ell} = \left\{ \begin{array}{ll} 1 \;, & \text{data exists at block } k, \ell \\ 0 \;, & \text{data does not exist at block } k, \ell \end{array} \right\}$$

The condition (5) leads to a typical "least-squares adjustment" problem for which we may write (2) as

$$L + V_X = AX \tag{6}$$

where L is a vector containing the "observations" $\bar{f}_{k\ell}$, V_X is the error vector containing $e_{k\ell}$, X is the vector of unknowns containing $\bar{a}_{nm}$ and $\bar{b}_{nm}$, and A is the design matrix. The solution X fulfilling (5) is

$$\begin{aligned} X &= (A^TA)^{-1}(A^TL) \\ &= N_X^{-1}U_X \end{aligned} \tag{7}$$

where $N_X = A^TA$ is the normal matrix and $U_X = A^TL$. The process of constructing the matrix N_X and the vector U_X is designated as the formation of the normal equations, which is the subject of this study. Based on (2), the elements of N_X, denoted as K_{nmrs}^j, can be classified into four types:

$$\left\{ \begin{array}{c} K_{nmrs}^1 \\ K_{nmrs}^2 \\ K_{nmrs}^3 \\ K_{nmrs}^4 \end{array} \right\} = \sum_{k=0}^{N-1} \frac{1}{\Delta\sigma_k^2} \overline{IP}_{nm}^k \overline{IP}_{rs}^k \sum_{\ell=0}^{2N-1} w_{k\ell} \left\{ \begin{array}{cc} IC_m^\ell & IC_s^\ell \\ IS_m^\ell & IS_s^\ell \\ IC_m^\ell & IS_s^\ell \\ IS_m^\ell & IC_s^\ell \end{array} \right\} \tag{8}$$

where n, r correspond to the degree and m, s correspond to the order of the spherical harmonics. Note that not every element in (8) will enter N_X. Similarly, the elements T_{nm}^j in vector U_X are

$$\left\{ \begin{array}{c} T_{nm}^1 \\ T_{nm}^2 \end{array} \right\} = \sum_{k=0}^{N-1} \frac{1}{\Delta\sigma_k} \overline{IP}_{nm}^k \sum_{\ell=0}^{2N-1} w_{k\ell} \bar{f}_{k\ell} \left\{ \begin{array}{c} IC_m^\ell \\ IS_m^\ell \end{array} \right\} \tag{9}$$

Again, we know that elements T_{n0}^2 will not enter U_X.

To evaluate the elements in (8) and (9), one can just accumulate the contributions from each latitude belt k in a straightforward manner, and take advantage of the property of associated Legendre function $\overline{P}_{nm}(-t) = (-1)^{n+m}\overline{P}_{nm}(t)$. Nevertheless, due to the use of the index function $w_{k\ell}$ which enables the regular forms of the elements in (8) and (9) to be formulated, a more efficient method can be developed. We shall exploit these forms using the FFT for the computational efficiency.

Let IE_m^ℓ be defined as

$$\begin{aligned} IE_m^\ell &= \int_{\lambda_\ell}^{\lambda_{\ell+1}} e^{im\lambda} d\lambda = h(m)e^{im\ell\Delta\lambda} \\ &= IC_m^\ell + i\, IS_m^\ell \end{aligned} \tag{10}$$

where

$$h(m) = \left\{ \begin{array}{ll} \Delta\lambda = \lambda_{\ell+1} - \lambda_\ell \;, & m = 0 \\ \dfrac{i(1 - e^{im\Delta\lambda})}{m} \;, \; i = \sqrt{-1} \;, & m \neq 0 \end{array} \right. \tag{11}$$

Furthermore, we define

$$\begin{aligned} \alpha &= IC_m^\ell\, IC_s^\ell \\ \beta &= IS_m^\ell\, IS_s^\ell \\ \gamma &= IC_m^\ell\, IS_s^\ell \\ \delta &= IS_m^\ell\, IC_s^\ell \end{aligned} \tag{12}$$

and

$$\begin{aligned} a &= IE_m^\ell\, IE_{-s}^\ell = h(m)h(-s)e^{i(m-s)\ell\Delta\lambda} \\ b &= IE_m^\ell\, IE_s^\ell = h(m)h(s)e^{i(m+s)\ell\Delta\lambda} \\ c &= IE_{-m}^\ell\, IE_{-s}^\ell = h(-m)h(-s)e^{-i(m+s)\ell\Delta\lambda} \\ d &= IE_{-m}^\ell\, IE_s^\ell = h(-m)h(s)e^{-i(m-s)\ell\Delta\lambda} \end{aligned} \tag{13}$$

If we expand a, b, c and d into real parts and imaginary parts using the relationship given in (10), we get

$$a = \alpha + \beta - i(\gamma - \delta)$$
$$b = \alpha - \beta + i(\gamma + \delta)$$
$$c = \alpha - \beta - i(\gamma + \delta) \qquad (14)$$
$$d = \alpha + \beta + i(\gamma - \delta)$$

Thus

$$\alpha = \frac{1}{4}\,(a + b + c + d)$$
$$\beta = \frac{1}{4}\,(a - b - c + d)$$
$$\gamma = \frac{1}{4i}\,(-a + b - c + d) \qquad (15)$$
$$\delta = \frac{1}{4i}\,(a + b - c - d)$$

which builds the connection between the products of two trigonometric functions and complex exponential functions (as used in an FFT process). From (14), we have

$$a = d^* \;,\; b = c^* \qquad (16)$$

or

$$a + b = (c + d)^* \;,\; a - b = (d - c)^* \qquad (17)$$

where * is the conjugate operator. Furthermore, we define the FFT (fast Fourier transform) of $w_{k\ell}$ at frequency (m + s) as

$$F^k(m, s) = \sum_{\ell=0}^{2N-1} w_{k\ell}\, e^{i(m+s)\ell\Delta\lambda} = \text{FFT}\,(w_{k\ell})\,,\ \Delta\lambda = \frac{2\pi}{2N} \qquad (18)$$

where k indicates the k^{th} latitude belt. Moreover, we define

$$U^k(m, s) = h(m)h(-s)F^k(m, -s) + h(m)h(s)F^k(m, s) \qquad (19)$$

$$V^k(m, s) = h(m)h(-s)F^k(m, -s) - h(m)h(s)F^k(m, s) \qquad (20)$$

Using (12) to (20), it is not difficult to see that

$$\begin{Bmatrix} K^1_{nmrs} \\ K^2_{nmrs} \\ K^3_{nmrs} \\ K^4_{nmrs} \end{Bmatrix} = \frac{1}{4}\sum_{k=0}^{N-1} \frac{1}{\Delta\sigma_k^2}\, \overline{IP}^k_{nm}\, \overline{IP}^k_{rs} \begin{Bmatrix} \left[U^k(m, s) + \left(U^k(m, s)\right)^*\right] \\ \left[V^k(m, s) + \left(V^k(m, s)\right)^*\right] \\ \left[-V^k(m, s) + \left(V^k(m, s)\right)^*\right]/i \\ \left[U^k(m, s) - \left(U^k(m, s)\right)^*\right]/i \end{Bmatrix}$$

$$= \frac{1}{2}\sum_{k=0}^{N-1} \frac{1}{\Delta\sigma_k^2}\, \overline{IP}^k_{nm}\, \overline{IP}^k_{rs} \begin{Bmatrix} \mathrm{Re}\left(U^k(m, s)\right) \\ \mathrm{Re}\left(V^k(m, s)\right) \\ \mathrm{Im}\left(-V^k(m, s)\right) \\ \mathrm{Im}\left(U^k(m, s)\right) \end{Bmatrix} \qquad (21)$$

where Re(·) and Im(·) are the real part and imaginary part of a complex number. Equation (21) shows the desired FFT form for evaluating the elements of the normal matrix N_X.

It is much easier to deal with the elements of vector U_X. First of all, let us define

$$F^k_m = \sum_{\ell=0}^{2N-1} w_{k\ell}\bar{f}_{k\ell}\, \mathrm{IE}^{\ell}_m = h(m)\text{FFT}\,(w_{k\ell}\bar{f}_{k\ell}) \qquad (22)$$

Using (9), (10) and (22), we get the desired FFT forms

$$\begin{Bmatrix} T^1_{nm} \\ T^2_{nm} \end{Bmatrix} = \sum_{k=0}^{N-1} \frac{1}{\Delta\sigma_k}\, \overline{IP}^k_{nm} \begin{Bmatrix} \mathrm{Re}\left(F^k_m\right) \\ \mathrm{Im}\left(F^k_m\right) \end{Bmatrix} \qquad (23)$$

The high efficiency of calculating the elements of normal matrix N_X and vector U_X lies in the fact that the major computational burden is relieved by the use of FFT in equations (18) and (22). Furthermore, since

$$\overline{P}_{nm}(-t) = (-1)^{n+m}\overline{P}_{nm}(t)$$
$$\overline{P}_{nm}(-t)\,\overline{P}_{rs}(-t) = (-1)^{n+m+r+s}\overline{P}_{nm}(t)\overline{P}_{rs}(t) \qquad (24)$$

the integrations of associated Legendre functions are only needed in one hemisphere and we can process data from two latitude belts at a time. If the data are real-valued (which is the case in this study), the algorithm for computing the FFT of two real functions simultaneously (Brigham, 1988, p. 188) may be used. For that algorithm, we are given two real-valued data arrays $h(\ell)$, $g(\ell)$, $\ell = 0, ..., 2N - 1$, from north and south latitudes. We then form the complex array $y(\ell)$ with $h(\ell)$ and $g(\ell)$ being its real and imaginary parts:

$$y(\ell) = h(\ell) + i\,g(\ell),\ \ \ell = 0, ..., 2N - 1 \qquad (25)$$

Now we can perform the Fourier transform for array $y(\ell)$ to get $Y(\ell)$, $\ell = 0, ..., 2N - 1$. Finally we obtain the Fourier transforms of $h(\ell)$ and $g(\ell)$, denoted as $H(\ell)$ and $G(\ell)$, from the relationships (see also ibid., p. 188)

$$H(\ell) = \frac{1}{2}\mathrm{Re}\,(Y(\ell) + Y\,(2N - \ell)) + \frac{1}{2}\,i\,\mathrm{Im}\,(Y(\ell) + Y\,(2N - \ell))$$
$$G(\ell) = \frac{1}{2}\mathrm{Re}\,(Y(\ell) + Y\,(2N - \ell)) - \frac{1}{2}\,i\,\mathrm{Im}\,(Y(\ell) + Y\,(2N - \ell)) \qquad (26)$$

for $\ell = 1, ..., 2N - 1$, and H(o) = Re (Y(o)), G(o) = Im (Y(o)).

Summarizing the above development, we present the algorithm for calculating the elements of matrix N_X and vector U_X as follows:

1. Compute h(m) in (11) and then the products h(m)h(s) for m = 0, ..., Nmax and s = - Nmax, ..., Nmax.
2. Compute the FFT of $w_{k\ell}$ and $(w_{k\ell}\,\bar{f}_{k\ell})$, $\ell = 0, ..., 2N - 1$, at the northern-most latitude belt and the southern-

most latitude belt using the algorithm shown in (25) and (26).

3. Compute $\overline{IP}_{nm}$ at the northern latitude belt.
4. Accumulate K^j_{nmrs}, j = 1, 2, 3, 4 and T^j_{nmrs}, j = 1, 2, up to n = m = v = s = Nmax at these two latitude belts. The values $U^k(m, s)$ and $V^k(m, s)$ in (19) and (20), and in (22) are found by identifying the products of h(m)h(s) (or h(m)) and the FFT components from step 2 through the orders m and s, or m alone.
5. Repeat 2 to 4 at the next northern and southern latitude belts, until all latitude belts are exhausted.
6. Identify the desired elements from the final K^j_{nmrs} and T^j_{nm} values.

As mentioned in the introduction, Colombo (1981) also developed an algorithm for the type of analysis discussed in this section. However, his algorithm is for a point data analysis. One can compare (2.91) in Colombo (ibid.) and (8) in this paper for the structure of the elements of the normal equations (Note: in this paper, mean values are assumed). The major difference between the two algorithms appears in the ways of treating the FFT processes. Compare (2.92) - (2.95) in Colombo (ibid.) with (18) - (21) in this paper. Both algorithms use the properties of Legendre's function (24). As far as the programming is concerned, the use of complex algebra in this paper has yielded a concise code in FORTRAN. Furthermore, due to the fact that no extensive logical judgements are required and due to the aforementioned computational procedure, vectorization of the program are easy to carry out in a supercomputer. These advantages in programming may not be found in Colombo's (ibid.) algorithm.

Geopotential Coefficients from Surface Gravity Anomalies

In this case, the geopotential coefficients are to be found from a set of gravity anomalies by a least-squares method. The formation of normal equations can also be made extremely efficient using the FFT approach illustrated above. Following Pavlis' (1988, eq. (4.11)) rigorous formulation, we have

$$\overline{\Delta g}_{k\ell} = \frac{1}{\Delta\sigma_k}\frac{GM}{\bar{r}^2_{k\ell}}\sum_{\substack{n=0\\n\neq 1}}^{Nmax}(n-1)\left[\frac{a}{\bar{r}_{k\ell}}\right]^n \cdot \sum_{m=0}^{n}\left(\overline{C}_{nm}\, IC^{\ell}_m + \overline{S}_{nm} IS^{\ell}_m\right)\overline{IP}^k_{nm} \tag{27}$$

where the precise definitions of $\overline{\Delta g}_{k\ell}$, $\bar{r}_{k\ell}$ and GM can be found in Pavlis (ibid., Chapter 2). In (27), the data $\overline{\Delta g}_{k\ell}$ are the mean gravity anomalies and the "unknowns" are the disturbing geopotential coefficients $\overline{C}_{nm}$, $\overline{S}_{nm}$. Assuming that the signal contribution above degree Nmax has been removed, $\overline{\Delta g}_{k\ell}$ may contain noise only. In such a case, the minimum variance solution will lead to the same form as in (7) where vector X contains the coefficients $\overline{C}_{nm}$, $\overline{S}_{nm}$. In this case the elements of the normal matrix are (cf. Pavlis, 1988, pp. 73-74).

$$\begin{pmatrix}\overline{K}^1_{nmrs}\\ \overline{K}^2_{nmrs}\\ \overline{K}^3_{nmrs}\\ \overline{K}^4_{nmrs}\end{pmatrix} = GM^2(n-1)(r-1)\sum_{k=0}^{N-1}\frac{1}{\Delta\sigma^2_k}\overline{IP}^k_{nm}\cdot\overline{IP}^k_{rs}\sum_{\ell=0}^{2N-1}\frac{1}{\bar{r}^4_{k\ell}}\left[\frac{a}{\bar{r}_{k\ell}}\right]^{n+r}P_{k\ell}\begin{pmatrix}IC^{\ell}_m\, IC^{\ell}_s\\ IS^{\ell}_m\, IS^{\ell}_s\\ IC^{\ell}_m\, IS^{\ell}_s\\ IS^{\ell}_m\, IC^{\ell}_s\end{pmatrix} \tag{28}$$

where the weight function $P_{k\ell}$ is different from the index function $w_{k\ell}$ in that

$$P_{k\ell} = \begin{cases}\dfrac{1}{\sigma^2_{k\ell}}, & \text{mean gravity anomaly exists at block } k\ell\\ 0, & \text{mean gravity anomaly does not exist at block } k\ell\end{cases}$$

where $\sigma_{k\ell}$ is the standard deviation of the mean gravity anomaly $\overline{\Delta g}_{k\ell}$. The elements of vector U_X in this case are (cf. ibid., pp. 73-74)

$$\begin{pmatrix}\overline{T}^1_{nm}\\ \overline{T}^2_{nm}\end{pmatrix} = GM(n-1)\sum_{k=0}^{N-1}\frac{1}{\Delta\sigma_k}\overline{IP}^k_{nm}\sum_{\ell=0}^{2N-1}\frac{1}{\bar{r}^2_{k\ell}}\left[\frac{a}{\bar{r}_{k\ell}}\right]^n\cdot P_{k\ell}\,\overline{\Delta g}_{k\ell}\begin{Bmatrix}IC^{\ell}_m\\ IS^{\ell}_m\end{Bmatrix} \tag{29}$$

As compared to the computer time for calculating the needed elements in the previous case, more computer time is expected in the current case due to the factors

$$\left[\frac{a}{\bar{r}_{k\ell}}\right]^{n+r}, \quad \left[\frac{a}{\bar{r}_{k\ell}}\right]^n$$

that involve degrees n, r at each latitude belt. Unlike the forms in (8) and (9) where only one FFT process is needed at one latitude belt for all degrees and orders, the forms in (28) and (29) also require the FFT processes for degree n and r at each latitude belt. To eliminate the degree dependence of the FFT process, one can reduce the mean gravity anomalies to a sphere with radius = a using downward continuation and ellipsoidal correction (see Rapp, 1986). By doing such a reduction, we can achieve exactly the same efficiency as we have in (21) and (23).

If we insist on the rigorous forms in (28) and (29) and consider the degree dependence of the FFT process, then the desired formulae for an FFT algorithm are

$$\begin{Bmatrix} \bar{K}^1_{nmrs} \\ \bar{K}^2_{nmrs} \\ \bar{K}^3_{nmrs} \\ \bar{K}^4_{nmrs} \end{Bmatrix} = \frac{1}{2} GM^2(n-1)(r-1) \sum_{k=0}^{N-1} \frac{1}{\Delta\sigma_k^2} \overline{IP}^k_{nm} \overline{IP}^k_{rs} \cdot \begin{Bmatrix} Re\left(\bar{U}^k(n+r, m, s)\right) \\ Re\left(\bar{V}^k(n+r, m, s)\right) \\ Im\left(-\bar{V}^k(n+r, m, s)\right) \\ Im\left(\bar{U}^k(n+r, m, s)\right) \end{Bmatrix} \tag{30}$$

and

$$\begin{Bmatrix} \bar{T}^1_{nm} \\ \bar{T}^2_{nm} \end{Bmatrix} = GM(n-1) \sum_{k=0}^{N-1} \frac{1}{\Delta\sigma_k} \overline{IP}^k_{nm} \begin{Bmatrix} Re\left(\bar{E}^k(n, m)\right) \\ Im\left(\bar{E}^k(n, m)\right) \end{Bmatrix} \tag{31}$$

where

$$\bar{U}^k(n+r, m, s) = h(m)h(-s)\bar{F}^k(n+r, m, -s) + h(m)h(s)\bar{F}^k(n+r, m, s) \tag{32}$$

$$\bar{V}^k(n+r, m, s) = h(m)h(-s)\bar{F}^k(n+r, m, -s) - h(m)h(s)\bar{F}^k(n+r, m, s) \tag{33}$$

and

$$\bar{F}^k(n+r, m, s) = \sum_{\ell=0}^{2N-1} \frac{1}{r^4_{k\ell}} \left[\frac{a}{r_{k\ell}}\right]^{n+r} P_{k\ell}\, e^{i(m+s)\ell\Delta\lambda} \tag{34}$$

$$\bar{E}^k(n, m) = h(m) \sum_{\ell=0}^{2N-1} \frac{1}{r^2_{k\ell}} \left[\frac{a}{r_{k\ell}}\right]^{n} P_{k\ell}\, \overline{\Delta g}_{k\ell}\, e^{im\ell\Delta\lambda} \tag{35}$$

It is clear that the FFT process takes place in (34) and (35). These formulae can be obtained by using exactly the same derivations for elements K^j_{nmrs} and T^j_{nm} in the previous case. The algorithm for computing the elements K^j_{nmrs} and T^j_{nm} is also exactly the same as that given in the previous case.

Results and Discussions

Two programs, called FFTSOL and FFTSOLA, have been developed separately to compute the elements of normal matrices and "U" vectors and solve for the coefficients by the FFT methods discussed above. Program FFTSOL is for the first case where data are given on a sphere; program FFTSOLA is for the second case where mean gravity anomalies are given on the surface of the earth. The counterparts of FFTSOL and FFTSOLA are ADJSST and ADJUST50 which exist in Professor Rapp's program library (Rapp, private communication) and use the conventional (straightforward) method for calculating the elements needed. Therefore, it is possible to compare the performance of these two algorithms (FFT algorithm vs. conventional algorithm) in terms of computer time. All experiments are carried out on the CRAY Y-MP/864 machine residing at the Ohio Supercomputer Center. Moreover, all the four programs have been vectorized to take advantage of the parallel computing hardware in a supercomputer. The FFT computations are made by the IMSL routine FFTCF.

In Table 1, we list the CPU times on the CRAY Y-MP/864 needed for the formation of normal matrices by programs FFTSOL and ADJSST. The data used are the 1° x 1° mean sea surface topography from Levitus as used in Engelis (1987). The CPU times for inverting the normal matrices, which are the same for both programs, are also listed. Nmax is the maximum degree of expansion. Due to limited computer resources, the computation by ADJSST with Nmax > 36 is not made. The results show that up to Nmax = 36, the maximum relative difference (the ratio between the magnitude of difference and the magnitude of quantity being compared) between the two solutions (X vectors) is around 10^{-14} which is the accuracy (in single-precision mode) of the computer used. However, as we can see from Table 1, substantial savings of computer time in the formation of normals has been achieved by the FFT method. One can also observe that the savings factor (CPU ratio in Table 1) grows with the Nmax used. This is due to the fact that larger Nmax enables a more efficient use of the FFT method.

Table 1. CPU Time Comparison of FFTSOL and ADJSST on a CRAY Y-MP/864

Normals (seconds)			CPU	Inversions†
Nmax	FFTSOL	ADJSST	ratios	(seconds)
10	0.33	7.02	21	0.04
15	0.71	41.96	60	0.24
24	3.09	186.12	81	2.16
36	7.42	890.41	120	17.29
50	21.86	~3272	~151	103.74
70	69.54	~12986	~186	~750

~ estimated
† same for both programs, Linpack's routines SPPCO and SPPDI (Dongarra et al., 1979) are used.

When Nmax is equal to the Nyquist frequency (2N/2), the FFT method achieves the maximum efficiency, since all the FFT components can be used in such a case. Roughly after Nmax = 30, the CPU time needed for inversion exceeds the CPU time needed for the formulation of the normal equations in the FFT method. In the conventional method, however, the latter is always larger than the former.

Next we shall compare the performances between FFTSOLA and ADJUST50. A simulated data set of mean gravity anomalies is generated from the OSU89B geopotential model (Rapp and Pavlis, 1990) to Nmax = 50 for 1° x 1° blocks (the ratio $a/r_{k\ell}$ is considered) on a global basis. An incomplete data set is derived from this global data set using the empty blocks in the Levituss SST data set (see above). This data set served as input file for both programs. Again due to limited computer resources, the computations using ADJUST50 are made only up to Nmax = 24. However, the runs with FFTSOLA were made to degrees as high as Nmax = 70. The results show that the maximum relative difference between the two solutions (the geopotential coefficients) reaches 10^{-12} when Nmax = 24. This relative difference is larger than that created by FFTSOL and ADJSST and should be due to different ways of handling the exponents of $\left(a/r_{k\ell}\right)^{n+r}$ in the FFT method and the conventional method. In Table 2, we list the CPU times on the CRAY Y-MP/864 for forming the normal equations from the mean gravity anomalies by the FFT method and the conventional method. Again, from Table 2 we can see that substantial reduction of computer times was achieved by the FFT method.

Table 2. CPU Time Comparison of FFTSOLA and ADJUST50 for Forming the Normal Equations on a CRAY Y-MP/864 in Seconds

Nmax	FFTSOL	ADJSST	CPU ratios
10	2.93	10.65	4
15	4.92	44.28	9
24	9.30	213.90	23
36	19.69	1227+	62
50	42.56	4839+	114
70	116.90	18246+	156

+ From N.K. Pavlis (private communication, 1992)

To do an expansion to Nmax = 100 from the surface gravity anomalies, it is estimated that 1040 seconds are needed for forming the normal equations, while 6400 seconds are needed for inverting the normal matrix (all CPU times are on the CRAY Y-MP/864). The memory required will be about 66 Megawords (for FFTSOLA). It is therefore clear that the major burdens in high degree expansion using "least-squares adjustment" method are represented by the inversion and the storage space of a computer, even if an FFT approach as illustrated above is used. To overcome this difficulty, one may consider the approximations as those suggested by Rapp (1989, p. 276).

Conclusion

The FFT method has shown its ability to dramatically reduce the computer time in forming the normal equations in the least-squares adjustment problems studied in this paper. If the data are given at regularly gridded points (not necessarily complete) instead of equiangular blocks, similar formulae for the formation of normal equations by FFT can also be developed (see Colombo, 1981). The computer time needed in such a method is independent of the number of empty blocks, due to the use of the index function $w_{k\ell}$ and the weight function $P_{k\ell}$ and the way the elements of the normal equations are computed. With this method, it is possible to test different data weighting strategies for the solutions experimentally without requiring inordinate amounts of computer time.

Acknowledgements

I am grateful to Professor R.H. Rapp for his continuous support and constructive suggestions. This study was supported through NASA's TOPEX Altimeter Research in Ocean Circulation Mission and funded through the Jet Propulsion Laboratory under contract 958121 with The Ohio State University Research Foundation. Extensive computer resources were provided by the Academic Computing Services of The Ohio State University and the Ohio Supercomputer Center, grant pas 160. The careful reviews from Dr. Jekeli and two anonymous reviewers have clarified some important issues and improved the quality of the paper.

References

Brigham EO (1988), The Fast Fourier Transform and its Applications, Prentice Hall, Englewood Cliffs, New Jersey.

Dongarra J.J., E.B. Moler, J.R. Bunch and G.W. Stewart (1979), Linpack User's Guide, The Society for Industrial and Applied Math. (SIAM), Philadelphia.

Colombo, O.L. (1981), Numerical Methods for Harmonic Analysis on the Sphere, Dept. of Geodetic Science and Surveying, Report No. 310, The Ohio State University, Columbus.

Engelis, T. (1987), Spherical Harmonic Expansion of the Levitus Sea Surface Topography, Dept. of Geodetic Science and Surveying, Report No. 385, The Ohio State University, Columbus.

Heiskanen, W. A., H. Moritz (1967), Physical Geodesy, W. H. Freeman, New York.

Hwang, C. (1991), Orthogonal Functions Over the Oceans and Applications to the Determination of Orbit Error, Geoid and Sea Surface Topography from Satellite Altimetry, Dept. of Geodetic Science and Surveying, Report No. 414, The Ohio State University, Columbus.

Pavlis, N.K. (1988), Modeling and Estimation of a Low Degree Geopotential Model from Terrestrial Gravity Data, Dept. of Geodetic Science and Surveying, Report No. 386, The Ohio State University, Columbus.

Rapp, R.H. (1986), Global Geopotential Solutions In: H. Sünkel (ed), Lecture Notes in Earth Sciences, Vol. 7, Springer-Verlag, New York, pp. 365-415.

Rapp, R.H. (1989), Combination of Satellite, Altimeter and Terrestrial Gravity Data In: F. Sansó and R. Rummel (eds), Lecture Notes in Earth Sciences, Vol. 25, Springer-Verlag, New York, pp. 261-284.

Rapp, R.H., N.K. Pavlis (1990), The Development and Analysis of Geopotential Coefficient Models to Spherical Harmonic Degree 360, J. Geophys. Res., Vol. 95, No. B3, pp. 21,889-21,911.